PSL UNIVERSITÉ PARIS

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PSL

Préparée au DMA, École Normale Supérieure

Homogénéisation quantitative sur l'amas de percolation et le système de particules

Soutenue par

Chenlin Gu

Le 1 Avril 2021

École doctorale n°386

École Doctorale de Sciences Mathématiques de Paris-Centre

Spécialité

Mathématiques

Composition du jury :

Mme Oriane BLONDEL Université Lyon 1 / CNRS	*Membre du jury*
M. Raphaël CERF ENS Paris	*Membre du jury*
M. Jian DING University of Pennsylvania	*Rapporteur*
Mme Isabelle GALLAGHER ENS Paris / Université de Paris	*Présidente du jury*
M. Yu GU Carnegie Mellon University	*Membre du jury*
M. Milton JARA IMPA	*Rapporteur*
M. Takashi KUMAGAI Kyoto University	*Membre du jury*
M. Jean-Christophe MOURRAT New York University / CNRS	*Directeur de thèse*

ENS ÉCOLE NORMALE SUPÉRIEURE

PSL RESEARCH UNIVERSITY PARIS

Acknowledgments

First and foremost I am extremely grateful to my supervisor Jean-Christophe Mourrat for his invaluable advice, continuous support, and patience during my Ph.D. study. He guides me into the world of "homogenization", shares with me his knowledge and ideas, and always encourages me to explore more in both academic research and daily life. Working with him is a wonderful and unforgettable experience.

Besides my supervisor, I would like to thank the rest of my thesis committee: Oriane Blondel, Raphaël Cerf, Jian Ding, Isabelle Gallagher, Yu Gu, Milton Jara, Takashi Kumagai, for their insightful comments and discussions, which inspire me a lot and also point out some perspectives in my future research.

In these years, I'm happy to have the opportunity to work with Vincent Bansaye, Paul Dario, Arianna Giunti, Wei Jiang, Maximilian Nitzschner, Alvin E. Roth, Qingyun Wu, Linglong Yuan, Tianyuan Zhao, Ban Zheng. I learned a lot from these collaborators and hope to continue our collaborations in the future. Besides, I would like to express my sincere gratitude to Scott Armstrong, Wei Wu, and Jiangang Ying for their helpful suggestions during my thesis.

Je pense que c'est la plus grande aventure dans ma vie de venir en France pour étudier les mathématiques. Je tiens à remercier Yijun Yao et Hongquan Li de l'Université Fudan, qui m'ont proposé cette idée car ils font également leurs thèses en France. J'adresse aussi remerciement à la Fondation de l'École Polytechnique et à la Fondation Mathématiques Jaques Hadamard pour leur financement pendant ma thèse et pendant mes études.

J'ai réalisé ma thèse dans deux laboratoires : DMA/ENS et l'Institut de Courant. J'aimerais remercier ces deux equipes pour m'avoir accueilli et donné une condition excellente de travail. J'aimerais aussi remercier Bénédicte Auffray et Zaïna Elmir pour leur disponibilité et leur aide dans les affaires administratives. Merci à mes collègues : Barbara, Cyril, Daniel, Louise, Hong-bin, Maxime, Mickaël, Michel, Nicolas, Paul, Pierre, Rémy, Shahriar, Shuangjia, Shuo, Thomas, Tobias, Tunan, Wei, Yijun, Yusuke, Yuzhe, et j'en profite beaucoup dans nos échanges et discussions.

Avant commencer ma thèse, j'ai eu la chance de suivre des formations de grande qualité en mathématiques à l'École Polytechnique et à l'Université Paris-Sud : Grégory Allaire, François Alouges, Vincent Bansaye, Thierry Bodineau, Jean-Philippe Bouchaud, Hugo-Duminil Copin, Nicolas Curien, Raphaël Danchin, Nathanaël Enriquez, Charles Favre, Emmanuel Gobet, Jean-François Le Gall, Pascal Massart, Sylvie Méléard, Pierre-Loïc Méliot, Pierre Raphaël, Gilles Stoltz, Nizar Touzi, pour ne pas tous les citer. Igor Kortchemski, Pascal Maillard, Pierre Henry-Labordere ont encadré mon stage de mémoire pendant cette période. J'ai appris non seulement les connaissances, mais aussi leurs passions pour les mathématiques dans ces cours. C'est grâce à cette expérience que je peux bien me former avant de me plonger dans la thèse.

Je remercie mes camarades de classes pour leur soutien et témoignage dans cette aventure : Waner, Honghao, Mingxing, Jingke, Shasha, Enhong, Li, Yiyang, Zhixin, Qi, Quan, Zekun, Wensi, Enzo, Axel, Benoit, Florian, Mathilde, Chloé, Oussama, Tien, Huy.

Je n'oublie pas le séjour chez Christine Dumas, où j'ai passé 4 mois magnifiques avec les cours de français, et les Yam's avec Diego et Louanne. Un merci tout particulier à mon parrain Pascal Correc et sa famille Adrianna, Victoria, Arthur pour les moments inoubliables avec des matchs de foot et la fête de la science.

I would like to thank Wei Jiang - you are the sunshine of my life.

Finally, let me thank my parents Xiaoxiang Gu and Huizhen Chen. They never told me to be like someone else, but just brought me up, support me, and let me define who I am.

Contents

Chapitre 0

Résumé de la thèse

Cette thèse consiste en les travaux de recherche [133, 134, 85, 135, 115] pendant mon doctorat et étudie l'interaction entre l'homogénéisation quantitative et deux modèles stochastiques : *le modèle de percolation surcritique* et *le système de particules en interaction*. Un objet fondamental de la théorie de l'homogénéisation stochastique est de comprendre l'équation

$$-\nabla \cdot (\mathbf{a} \nabla u) = f \qquad \text{dans } B_r, \tag{1}$$

avec $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^{d\times d}_{sym}$ un coefficient symétrique, $\mathbb{Z}^d$-stationnaire, ergodique satisfaisant l'ellipticité uniforme $|\xi|^2 \leqslant \xi \cdot \mathbf{a}\xi \leqslant \Lambda|\xi|^2$, et où B_r est la boule euclidienne de rayon r centrée à l'origine. Pour r très grand, sa solution peut être approximée par *la solution effective* $\bar{u}$ satisfaisant

$$-\nabla \cdot (\bar{\mathbf{a}} \nabla \bar{u}) = f \qquad \text{dans } B_r, \tag{2}$$

avec la même condition au bord. Ici, $\bar{\mathbf{a}}$ est appelé *le coefficient effectif*, qui est une matrice constante. La quantité $\bar{\mathbf{a}}$ caractérise non seulement le comportement asymptotique à grande échelle du problème elliptique, mais saisit également le comportement à grande échelle et à long terme du problème parabolique. C'est le lien entre l'homogénéisation et divers modèles de diffusion en probabilité, où la théorie quantitative de l'homogénéisation fournit des outils pour les estimations. Réciproquement, les deux modèles étudiés dans cette thèse ramènent des nouvelles techniques à la théorie de l'homogénéisation : le modèle de percolation va au-delà du cadre de l'ellipticité uniforme, tandis que le système de particules explore une analyse des EDP en dimension infinie avec l'environnement dynamique.

Malgré ces applications, nous devons garder à l'esprit une motivation importante de l'homogénéisation dès le début : une approximation numérique efficace. En fait, la solution numérique dans l'eq. (1) est très coûteuse à calculer numériquement pour un grand r si l'on la résout naïvement avec l'algorithme des différences finies, alors que $\bar{u}$ dans l'eq. (2) peut être calculée rapidement car le coefficient constant fournit une très grande régularité. Cependant, pour un r fixe, il y a toujours un écart entre la solution réelle u et la solution effective $\bar{u}$. Récemment, Armstrong, Hannukainen, Kuusi et Mourrat ont proposé un nouvel algorithme itératif (AHKM) qui peut approximer u avec une précision arbitraire dans H^1, et le coût est proche de celui du calcul de $\bar{u}$. Dans le chapitre 2, nous présenterons cet algorithme et prouverons sa cohérence numérique.

Dans le chapitre 3, l'algorithme itératif AHKM est appliqué aux amas de percolation de $\mathbb{Z}^d$-Bernoulli ($d \geqslant 2$), qui est un modèle fondamental de milieu perforé. Plus précisément, nous échantillonnons des variables aléatoires i.i.d. de Bernoulli avec le paramètre $\mathfrak{p} \in (\mathfrak{p}_c, 1]$ où $\mathfrak{p}_c$ est le point critique et $\mathfrak{p} > \mathfrak{p}_c$ assure un unique amas infini $\mathscr{C}_\infty$. Nous étudions ensuite le problème de Dirichlet l'eq. (1) sur l'amas maximal de type $\mathscr{C}_\infty$ dans une grande boîte. Comme la condition d'ellipticité uniforme n'est plus satisfaite et la géométrie de l'amas est fractale, l'analyse devient plus difficile. La théorie de l'homogénéisation quantitative sur la percolation est initiée par Armstrong et Dario et une technique importante est une décomposition de type Calderón-Zygmund. Sur la base de ces résultats et techniques, nous prouvons une méthode numérique rigoureuse pour obtenir une approximation efficace à la fois du potentiel u et du gradient ∇u.

Le chapitre 4 se concentre sur la fonction de Green parabolique sur l'amas infini de percolation $\mathscr{C}_\infty$, c'est-à-dire $p(\cdot,\cdot,y) : \mathbb{R}^+ \times \mathscr{C}_\infty \to [0,1]$ en résolvant

$$\begin{cases} \partial_t p(\cdot,\cdot,y) - \nabla \cdot \mathbf{a} \nabla p(\cdot,\cdot,y) = 0 & \text{dans } (0,\infty) \times \mathscr{C}_\infty, \\ p(0,\cdot,y) = \delta_y & \text{dans } \mathscr{C}_\infty, \end{cases} \tag{3}$$

qui est la probabilité de transition du processus de saut à partir de $y \in \mathscr{C}_\infty$ associé au générateur $\nabla \cdot \mathbf{a} \nabla$. Ce sujet est très étudié par de nombreux pionniers et les résultats tels que la limite gaussienne, le principe d'invariance, et le théorème central limite local asymptotique ont été prouvés. Tous ces résultats nous indiquent que $p(t,\cdot,y)$ est proche d'une densité gaussienne pour t grand. Avec la collaboration de Dario, nous allons un pas plus loin pour prouver un taux de convergence quasi optimal, qui peut être interprété comme un théorème limite central quantitatif. La preuve fait appel à plusieurs résultats des travaux précédents d'Armstrong et de Dario, ainsi qu'à l'estimation du flux prouvée dans le chapitre 3.

Bien que la marche aléatoire sur l'amas infini de percolation soit compliquée, elle peut toujours être considérée comme la diffusion d'une particule dans un environnement aléatoire statique. Dans les chapitres 5 et 6, nous nous tournons vers les systèmes de particules en interaction, où l'environnement est dynamique et où il y a une infinité de particules au lieu d'une seule. Notre modèle peut être considéré comme *le processus d'exclusion symétrique généralisé* dans un espace continu. Il ne satisfait pas *la condition de gradient*, et il faut lever l'espace des fonctions défini sur la configuration des particules $\sum_{i=1}^{\infty} \delta_{x_i}$. Dans le chapitre 5, nous prouvons une limite pour la relaxation vers l'équilibre de type $t^{-\frac{d}{2}}$.

Afin de décrire le comportement asymptotique à long terme de ce nuage de particules, il faut identifier *le coefficient de diffusion* $\bar{\mathbf{a}}$, qui est une analogie du coefficient effectif pour les systèmes de particules. Le chapitre 6 présentera un travail conjoint avec Giunti et Mourrat sur l'approximation en volume fini du coefficient volumique $\bar{\mathbf{a}}$. Nous remarquons que pour les équations elliptiques, comprendre la convergence de l'approximation en volume fini de $\bar{\mathbf{a}}$ est la base de l'homogénéisation quantitative si l'on adopte l'approche de renormalisation d'Armstrong, Kuusi, Mourrat et Smart. Notre contribution consiste à généraliser cette méthode à l'analyse en dimension infinie, avec plusieurs inégalités fonctionnelles sans dimension (*l'inégalité de Poincar'e multi-échelle*, *l'inégalité de Caccioppoli* etc.) Les espaces de fonctions permettant d'étudier le système de particules sont très différents des espaces de fonctions classiques sur $\mathbb{R}^d$.

Le reste du chapitre 1 est organisé comme suit. Dans la section 0.1, nous donnerons un aperçu de la théorie de l'homogénéisation, en particulier de la méthode quantitative clé utilisée tout au long de la thèse. Ensuite, dans la section 0.2, nous exposerons les détails de l'homogénéisation dans les algorithmes numériques, et le résultat principal des chapitres 2 et 3 concernant l'algorithme AHKM. Nous passons en revue le modèle de percolation dans la section 0.3, puis nous présentons notre contribution dans le chapitre 4. La section 0.4 a pour but d'introduire les résultats des chapitres 5 et 6, et nous rappellerons aussi quelques résultats classiques dans le système de particules.

0.1 Un panorama de l'homogénéisation

La théorie de l'homogénéisation a une longue histoire et est très étudiée dans diverses directions. Le sujet le plus classique consiste à étudier le comportement de l'opérateur de forme de divergence $-\nabla \cdot (\mathbf{a}(\frac{\cdot}{\varepsilon})\nabla)$ lorsque $\varepsilon \to 0$. Les deux situations les plus typiques sont de supposer que $\mathbf{a}$ est soit périodique, soit stationnaire et ergodique. Du point de vue mathématique, il existe des résultats *qualitatifs* et *quantitatifs*. Comme on peut s'y attendre, les résultats qualitatifs ont été obtenus en premier, et ont permis d'identifier l'opérateur effectif $-\nabla \cdot (\bar{\mathbf{a}}\nabla)$, où $\bar{\mathbf{a}}$ est constant, mais $\bar{\mathbf{a}}$ n'est pas la moyenne ou l'espérance de $\mathbf{a}$. Les résultats quantitatifs ont été obtenus bien plus tard, et visaient à déterminer les taux de convergence. En fait, l'homogénéisation est également une méthode numérique utile, et les estimations d'erreurs sont des questions naturelles du point de vue de l'analyse numérique. De plus, l'homogénéisation fournit des outils pratiques pour d'autres sujets en EDP et en probabilité. Ceux-ci seront discutés en détail dans les autres sections de ce chapitre. Enfin, quel que soit le cadre (périodique, stochastique) et les objectifs (qualitatifs, quantitatifs), la plupart des résultats de la théorie de l'homogénéisation sont construits autour les trois objets clés : *la matrice de coefficient effectif*, *les correcteurs* et *l'expansion à deux échelles*.

Dans cette section, nous passons en revue certains des résultats généraux de la théorie de l'homogénéisation. Nous parlerons d'abord des résultats dans le cadre périodique, puis nous nous concentrerons sur le cas stochastique. Des monographies excellentes [47, 219, 145, 199, 210, 25] et des exposés [7, 185] sur la théorie de l'homogénéisation sont de bonnes références.

0.1.1 Homogénéisation périodique

Dans ce paragraphe, nous supposons que $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^{d\times d}_{sym}$ est $\mathbb{Z}^d$-matrice périodique, symétrique avec une condition d'ellipticité uniforme $|\xi|^2 \leqslant \xi \cdot \mathbf{a}\xi \leqslant \Lambda|\xi|^2$. Nous étudions le problème de Dirichlet pour $u^\varepsilon \in g + H^1_0(U)$

$$\begin{cases} -\nabla \cdot \left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon\right) = f & \text{dans } U, \\ u^\varepsilon = g & \text{sur } \partial U, \end{cases} \tag{4}$$

avec $f \in H^{-1}(U), g \in H^1(U)$, et $U \subseteq \mathbb{R}^d$ avec une frontière de Lipschitz. Pour $\varepsilon \to 0$, le comportement de la solution u^ε peut être approximé par la solution homogénéisée

$$\begin{cases} -\nabla \cdot (\bar{\mathbf{a}}\nabla \bar{u}) = f & \text{dans } U, \\ \bar{u} = g & \text{sur } \partial U. \end{cases} \tag{5}$$

Nous donnons son énoncé précis :

Théorème 0.1.1 ([47, 212, 187]). *Étant donné $(\mathbf{a}(x))_{x\in\mathbb{R}^d}$ un champ matriciel symétrique périodique de $\mathbb{Z}^d$ avec la condition d'ellipticité uniforme, il existe une matrice effective constante $\bar{\mathbf{a}}$, telle que la solution $(u^\varepsilon)_{\varepsilon>0}$ du problème de Dirichlet l'eq. (4) admet une solution homogénéisée $\bar{u}$ résolvant l'eq. (5) et, lorsque ε tend vers zéro,*

$$u^\varepsilon \xrightarrow{L^2(U)} \bar{u}, \qquad \nabla u^\varepsilon \xrightharpoonup{L^2(U)} \nabla \bar{u}, \qquad \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon \xrightharpoonup{L^2(U)} \bar{\mathbf{a}}\nabla \bar{u},$$

où $\nabla u^\varepsilon \xrightharpoonup{L^2(U)} \nabla \bar{u}$ et $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon \xrightharpoonup{L^2(U)} \bar{\mathbf{a}}\nabla \bar{u}$ sont compris comme la convergence faible.

Nous donnons ici une esquisse de sa preuve. Par la borne de l'estimation de l'énergie, la faible compacité de $H^1(U)$, et le théorème de Rellich, à une sous-suite près, nous avons

$$\varepsilon \to 0, \qquad u^\varepsilon \xrightarrow{L^2(U)} \bar{u}, \qquad \nabla u^\varepsilon \xrightarrow{L^2(U)} \nabla \bar{u}, \qquad \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon \xrightarrow{L^2(U)} q, \tag{6}$$

où la quantité $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon$ et sa limite faible q sont quelques fois appelées *le flux*. La question principale est de caractériser $\bar{u}, \bar{\mathbf{a}}$ et q. Une méthode heuristique classique pour ce problème est *l'ansatz d'expansion asymptotique à deux échelles*. (voir [47]) : nous écrivons de manière informelle u^ε comme suit

$$u^\varepsilon(x) = u_0\left(x, \frac{x}{\varepsilon}\right) + \varepsilon u_1\left(x, \frac{x}{\varepsilon}\right) + \varepsilon^2 u_1\left(x, \frac{x}{\varepsilon}\right) + \cdots, \tag{7}$$

où dans chaque terme $u_i : U \times \mathbb{T}^d \to \mathbb{R}$, et $u_i(x, \cdot)$ est $\mathbb{Z}^d$-périodique. L'intuition ici est de développer la fonction dans différents ordres de ε comme une série de Taylor, et d'utiliser la première coordonnée x pour décrire le comportement macroscopique, et la seconde coordonnée $\frac{x}{\varepsilon}$ pour le comportement oscillant microscopique. En comparant chaque ordre de ε, on verra pour l'ordre zéro $u_0 = \bar{u}$; pour l'ordre ε, il est décrit par *les correcteurs du premier ordre* $\{\phi_{e_i}\}_{1 \leqslant i \leqslant d}$ satisfaisant l'équation du *problème cellulaire*

$$\begin{cases} -\nabla \cdot \mathbf{a}(e_i + \nabla \phi_{e_i}) = 0 & \text{dans } \mathbb{T}^d, \\ \int_{\mathbb{T}^d} \phi_{e_i} = 0 & , \end{cases} \tag{8}$$

et $u_1\left(x, \frac{x}{\varepsilon}\right) = \sum_{i=1}^d (\partial_{x_i}\bar{u}(x))\phi_{e_i}\left(\frac{x}{\varepsilon}\right)$. Ensuite, nous calculons le flux

$$\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon = \sum_{i=1}^d \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\left(e_i + \nabla\phi_{e_i}\left(\frac{\cdot}{\varepsilon}\right)\right)\partial_{x_i}\bar{u} + O(\varepsilon), \tag{9}$$

ce qui implique la définition du coefficient homogénéisé

$$\bar{\mathbf{a}} e_i := \int_{\mathbb{T}^d} \mathbf{a}(e_i + \nabla \phi_{e_i}), \tag{10}$$

car elle nous permet de voir la limite faible dans l'eq. (9) en passant $\varepsilon \to 0$. Enfin, par le même argument, la limite faible de $\nabla \cdot \left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon\right)$ est $\nabla \cdot (\bar{\mathbf{a}}\nabla\bar{u})$ et cela donne l'eq. (5).

Cet ansatz contient de nombreux ingrédients et inspire de nombreux développements dans la théorie de l'homogénéisation. Il nous aide à dériver la définition des correcteurs dans l'eq. (8), de la matrice de coefficient effectif dans l'eq. (10) et de *l'expansion à deux échelles*.

$$w^\varepsilon := \bar{u} + \varepsilon \sum_{i=1} (\partial_{x_i}\bar{u})\phi_{e_i}\left(\frac{\cdot}{\varepsilon}\right). \tag{11}$$

Cependant, cet ansatz n'est pas rigoureux, car l'ordre d'erreur est $\|w^\varepsilon - u^\varepsilon\|_{H^1(U)} \simeq \sqrt{\varepsilon}$ en raison de *l'effet de couche limite*. Voir la discussion dans [48, 8, 26].

La première preuve rigoureuse du théorème 0.1.1 est due à De Giorgi et Spanolo [212, 213, 88], où l'argument est une méthode de style compacité pour l'opérateur différentiel $-\nabla \cdot (\mathbf{a}^\varepsilon \nabla)$. De plus, cette méthode suppose seulement la condition de matrices à coefficients symétriques $(\mathbf{a}^\varepsilon)_{\varepsilon \geqslant 0}$ avec une estimation uniforme $|\xi|^2 \leqslant \xi \cdot \mathbf{a}^\varepsilon \xi \leqslant \Lambda|\xi|^2$, elle s'applique donc à des paramètres plus généraux que le coefficient périodique ou stationnaire. Plus tard, cette méthode est étendue aux matrices asymétriques par Murat et Tartar dans [187].

Il existe également des méthodes permettant de rendre rigoureux l'ansatz d'expansion asymptotique. Une approche élégante et robuste est *la méthode de la fonction test oscillante* (également appelée *la méthode de l'énergie*) proposée par Tartar. L'idée principale est de tester l'eq. (4) avec la fonction oscillante comme $v^\varepsilon = v + \varepsilon \sum_{i=1}^d (\partial_{x_i} v)\phi_{e_i}\left(\frac{\cdot}{\varepsilon}\right)$ avec $v \in C_c^\infty(U)$ et ensuite passer ε à 0. Dans cette procédure, on a besoin de la convergence faible du produit de ∇u^ε et de $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla v^\varepsilon$, et ceci est *le théorème de compacité compensée* développé par Tartar dans [218] et par Murat dans [186]. Un autre cadre pratique pour l'homogénéisation périodique est *la convergence en deux échelles* de Nguetseng dans [194] et par Allaire dans [6], où ils définissent une topologie avec plus d'informations que la convergence faible classique.

Divers autres résultats sont développés dans l'homogénéisation périodique. Dans l'ouvrage célèbre [33, 35, 34] d'Avellaneda et Lin, ils prouvent les résultats de régularité, les théorèmes de Liouville et l'estimation de Calderón-Zygmund. Dans l'ouvrage [146, 147, 148], Kenig, Lin et Shen développent l'homogénéisation quantitative pour les systèmes elliptiques à coefficient périodique, y compris le taux de convergence pour les problèmes de Dirichlet et de Neumann et le taux de convergence pour la fonction de Green. Voir également [210] pour une revue complète.

0.1.2 Homogénéisation stochastique

La théorie de l'homogénéisation stochastique qualitative est développée dans les années 80, avec les travaux de Kozlov [160], Papanicolaou et Varadhan [198] et Yurinskiĭ [225]. Le réglage du coefficient $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^{d\times d}_{sym}$ satisfait aux conditions suivantes

1. $\mathbf{a}$ est une matrice symétrique avec une condition d'ellipticité uniforme $|\xi|^2 \leqslant \xi\cdot\mathbf{a}\xi \leqslant \Lambda|\xi|^2$;
2. $\mathbf{a}$ est un champ aléatoire ergodique à valuer dans $\mathbb{Z}^d$.

Théorème 0.1.2 ([160], [198], [225])**.** *Étant donné un champ de coefficient* $(\mathbf{a}(x))_{x\in\mathbb{R}^d}$ *satisfaisant les conditions ci-dessus, alors il existe une matrice effective constante* $\bar{\mathbf{a}}$*, telle que la solution* $(u^\varepsilon)_{\varepsilon>0}$ *du problème de Dirichlet l'eq.* (4) *admet une solution homogénéisée* $\bar{u}$ *résolvant l'eq.* (5) *et, lorsque* ε *tend vers zéro,*

$$u^\varepsilon \xrightarrow{L^2(U)} \bar{u}, \qquad \nabla u^\varepsilon \xrightarrow{L^2(U)} \nabla\bar{u}, \qquad \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon \xrightarrow{L^2(U)} \bar{\mathbf{a}}\nabla\bar{u}.$$

On peut répéter la preuve des fonctions tests oscillantes, mais une différence majeure est la construction du correcteur, car le correcteur n'est plus défini par le problème cellulaire l'eq. (8). En fait, comme la solution du problème cellulaire peut être vue comme une solution périodique dans $\mathbb{R}^d$, il est naturel de définir le correcteur ϕ_{e_i} en résolvant

$$-\nabla\cdot\mathbf{a}(e_i + \nabla\phi_{e_i}) = 0 \text{ dans } \mathbb{R}^d. \tag{12}$$

Cependant, cette équation n'est pas bien définie si on ne donne pas l'espace des fonctions. Une approche consiste à ajouter une régularisation $\lambda > 0$

$$\lambda\phi^\lambda_{e_i} - \nabla\cdot\mathbf{a}(e_i + \nabla\phi^\lambda_{e_i}) = 0 \text{ dans } \mathbb{R}^d. \tag{13}$$

Nous pouvons prendre $\lambda \searrow 0$ et extraire une sous-séquence de $\nabla\phi^\lambda_{e_i}$ qui admet une limite de $\nabla\phi_{e_i}$ en tant que champ de gradient $\mathbb{Z}^d$-stationnaire en résolvant l'eq. (12). Il suffit alors de définir

$$\bar{\mathbf{a}}e_i := \mathbb{E}\left[\int_{[0,1]^d}\mathbf{a}(e_i + \nabla\phi_{e_i})\right], \tag{14}$$

et justifier l'argument de la convergence faible par *le théorème ergodique de Birkhoff*

$$\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\left(e_i + \nabla\phi_{e_i}\left(\frac{\cdot}{\varepsilon}\right)\right) \xrightarrow{L^2} \bar{\mathbf{a}}e_i.$$

L'homogénéisation stochastique qualitative a ensuite diverses applications, mais certains aspects ne sont pas pratiques à utiliser.:

- Nous rappelons que la solution de l'eq. (12) est résolue pour $\nabla\phi_{e_i}$ au lieu de ϕ_{e_i}, donc ϕ_{e_i} est défini à une constante près et *a priori* il n'est pas stationnaire. Ceci est très différent de l'homogénéisation périodique, où ϕ_{e_i} lui-même est périodique.
- Pour obtenir $\nabla\phi_{e_i}$, nous devons résoudre le problème dans l'espace entier $\mathbb{R}^d$, ce qui est impossible en pratique. En revanche, le problème cellulaire l'eq. (8) nécessite seulement de résoudre le problème dans un tore unitaire.
- Obtenir $\bar{\mathbf{a}}$ en pratique hérite également de la difficulté de celle de $\nabla\phi_{e_i}$.

Une méthode pratique pour calculer $\bar{\mathbf{a}}$ est d'utiliser le théorème 0.1.2 dans un cube unitaire $\square := \left(-\frac{1}{2}, \frac{1}{2}\right)^d$ avec une condition limite affine

$$\begin{cases} -\nabla\cdot\left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon\right) = 0 & \text{dans } \square, \\ u^\varepsilon(x) = e_i\cdot x & \text{sur } \partial\square. \end{cases} \tag{15}$$

Puisque sa solution homogénéisée est $\bar{u}(x) = e_i\cdot x$, la convergence faible du flux est $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon \xrightarrow{L^2} \bar{\mathbf{a}}e_i$, donc nous pouvons utiliser la moyenne spatiale pour approximer $\bar{\mathbf{a}}$.

$$\int_\square \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon \xrightarrow{\varepsilon\to 0} \bar{\mathbf{a}}e_i.$$

Après un changement d'échelle, cela équivaut à approximer l'eq. (12) dans un grand cube $\square_m := \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d$ avec $\phi_{e_i,m} \in H^1_0(\square_m)$ en résolvant

$$-\nabla\cdot\mathbf{a}(e_i + \nabla\phi_{e_i,m}) = 0 \text{ dans } \square_m, \tag{16}$$

et la moyenne spatiale à grande échelle devient

$$\mathbf{a}(\square_m)e_i := \frac{1}{|\square_m|}\int_{\square_m}\mathbf{a}(e_i + \nabla\phi_{e_i,m}), \qquad \mathbf{a}(\square_m) \xrightarrow{m\to\infty} \bar{\mathbf{a}}. \tag{17}$$

Cette méthode est appelée *le volume élémentaire représentatif*, et est largement utilisée comme méthode numérique. Dans l'ouvrage [63] de Bourgeat et Piatnitski, ils prouvent la cohérence de cette méthode pour l'eq. (16) avec une condition aux limites de Dirichlet, Neumann, ou périodique. Ils obtiennent également un taux de convergence non explicite pour $\mathbb{E}[\mathbf{a}(\square_m)]$ sous certaines conditions de mélange.

La théorie quantitative de l'homogénéisation stochastique est développée ces dernières années. L'une des approches consiste à utiliser l'inégalité Efron-Stein en s'appuyant sur les idées de Naddaf et Spencer dans [188]. Dans les travaux de Gloria et Otto [123, 124], ils étudient le problème défini sur un graphe en treillis $(\mathbb{Z}^d, E_d)$ et supposent que

$$\{\mathbf{a}(e)\}_{e\in E_d} \text{ i.i.d. et } 0 < \alpha < \mathbf{a}(e) < \beta < \infty.$$

Ensuite, pour le problème résolvant l'eq. (13), ils obtiennent une estimation uniforme pour $d \geqslant 3$ [123, la proposition 2.1]

$$\mathbb{E}\left[|\phi_{e_i}^{\lambda}|^p\right] \leqslant C_p.$$

Ceci ([123, la corollaire 2.1]) répond à la question longtemps ouverte : pour $d \geqslant 3$, il existe un unique ϕ_{e_i} stationnaire résolvant l'eq. (12) tel que $\mathbb{E}[\phi_{e_i}] = 0$. Par la suite, cette méthode est également généralisée au cadre $\mathbb{R}^d$ en supposant *la condition du trou spectral* pour $\mathbf{a}$ en Gloria et Otto [125] et Gloria, Neukamm, Otto [121].

Une autre approche de l'homogénéisation quantitative est *l'approche de renormalisation* initiée par Armstrong et Smart en [31], qui ont étendu les techniques d'Avellaneda et Lin [33, 35] et celles de Dal Maso et Modica [80, 81]. Ces résultats ont ensuite été améliorés dans une série de travaux [30, 23, 24] par Armstrong, Kuusi et Mourrat, et maintenant reformulés dans la monographie [25] des mêmes auteurs. Ils travaillent sur le cadre $\mathbb{R}^d$ et supposent

$$(\mathbf{a}(x))_{x\in\mathbb{R}^d} \text{ a une corrélation de portée finie.}$$

Une corrélation en distance unitaire signifie que, pour deux ensembles quelconques $U, V \subseteq \mathbb{R}^d$ tels que $\operatorname{dist}(U,V) \geqslant 1$, les coefficients $(\mathbf{a}(x))_{x\in U}$ et $(\mathbf{a}(x))_{x\in V}$ sont indépendants. En fait, cette méthode est robuste et s'applique également aux champs de coefficients généraux avec une condition de mélange polynomiale. Comme cette thèse utilise aussi beaucoup l'approche de renormalisation, nous faisons un bref rappel dans les paragraphes suivants.

L'idée principale est similaire à l'eq. (17) et nous avons besoin du taux de convergence. Soit $\square_m = \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d$ et $\ell_p(x) := p \cdot x$, nous définissons la densité d'énergie de Dirichlet dans le volume fini

$$\nu(\square_m, p) := \inf_{v \in \ell_p + H_0^1(\square_m)} \frac{1}{|\square_m|} \int_{\square_m} \frac{1}{2} \nabla v \cdot \mathbf{a} \nabla v. \tag{18}$$

Nous désignons par $v(\square_m, p, \cdot)$ son minimiseur, et $\nu(\square_m, p) = \frac{1}{2} p \cdot \mathbf{a}(\square_m) p$ de la définition dans l'eq. (16) et l'eq. (17). On observe que $\nu(\square_m, p)$ est une *quantité sous-additive*, car pour une échelle $n < m$

$$\widetilde{v}(x) = \sum_{z \in 3^n \mathbb{Z}^d \cap \square_m} v(z + \square_n, p, x) \mathbf{1}_{\{x \in z + \square_n\}},$$

fournit un sous-minimiseur pour le problème d'optimisation de $\nu(\square_m, p)$. Nous avons alors

$$\nu(\square_m, p) \leqslant \frac{1}{|\square_m|} \int_{\square_m} \frac{1}{2} \nabla \widetilde{v} \cdot \mathbf{a} \nabla \widetilde{v} = 3^{-d(m-n)} \sum_{z \in 3^n \mathbb{Z}^d \cap \square_m} \nu(z + \square_n, p).$$

Par stationnarité, on prend l'espérance et on obtient que $\mathbb{E}[\mathbf{a}(\square_m)] \leqslant \mathbb{E}[\mathbf{a}(\square_n)]$, donc la suite décroissante $\{\mathbb{E}[\mathbf{a}(\square_m)]\}_{m \geqslant 1}$ admet une limite. Nous définissons

$$\bar{\mathbf{a}} := \lim_{m \to \infty} \mathbb{E}[\mathbf{a}(\square_m)], \tag{19}$$

et à partir de l'eq. (17), nous savons que les définitions dans l'eq. (19) et l'eq. (14) coïncident.

Afin d'obtenir le taux de convergence de $\bar{\mathbf{a}}(\square_m)$ vers $\bar{\mathbf{a}}$, nous considérons *le problème dual*

$$\nu^*(\square_m, q) := \sup_{u \in H^1(\square_m)} \frac{1}{|\square_m|} \int_{\square_m} \left(-\frac{1}{2} \nabla u \cdot \mathbf{a} \nabla u + q \cdot \nabla u\right). \tag{20}$$

Nous désignons par $u(\square_m, q, \cdot)$ le maximiseur, et $\nu^*(\square_m, q) = \frac{1}{2} q \cdot \mathbf{a}_*^{-1}(\square_m) q$ puisqu'on peut vérifier que $q \mapsto \nu^*(\square_m, q)$ est aussi une forme quadratique. Par un argument similaire et en remarquant que $u(\square_m, q, \cdot)$ est un sous-maximiseur pour tout problème $\nu^*(z+\square_n, q)$, $z \in 3^n\mathbb{Z}^d \cap \square_m$, nous avons

$$\begin{aligned}\nu^*(\square_m, q) &= 3^{-d(m-n)} \sum_{z\in 3^n\mathbb{Z}^d\cap\square_m} \frac{1}{|\square_n|}\int_{z+\square_n}\left(-\frac{1}{2}\nabla u(\square_m,q,\cdot)\cdot \mathbf{a}\nabla u(\square_m,q,\cdot) + q\cdot\nabla u(\square_m,q,\cdot)\right)\\ &\leqslant 3^{-d(m-n)} \sum_{z\in 3^n\mathbb{Z}^d\cap\square_m} \nu^*(z+\square_n, q).\end{aligned}$$

Par conséquent, $\nu^*(\square_m, q)$ est également une quantité sous-additive, et $\{\mathbb{E}[\mathbf{a}_*(\square_m)]\}_{m\geqslant 1}$ est une suite croissante. La quantité duale aide à contrôler le taux de convergence car nous pouvons tester $\nu^*(\square_m, q)$ avec le minimiseur $v(\square_m, p, \cdot)$ de $\nu(\square_m, p)$ et obtenir

$$-\frac{1}{2}p\cdot\mathbf{a}(\square_m)p + p\cdot q \leqslant \frac{1}{2}q\cdot\mathbf{a}_*^{-1}(\square_m)q.$$

En fixant $q = \mathbf{a}_*(\square_m)p$, on obtient

$$\mathbf{a}_*(\square_m) \leqslant \mathbf{a}(\square_m). \tag{21}$$

Le taux de convergence peut être majoré par

$$|\mathbb{E}[\mathbf{a}(\square_m)] - \bar{\mathbf{a}}| \leqslant |\mathbb{E}[\mathbf{a}(\square_m)] - \mathbb{E}[\mathbf{a}_*(\square_m)]|. \tag{22}$$

En pratique, l'eq. (22) peut être très utile, car les quantités $\mathbf{a}(\square_m), \mathbf{a}_*(\square_m)$ peuvent toujours être calculées localement dans $\square_m$, et la fluctuation peut être estimée par le TCL ou inégalité de concentration. Ainsi, si nous observons que $|\mathbf{a}(\square_m) - \mathbf{a}_*(\square_m)|$ est très petit, alors nous pouvons affirmer que l'approximation $|\mathbf{a}(\square_m) - \bar{\mathbf{a}}|$ est également très précise.

La preuve théorique que $\lim_{m\to\infty}|\mathbb{E}[\mathbf{a}(\square_m)] - \mathbb{E}[\mathbf{a}_*(\square_m)]| = 0$ demande plus de travail. On peut trouver sa preuve originale dans [31], ou une preuve simplifiée dans [25, le chapitre 2] où *l'inégalité de Poincaré à multi-échelle* est utilisée. Nous pouvons non seulement prouver la convergence de l'espérance, mais aussi contrôler la fluctuation : il existe un exposant $\alpha(d,\Lambda) \in \left(0, \frac{1}{2}\right]$ et, pour tout $s \in (0,d)$, une constante $C(s,d,\Lambda) < \infty$ telle que

$$|\mathbf{a}(\square_m) - \bar{\mathbf{a}}| + |\mathbf{a}_*(\square_m) - \bar{\mathbf{a}}| \leqslant C3^{-\alpha(d-s)m} + \mathcal{O}_1(3^{-sm}), \tag{23}$$

où la notation $\mathcal{O}_s$ est définie comme suit

$$X \leqslant \mathcal{O}_s(\theta) \iff \mathbb{E}\left[\exp((\theta^{-1}X)_+^s)\right] \leqslant 2. \tag{24}$$

En général, la notation $\mathcal{O}_s(\theta)$ décrit une variable aléatoire de taille typique θ avec une queue sous- ou sur-exponentielle. Lorsque l'on prend s proche de d pour réduire la part de fluctuation, l'eq. (23) permet de contrôler très étroitement la probabilité de grands écarts de $\mathbf{a}(\square_m)$. Au prix de la réduction de l'exposant s, on peut par la suite améliorer l'exposant α jusqu'à sa valeur optimale, voir [25, le chapitre 4]. Le taux de convergence pour $|\mathbf{a}(\square_m) - \bar{\mathbf{a}}|$ mesure également la convergence des correcteurs, du flux et de la solution homogénéisée, voir [25, le chapitre 1].

Enfin, l'approche de renormalisation est très robuste et s'applique à l'homogénéisation des équations paraboliques [18], aux équations à différences finies sur les amas de percolation [19, 83, 85], les formes différentielles [84], le modèle d'interface « $\nabla\phi$ » [82, 29], le modèle de Villain [86], les gaz de Coulomb [28], et les systèmes de particules en interaction [115].

0.2 Homogénéisation et algorithmes numériques

Dans cette partie, nous parlerons tout d'abord de l'intérêt de l'homogénéisation pour la résolution numérique des EDP, puis dans la section 0.2.1 nous présenterons la contribution de cette thèse (les chapitres 2 et 3) dans cette direction.

La question principale que nous espérons aborder est la méthode numérique pour le problème de Dirichlet à grande échelle : soit $U \subseteq \mathbb{R}^d$ avec une frontière de Lipschitz et $U_r := rU$.

$$\begin{cases} -\nabla \cdot (\mathbf{a}\nabla u) = f & \text{dans } U_r, \\ u = g & \text{sur } \partial U_r, \end{cases} \tag{25}$$

avec le coefficient $(\mathbf{a}(x))_{x\in\mathbb{R}^d}$ matrice symétrique satisfaisant à la condition d'ellipticité uniforme, $\mathbb{Z}^d$-périodique ou $\mathbb{Z}^d$-stationnaire et ergodique. Ce problème peut également être reformulé dans U un domaine fixé avec l'échelle ε comme l'eq. (4). Le défi ici est la nécessité de raffiner le maillage lorsque $r \to \infty$ ou $\varepsilon \to 0$, donc le coût numérique augmente et nous espérons trouver des algorithmes efficaces. La réponse à cette question dépend également de la situation concrète et nous en donnons ici un bref aperçu.

Solution en un point

Si l'on veut seulement obtenir la solution de l'eq. (25) en un point, par exemple $u(x_0), x_0 \in U_r$, alors la méthode la plus pratique est utiliser *la chaîne de Markov de Monte-Carlo (MCMC)*. Prenons un exemple simple : $f = 0$ et $g \in C^1(U_r)$. Il suffit d'exécuter une diffusion $(X_t)_{t\geqslant 0}$ associée à l'opérateur $-\nabla\cdot(\mathbf{a}\nabla)$ à partir de x_0, et que τ soit le temps d'atteinte sur la frontière ∂U_r, alors nous avons

$$u(x_0) = \mathbb{E}[g(X_\tau)]. \tag{26}$$

Cette représentation probabiliste génère un algorithme MCMC qui est également dimension libre. Il n'utilise même pas la condition périodique ou la stationnarité de $\mathbf{a}$, et fonctionne également pour un grand domaine général U_r avec une certaine régularité de la frontière. (Une condition suffisante générale est *la condition du cône*, voir les discussions dans [78].) Bien sûr, nous devons aussi faire une approximation discrète pour la diffusion, voir [76, 37, 38, 158, 176] pour l'estimation de l'erreur d'approximation.

Solution en tout point

Le défi principal consiste à résoudre l'eq. (25) pour chaque point du domaine U_r. Dans ce cas, l'algorithme MCMC nécessite également de nombreuses simulations de diffusion émises à partir de différents points de départ, ce qui augmente la complexité. Si l'on résout l'eq. (25) par la *méthode des différences finies* classique, cela revient à résoudre un grand système linéaire. Un algorithme naïf est la méthode itérative de Jacobi : après la discrétisation de l'eq. (25) dans $(\mathbb{Z}^d, E_d)$, on définit

$$P(x,y) := \frac{\mathbf{a}(\{x,y\})}{\sum_{z\sim x}\mathbf{a}(\{x,z\})}, \qquad \widetilde{f}(x) = f(x)/(\sum_{z\sim x}\mathbf{a}(\{x,z\})), \tag{27}$$

puis nous faisons l'itération

$$u_0 = g, \qquad u_{n+1} = J(u_n, \widetilde{f}), \qquad J(u_n, \widetilde{f}) := Pu_n + \widetilde{f}. \tag{28}$$

Du point de vue probabiliste, cela suit le même esprit que la méthode MCMC, mais nous faisons des itérations pour le semigroupe du problème de Dirichlet au lieu des simulations de trajectoires. Le taux de contraction dans l'eq. (28) dépend du *trou spectral*, et pour le domaine U_r, il peut être d'environ $\left(1 - \frac{1}{r^2}\right)$. Par conséquent, pour une précision de ε_0, il faut $O(r^2|\log \varepsilon_0|)$ tours d'itérations. Cet algorithme peut être légèrement accéléré par *la méthode du gradient conjugué (CGM)*, qui atteint un taux de contraction $\left(1 - \frac{1}{r}\right)$; il suffit donc de $O(r|\log \varepsilon_0|)$ tours de CGM (voir [207, le théorème 6.29, l'eq.(6.128)]). Comme le coût numérique d'une itération de CGM est proche de celui de la méthode de Jacobi, toutes ces méthodes auront une grande complexité lorsque r augmente.

Solution de l'équation à coefficient constant en tout point

L'algorithme multigrille est une méthode puissante pour le problème de Dirichlet dans un grand domaine avec coefficient constant. On peut trouver l'étude complète de cette méthode dans [137, 223, 99, 64] et ici nous donnons une version dans notre contexte. Supposons que nous voulions résoudre $-\Delta u = f$ avec $u \in g + H_0^1(U_r)$, l'algorithme peut être énoncé comme suit : définir la grille la plus fine d'échelle $\frac{r}{M}$, et désigner par J^M la méthode de Jacobi dans l'eq. (28) pour cette grille.

1. Commencez par une estimation initiale $u_0 = g$.
2. Implémentez une étape d'itération multigrille avec la méthode de Jacobi
   (a) $u_1 = J^M(u_0, f)$;
   (b) $f_1 = f - (-\Delta u_1)$, grossir la grille de 2, et $u_2 = J^{M/2}(0, f_1)$;
   (c) $f_2 = f_1 - (-\Delta u_2)$, rend la grille plus grossière de 2, et $u_3 = J^{M/4}(0, f_2)$.
3. Définissez $\hat{u} := u_1 + u_2 + u_3$ et mettez $\hat{u}$ à la place de u_0. Revenir à l'étape 2 et répéter cette procédure d'itérations.

En pratique, il faut ajouter plusieurs échelles intermédiaires dans l'étape d'itération multigrille. Remarquons que la grille grossière n'est pas précise, mais elle peut récupérer le comportement macroscopique de la solution avec un coût numérique moindre ; la grille fine peut calculer la solution à l'échelle microscopique, mais la valeur se propage lentement dans la méthode de Jacobi et nécessite de nombreuses étapes de calcul. Par conséquent, nous combinons différentes grilles et pouvons résoudre cette solution plus efficacement. Pour la précision ε_0, le coût numérique est d'environ $O(|\log \varepsilon_0|)$ tours de CGM ([64, le chapitre 4]) - nous pouvons toujours remplacer la méthode de Jacobi dans l'algorithme par la méthode du gradient conjugué, mais la première est plus facile à énoncer. Enfin, nous remarquons que certaines opérations sont nécessaires pour le passage des fonctions entre la grille fine et la grille grossière. Elles sont appelées *l'opérateur de grossissement* et *l'opérateur de projection*, qui sont des idées très importantes dans l'algorithme multigrille. Dans notre cadre, l'opérateur de grossissement et l'opérateur de projection ne sont que des échantillons de grilles et des interpolations linéaires, car Δ donne plus de régularité à la solution par rapport à $-\nabla \cdot (\mathbf{a}\nabla)$. Cela explique également pourquoi l'algorithme multigrille classique requiert la condition de coefficient constant.

Solution homogénéisée

L'algorithme multigrille ci-dessus explique l'intérêt de l'homogénéisation pour la résolution numérique du problème de Dirichlet. Au lieu de résoudre directement le l'eq. (25), nous pouvons résoudre sa solution homogénéisée $\bar{u}$ avec l'algorithme multigrille. Nous n'avons alors à payer qu'une petite erreur $\|u - \bar{u}\|_{L^2(U_r)}$, et par la théorie de l'homogénéisation, cette erreur est assez faible comparée à $\|u\|_{L^2(U_r)}$ pour les grands r. Nous renvoyons aux références [42, 36, 101, 128, 197, 178, 159, 196], ainsi qu'à [129, 154, 103, 104] pour cette idée.

Par conséquent, lorsque nous combinons la solution homogénéisée et l'algorithme multigrille, il suffit d'obtenir le coefficient effectif $\bar{\mathbf{a}}$. Comme nous en avons beaucoup discuté dans la section 0.1, cette tâche est plus compliquée dans l'homogénéisation stochastique que dans l'homogénéisation périodique.

- Pour $\mathbf{a}$ un coefficient $\mathbb{Z}^d$-périodique, nous pouvons obtenir $\bar{\mathbf{a}}$ en résolvant le problème cellulaire l'eq. (8) mentionné dans la section 0.1.

- Pour le paramètre de coefficient stochastique, nous utilisons le volume élémentaire représentatif (REV) mentionné dans l'eq. (17). Plus précisément, nous divisons les données $(\mathbf{a}(x))_{x \in U_r}$ en sous-ensembles d'échelle l, appliquons l'eq. (17) dans chaque sous-ensemble, puis faisons la moyenne sur les $\left(\frac{r}{l}\right)^d$ copies pour réduire la fluctuation.

Pour le cas du coefficient stochastique, il existe de nombreuses références [119, 102, 184, 107, 138] discutant des erreurs et des coûts numériques. Parmi eux, [102] a étudié le modèle sur $(\mathbb{Z}^d, E_d)$ avec une conductance i.i.d. $\{\mathbf{a}(e)\}_{e \in E_d}$. Son résultat principal dit qu'avec un choix $l = r^{\frac{1}{2}}$ dans le REV, on a une précision $r^{-\frac{d}{2}}$ avec une complexité $O(r^{\frac{1}{2}})$ rounds de CGM. Plus tard, [184] a proposé un autre algorithme efficace qui nous permet d'obtenir la complexité optimale dans le cadre général de l'homogénéisation stochastique : une précision $r^{-\frac{d}{2}}$ avec la complexité $O(\log r)$ tours de CGM. Quelle que soit la méthode utilisée, obtenir $\bar{\mathbf{a}}$ avec une bonne précision n'est pas très coûteux.

Au-delà de la solution homogénéisée

Bien que la solution homogénéisée $\bar{u}$ soit une bonne approximation de l'eq. (25), elle est trop lisse pour récupérer les détails microscopiques. Pour aller plus loin, une stratégie consiste à utiliser la méthode d'expansion à deux échelles. Comme mentionné dans l'eq. (11), $\bar{u}$ ne converge vers u que dans L^2, mais $w = \bar{u} + \sum_{i=1}^d (\partial_{x_i}\bar{u})\phi_{e_i}$ donne une approximation de la solution de l'eq. (25) au sens H^1. Ainsi, dans le cadre des coefficients périodiques, nous pouvons attaquer en premier le problème cellulaire l'eq. (8) pour obtenir à la fois $\bar{\mathbf{a}}$ et tous les correcteurs du premier ordre $\{\phi_{e_i}\}1 \leqslant i \leqslant d$, puis nous résolvons la solution homogénéisée $\bar{u}$. En combinant les correcteurs et la solution homogénéisée, w nous donne une meilleure approximation. Cette méthode peut aussi être un peu améliorée en utilisant un correcteur modifié, à savoir

$$\widetilde{w} = \bar{u} + \sum_{i=1}^{d} (\partial_{x_i}\bar{u})\phi_{e_i}\eta, \tag{29}$$

avec $\eta \in C_c^\infty(U_r)$ une fonction de coupure lisse afin d'éliminer les principales sources d'erreur - *l'effet de couche limite*. Voir la discussion de ce sujet dans [48, 8, 26]. Malheureusement, cette idée peut difficilement être utilisée dans l'homogénéisation stochastique car la complexité pour calculer les correcteurs est la même que pour calculer la solution originale u.

Remarquons que pour un r fixe, il y a toujours une limite de précision entre les solutions approchées $\bar{u}, w, \widetilde{w}$ et la solution réelle u, il est donc naturel de chercher un algorithme efficace avec une résolution au-delà de cette limite. D'après la discussion ci-dessus, il semble qu'il y ait un compromis entre la précision et le coût numérique. En fait, il existe une troisième dimension : la probabilité. Nous allons voir que nous pouvons payer une certaine probabilité de cohérence pour gagner à la fois en précision et en efficacité numérique. Un exemple est l'algorithme itératif AHKM [22], qui sera discuté en détail dans la section 0.2.1.

0.2.1 Résumé des chapitres 2 et 3

L'algorithme itératif AHKM est le sujet principal étudié dans les chapitres 2 et 3. Il est inventé par Armstrong, Hannukainen, Kuusi et Mourrat dans [22], qui vise à obtenir une approximation de u au-delà de la précision de la solution homogénéisée $\bar{u}$ avec des coûts numériques raisonnables. Il suit l'esprit de l'algorithme multigrille et fait également appel à la théorie de l'homogénéisation.

Présentons d'abord la structure de l'algorithme AHKM pour l'eq. (25).

1. Nous commençons par une estimation initiale $u_0 = g$, et choisissons un paramètre de régularisation $\lambda \in \left(\frac{1}{r}, \frac{1}{2}\right)$.
2. Nous résolvons les systèmes suivants
$$\begin{cases} (\lambda^2 - \nabla \cdot \mathbf{a}\nabla)u_1 & = f + \nabla \cdot \mathbf{a}\nabla u_0 & \text{dans } U_r, \\ -\nabla \cdot \bar{\mathbf{a}}\nabla \bar{u} & = \lambda^2 u_1 & \text{dans } U_r, \\ (\lambda^2 - \nabla \cdot \mathbf{a}\nabla)u_2 & = (\lambda^2 - \nabla \cdot \bar{\mathbf{a}}\nabla)\bar{u} & \text{dans } U_r. \end{cases} \tag{30}$$
3. Nous fixons $\hat{u} := u_0 + u_1 + u_2$ et le remettons à la place de u_0 pour répéter les itérations de l'étape 2.

Cela ressemble beaucoup à l'algorithme multigrille : dans la première équation de l'eq. (30), nous résolvons le problème de Dirichlet en grille fine. Mais nous ajoutons une certaine régularisation pour réduire les tours de CGM. Puisque $(u_0 + u_1)$ ne peut pas récupérer toute la solution, nous mettons le résidu

$$\lambda^2 u_1 = f - \left(-\nabla \cdot \mathbf{a}\nabla\left(u_0 + u_1\right)\right),$$

comme source dans le côté droit de la deuxième équation de l'eq. (30). Dans la deuxième équation de l'eq. (30), nous résolvons simplement le problème sur une grille grossière avec la solution homogénéisée. Cependant, la solution homogénéisée est trop lisse pour la grille fine. Ainsi, dans la troisième équation de l'eq. (30), nous effectuons un post-traitement et on peut considérer u_2 comme la projection de $\bar{u}$ dans la grille de recherche pour l'opérateur $\lambda^2 - \nabla \cdot \mathbf{a}\nabla$.

Pour prouver la cohérence de l'algorithme AHKM, l'ingrédient principal est l'expansion à deux échelles $w := \bar{u} + \sum_{k=1}^{d}(\partial_{x_k}\bar{u})\phi_{e_k}$. En combinant la première équation, la deuxième équation de l'eq. (30) et l'eq. (25), nous pouvons obtenir que

$$-\nabla \cdot \bar{\mathbf{a}}\nabla\bar{u} = -\nabla \cdot \mathbf{a}\nabla(u - u_0 - u_1) \qquad \text{dans } U_r,$$

qui est une équation d'homogénéisation, on a donc $(u - u_0 - u_1) \simeq w$. De plus, la troisième équation dans l'eq. (30) suit également la forme d'homogénéisation. Ainsi, nous avons

$$(u - u_0 - u_1) \simeq w \simeq u_2,$$

à une petite erreur près dans le sens $H^1(U_r)$, donc nous pouvons estimer $\|\hat{u} - u\|_{H^1(U_r)}$ en étudiant

$$\|\hat{u} - u\|_{H^1(U_r)} = \|u - (u_0 + u_1 + u_2)\|_{H^1(U_r)} \leqslant \|(u - u_0 - u_1) - w\|_{H^1(U_r)} + \|w - u_2\|_{H^1(U_r)}.$$

La plupart de l'idée ci-dessus a déjà été incluse dans l'article [22], mais comme l'environnement est aléatoire, le taux de contraction est également une variable aléatoire. Dans [22], les auteurs obtiennent une borne pour ce taux de contraction d'un pas, mais cette estimation ne peut pas être itérée. La contribution dans le chapitre 2 est une *borne uniforme* pour le taux de contraction. Cette borne uniforme peut ensuite être itérée pour justifier la validité de l'algorithme. Dans l'énoncé suivant, la notation $\mathcal{O}_s$ est définie dans l'eq. (24) et $\ell(\lambda)$ est défini comme suit

$$\ell(\lambda) := \begin{cases} (\log(1+\lambda^{-1}))^{\frac{1}{2}} & d = 2, \\ 1 & d > 2. \end{cases}$$

Théorème 0.2.1 (Le théorème principal dans le chapitre 2)**.** *Pour tout domaine borné $U \subseteq \mathbb{R}^d$ avec un bord de $C^{1,1}$ et tout $s \in (0,2)$, il existe une constante finie positive $C(U,\Lambda,s,d)$ et, pour tout $r \geqslant 2$ et $\lambda \in \left(\frac{1}{r}, \frac{1}{2}\right)$, une variable aléatoire $\mathcal{Z}$ satisfaisant à*

$$\mathcal{Z} \leqslant \mathcal{O}_s\left(C\ell(\lambda)^{\frac{1}{2}}\lambda^{\frac{1}{2}}(\log r)^{\frac{1}{s}}\right), \tag{31}$$

tel que ce qui suit est établi. Soit $U_r := rU$, $f \in H^{-1}(U_r)$, $g \in H^1(U_r)$, $u_0 \in g + H^1_0(U_r)$, $u \in g + H^1_0(U_r)$ la solution de l'eq. (25), et laissons $u_1, \bar{u}, u_2 \in H^1_0(U_r)$ résoudre l'eq. (30) avec une condition au bord de Dirichlet nulle. Alors pour $\hat{u} := u_0 + u_1 + u_2$, nous avons l'estimation de contraction

$$\|\nabla(\hat{u} - u)\|_{L^2(U_r)} \leqslant \mathcal{Z}\|\nabla(u_0 - u)\|_{L^2(U_r)}. \tag{32}$$

Par conséquent, le taux de contraction de l'algorithme AHKM peut être limité par une variable aléatoire $\mathcal{Z}$ de l'ordre de $\lambda^{\frac{1}{2}}$, et plus précisément,

$$\mathbb{P}[\mathcal{Z} \geqslant x] \leqslant 2\exp\left(-\left(\frac{x}{C\ell(\lambda)^{\frac{1}{2}}\lambda^{\frac{1}{2}}(\log r)^{\frac{1}{s}}}\right)^s\right).$$

Par un choix raisonnable $\lambda \simeq (\log r)^{-1}$, pour une précision ε_0 la complexité de l'algorithme AHKM est $O(\log r|\log\varepsilon_0|^2)$. En conclusion, l'algorithme AHKM permet d'obtenir à la fois une grande précision et un coût faible en numérique, au prix de l'exclusion d'un événement de probabilité très faible.

L'algorithme AHKM est une méthode assez robuste et il s'applique également à d'autres problèmes de Dirichlet dans un environnement aléatoire dégénéré. La principale contribution dans le chapitre 3 est un exemple pour son application sur *l'amas de percolation*, qui peut être utilisé pour simuler le modèle en milieu poreux de deux types de composites à fort contraste. Voir [220] pour une introduction complète et [95, 163, 175] pour quelques exemples de ses applications dans les nanomatériaux.

Nous donnons ici une brève introduction du modèle de percolation. Vous trouverez plus de détails dans la section 0.3. Sur le graphe en réseau $(\mathbb{Z}^d, E_d)$, soit $\mathbf{a} : E_d \to \{0\} \cup [\Lambda^{-1}, 1]$ telle que les variables aléatoires $\{\mathbf{a}(e)\}_{e \in E_d}$ sont indépendantes et identiquement distribuées. La percolation de Bernoulli est définie par la conductance aléatoire $\{\mathbf{a}(e)\}_{e \in E_d}$: pour toute liaison $e \in E_d$, on dit que e est une *liaison ouverte* si $\mathbf{a}(e) > 0$, et que e est une *liaison fermée* sinon. Les composantes connectées sur $(\mathbb{Z}^d, E_d)$ générées par les liaisons ouvertes sont appelées *amas*. Pour $d \geqslant 2$, il existe un paramètre $\mathfrak{p}_c(d)$ tel que pour $\mathfrak{p} := \mathbb{P}[\mathbf{a}(e) > 0] > \mathfrak{p}_c$, il existe un unique amas de percolation infini $\mathscr{C}_\infty$ [149]. Ce cas est appelé *la percolation surcritique*, et sous ce paramètre dans un cube fini $\square_m := \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d \cap \mathbb{Z}^d$, typiquement nous verrons un amas géant $\mathscr{C}_*(\square_m)$. Il s'agit d'un analogie de $\mathscr{C}_\infty$ (voir le figure 1 pour une illustration) et nous appelons ce cas « $\square_m$ *est un bon cube* ». Les définitions rigoureuses de « $\square_m$ est un bon cube » et de « l'amas maximal $\mathscr{C}_*(\square_m)$ » seront données dans la section 3.2, et elles sont typiques puisqu'il existe une constante positive $C(d,\mathfrak{p})$ telle que

$$\mathbb{P}[\square_m \text{ is a good cube}] \geqslant 1 - C(d,\mathfrak{p})\exp(-C(d,\mathfrak{p})^{-1}3^m).$$

De manière informelle, on peut simplement traiter $\mathscr{C}_*(\square_m)$ comme $\mathscr{C}_\infty \cap \square_m$. Notre objectif est de trouver un algorithme pour résoudre le problème de Dirichlet sur $\mathscr{C}_*(\square_m)$

$$\begin{cases} -\nabla \cdot \mathbf{a}\nabla u = f & \text{dans } \mathscr{C}_*(\square_m), \\ u = g & \text{sur } \mathscr{C}_*(\square_m) \cap \partial\square_m, \end{cases} \tag{33}$$

où l'opérateur de forme de divergence est défini comme suit

$$-\nabla \cdot \mathbf{a}\nabla u(x) \coloneqq \sum_{y\sim x} \mathbf{a}(\{x,y\})\,(u(x)-u(y))\,. \tag{34}$$

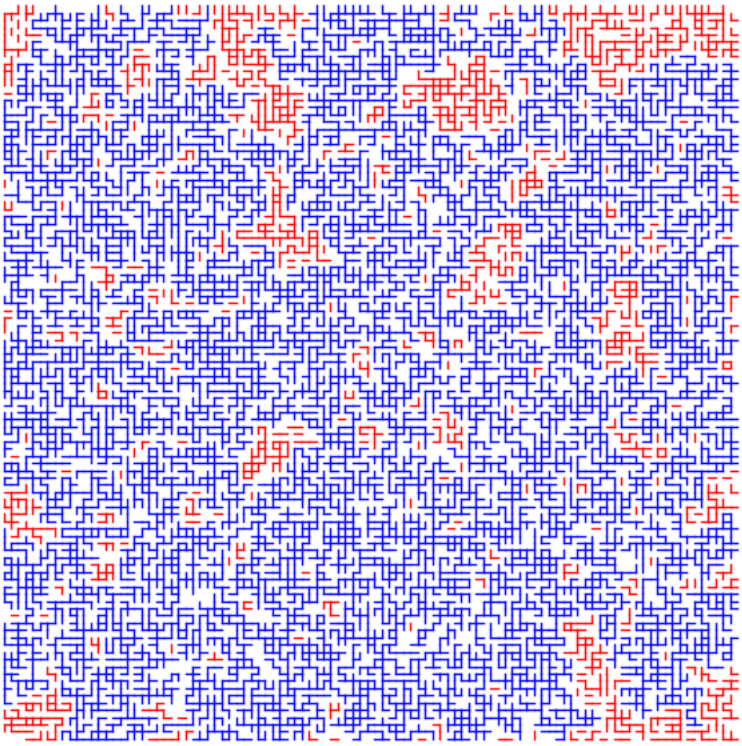

Figure 1: Une simulation de percolation de Bernoulli $2D$ avec $\mathfrak{p} = 0.51$ dans un cube $\square$ de taille 100×100. Le composant en bleu est l'amas maximal $\mathscr{C}_*(\square)$ tandis que les composants en rouge sont les autres petits amas.

L'algorithme AHKM sur l'amas de percolation est le suivant : nous désignons par $C_0(\square_m)$ les fonctions à condition limite nulle sur $\square_m$ et $\lambda_{\mathscr{C},m} \coloneqq \lambda \mathbf{1}_{\{\mathscr{C}_*(\square_m)\}}$.

Théorème 0.2.2 (Le théorème principal dans le chapitre 3)**.** *Il existe deux constantes positives finies $s \coloneqq s(d,\mathfrak{p},\Lambda), C \coloneqq C(d,\mathfrak{p},\Lambda,s)$, et pour tout entier $m > 1$ et $\lambda \in \left(\frac{1}{3^m},\frac{1}{2}\right)$, une variable aléatoire $\mathcal{Z}$ mesurable par $\mathcal{F}$ satisfaisant à*

$$\mathcal{Z} \leqslant \mathcal{O}_s\left(C\ell(\lambda)^{\frac{1}{2}}\lambda^{\frac{1}{2}}m^{\frac{1}{s}+d}\right),$$

de telle sorte que la règle suivante s'applique. Soit $f, g : \square_m \to \mathbb{R}, u_0 \in g + C_0(\square_m)$ *et* $u \in g + C_0(\square_m)$ *la solution de l'eq. (33). Au cas où* $\square_m$ *est un bon cube, pour* $u_1, \bar{u}, u_2 \in C_0(\square_m)$ *qui résolvent (avec une condition au bord de Dirichlet nulle)*

$$\begin{cases} (\lambda^2 - \nabla \cdot \mathbf{a} \nabla) u_1 &= f + \nabla \cdot \mathbf{a} \nabla u_0 & \textit{dans } \mathscr{C}_*(\square_m) \backslash \partial \square_m, \\ -\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} &= \lambda^2_{\mathscr{C},m} u_1 & \textit{dans } \operatorname{int}(\square_m), \\ (\lambda^2 - \nabla \cdot \mathbf{a} \nabla) u_2 &= (\lambda^2 - \nabla \cdot \bar{\mathbf{a}} \nabla) \bar{u} & \textit{dans } \mathscr{C}_*(\square_m) \backslash \partial \square_m, \end{cases} \tag{35}$$

et pour $\hat{u} := u_0 + u_1 + u_2$*, nous avons l'estimation de contraction*

$$\left\| \nabla (\hat{u} - u) \mathbf{1}_{\{\mathbf{a} \neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))} \leqslant \mathcal{Z} \left\| \nabla (u_0 - u) \mathbf{1}_{\{\mathbf{a} \neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))}. \tag{36}$$

La nouveauté de cette application sur l'amas de percolation est de définir un algorithme multigrille sur l'environnement aléatoire singulier car $\mathbf{a}$ n'a pas une ellipticité uniforme. Par conséquent, dans la première et la troisième équation de l'eq. (35), la grille fine est définie sur l'amas de percolation, tandis que la grille grossière de la deuxième équation de l'eq. (35) est définie sur $\square_m$. Cela implique que non seulement le coefficient aléatoire, mais aussi la géométrie aléatoire est homogénéisée. Pour voir plus précisément que l'eq. (35) définit l'opérateur d'approximation et de projection approprié, nous observons que u_1, u_2 résout également les itérations équivalentes suivantes avec *n'importe quelle extension arbitraire* de valeur sur $\square_m \backslash \mathscr{C}_*(\square_m)$

$$\begin{cases} (\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla) u_1 &= f_{\mathscr{C},m} + \nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla u_0 & \text{dans } \operatorname{int}(\square_m), \\ -\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} &= \lambda^2_{\mathscr{C},m} u_1 & \text{dans } \operatorname{int}(\square_m), \\ (\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla) u_2 &= (\lambda^2_{\mathscr{C},m} - \nabla \cdot \bar{\mathbf{a}} \nabla) \bar{u} & \text{dans } \operatorname{int}(\square_m), \end{cases} \tag{37}$$

où $\mathbf{a}_{\mathscr{C},m}(\{x,y\}) = \mathbf{a}(\{x,y\}) \mathbf{1}_{\{x,y \in \mathscr{C}_*(\square_m)\}}$ et $f_{\mathscr{C},m} = f \mathbf{1}_{\{\mathscr{C}_*(\square_m)\}}$. Voir plus de détails dans la proposition 3.1.1. Un deuxième défi dans cette application est l'analyse de cohérence, car la théorie d'homogénéisation quantitative sur les grappes de percolation est absente depuis longtemps jusqu'aux travaux récents de [19, 83].

0.3 Homogénéisation sur l'amas de percolation

Dans cette partie, nous présenterons tout d'abord *le modèle de percolation de* $\mathbb{Z}^d$*-Bernoulli*, puis nous passerons en revue les résultats de la marche aléatoire sur celui-ci. Nous soulignerons également ses liens avec la théorie de l'homogénéisation, et présenterons notre contribution du chapitre 4 dans la section 0.3.

Le modèle de percolation $\mathbb{Z}^d$-Bernoulli est d'abord introduit par Broadbent et Hammersley afin d'étudier les milieux poreux. Nous donnons ici sa définition dans notre contexte : soit $(\mathbb{Z}^d, E_d)$ un graphe en réseau, *la conductance aléatoire* $\mathbf{a} : E_d \to \{0\} \cup [\lambda, 1]$ et $\{\mathbf{a}(e)\}_{e \in E_d}$ variables aléatoires qui sont identiquement et indépendamment distribuées. On dit qu'une arête e est *ouverte* si $\mathbf{a}(e) > 0$ et e est *fermée* si $\mathbf{a}(e) = 0$. Les composantes connectées définies par les arêtes ouvertes sont appelées *les amas*, et nous désignons par $x \longleftrightarrow y$ si x et y sont dans le même amas. Un cas particulier $x \longleftrightarrow \infty$ implique un *amas infini* $\mathscr{C}_\infty$ contenant x. Le comportement des amas dépend du paramètre $\mathfrak{p} := \mathbb{P}[\mathbf{a}(e) > 0]$ et nous désignons par $\theta(\mathfrak{p}) := \mathbb{P}[0 \longleftrightarrow \infty]$ *le paramètre de connectivité*. Pour $d = 1$ le comportement des amas est trivial, et pour $d \geqslant 2$ il existe une *transition de phase* dans ce modèle : il existe un *point critique* $\mathfrak{p}_c \in (0,1)$ tel que

1. *Phase souscritique* : $\mathfrak{p} \in [0, \mathfrak{p}_c)$, il n'y a pas d'amas infini et $\theta(\mathfrak{p}) = 0$.

2. *Phase surcritique* : $\mathfrak{p} \in (\mathfrak{p}_c, 1]$, il existe un unique amas infini $\mathscr{C}_\infty$ et $\theta(\mathfrak{p}) > 0$.

3. *Phase critique* : $\mathfrak{p} = \mathfrak{p}_c$, on sait que $\theta(\mathfrak{p}_c) = 0$ pour $d = 2$ et $d \geqslant 11$, mais pour $3 \leqslant d \leqslant 10$ c'est encore une conjecture que $\theta(\mathfrak{p}_c) = 0$.

Voir la monographie [131] et l'enquête récente [97] pour plus d'informations sur la percolation. Nous nous intéressons à la marche aléatoire sur le modèle de percolation surcritique. Ce modèle peut être utilisé pour décrire la diffusion dans des milieux poreux, ou dans des matériaux bicomposés à fort contraste. Plus précisément, laissez $\mathfrak{p} > \mathfrak{p}_c(d)$ et nous considérons *la marche aléatoire à vitesse variable.* (VSRW), qui est un processus de saut de Markov à temps continu $(X_t)_{t\geqslant 0}$ commençant par un certain $y \in \mathscr{C}_\infty$, et associé au *générateur*

$$\mathcal{L}u(x) = \nabla \cdot \mathbf{a}\nabla u(x) := \sum_{z\sim x} \mathbf{a}(\{x,z\})\,(u(z) - u(x))\,. \tag{38}$$

Nous désignons *le semigroupe* (ou *la probabilité de transition*) de la marche aléatoire par

$$p\,(t,x,y) = p^{\mathbf{a}}\,(t,x,y) := \mathbb{P}^{\mathbf{a}}_y\,(X_t = x)\,,$$

qui est définie comme la solution de l'équation parabolique

$$\begin{cases} \partial_t p(\cdot,\cdot,y) - \nabla\cdot\mathbf{a}\nabla p(\cdot,\cdot,y) = 0 & \text{dans } (0,\infty)\times\mathscr{C}_\infty, \\ p(0,\cdot,y) = \delta_y & \text{dans } \mathscr{C}_\infty. \end{cases} \tag{39}$$

En raison de cette caractérisation, nous faisons souvent référence au semigroupe $p(t,\cdot,y)$ comme étant *le noyau de chaleur* ou *la fonction de Green parabolique.*

Nous remarquons que la VSRW définie ci-dessus n'est qu'une façon possible de construire la marche aléatoire sur $\mathscr{C}_\infty$ et qu'il existe d'autres modèles similaires. Deux des modèles les plus connus sont :

1. *La marche aléatoire à vitesse constante* (CSRW) : il s'agit d'un processus de saut de Markov à temps continu commençant par $y \in \mathscr{C}_\infty$, avec un taux de saut de 1 et la probabilité de transition
$$P(x,z) = \frac{\mathbf{a}\,(\{x,z\})}{\sum_{w\sim x}\mathbf{a}\,(\{x,w\})}. \tag{40}$$
En d'autres termes, son générateur associé est
$$\mathcal{L}u(x) = \frac{1}{\sum_{w\sim x}\mathbf{a}\,(\{x,w\})}\sum_{z\sim x}\mathbf{a}(\{x,z\})\,(u(z)-u(x))\,.$$

2. *La marche aléatoire en temps discret* (DTRW) : la marche aléatoire $(X_n)_{n\in\mathbb{N}}$ est indexée sur les entiers. Elle part d'un point $y \in \mathscr{C}_\infty$, et lorsque $X_n = x$, la valeur de X_{n+1} est choisie aléatoirement parmi tous les voisins de x suivant la probabilité de transition (40).

Ces processus ont des propriétés similaires, mais pas identiques, et ont fait l'objet d'un intérêt dans la littérature.

La marche aléatoire sur l'amas de percolation est un sujet parmi les *modèles de conductance aléatoire* plus généraux, où de nombreux modèles appartiennent à *l'universalité brownienne.* Par exemple, pour la VSRW sur $(\mathbb{Z}^d, E_d)$ avec $\{\mathbf{a}(e)\}_{e\in E_d}$ i.i.d. satisfaisant la condition d'ellipticité uniforme $0 < \lambda \leqslant \mathbf{a} \leqslant 1$, alors son semigroupe $p(t, \cdot, y)$ a une *borne gaussienne*

$$\forall |x-y| \leqslant t, \qquad \frac{C_1}{(2\pi t)^{\frac{d}{2}}} \exp\left(-\frac{|x-y|^2}{2C_1 t}\right) \leqslant p(t,x,y) \leqslant \frac{C_2}{(2\pi t)^{\frac{d}{2}}} \exp\left(-\frac{|x-y|^2}{2C_2 t}\right). \tag{41}$$

De plus, le processus a presque sûrement une limite d'échelle du mouvement brownien dans la topologie de Skorokhod.

$$\left(\frac{1}{\sqrt{n}} X_{nt}\right)_{t \geqslant 0} \overset{n\to\infty}{\Longrightarrow} (\bar{\sigma} B_t)_{t \geqslant 0}.$$

C'est *le principe d'invariance presque sûr* pour $(X_t)_{t\geqslant 0}$, où l'environnement $\{\mathbf{a}(e)\}_{e \in E_d}$ est fixé dans l'énoncé. Il est généralement plus facile d'établir *le principe d'invariance moyenné* en faisant une moyenne sur l'environnement.

Faisons quelques remarques supplémentaires sur ces deux résultats. La borne gaussienne pour l'opérateur de type divergence est initiée par les travaux de De Giorgi, Moser et Nash sur $\mathbb{R}^d$, puis elle est généralisée aux collecteurs par Grigor'yan dans [130] et par Saloff-Coste dans [208]. Pour le CSRW sur $\mathbb{Z}^d$, sa preuve peut être trouvée dans le travail de Delmotte [92], où le théorème est connu sous le nom de *« la condition de doublement du volume et l'inégalité de Poincaré impliquent la borne gaussienne »*. La condition $|x-y| \leqslant t$ dans l'eq. (41) est nécessaire pour la borne gaussienne sur $(\mathbb{Z}^d, E_d)$, car le générateur est un opérateur de différence finie au lieu d'un opérateur différentiel. Pour le régime $|x-y| \geqslant t$, la queue est exponentielle plutôt que gaussienne ; voir les travaux de Davies [87]. Le principe d'invariance éteint a un lien très étroit avec la théorie de l'homogénéisation. Un outil puissant pour le prouver est la méthode des correcteurs initiée par Kozlov dans [160] : laissez $\{\phi_{e_i}\}_{1\leqslant i \leqslant d}$ être les correcteurs du premier ordre associés à la base canonique $\{e_i\}_{1\leqslant i\leqslant d}$ et au générateur $\nabla \cdot \mathbf{a} \nabla$, alors

$$M_t = (X_t \cdot e_1 + \phi_{e_1}(X_t), \cdots, X_t \cdot e_d + \phi_{e_d}(X_t)),$$

est une martingale. Maintenant le théorème de convergence des martingales [139] s'applique

$$\left(\frac{1}{\sqrt{n}} M_{nt}\right)_{t \geqslant 0} \overset{n\to\infty}{\Longrightarrow} (\bar{\sigma} B_t)_{t \geqslant 0}.$$

Il suffit de prouver que la partie du correcteur disparaît presque sûrement $\frac{\phi_{e_1}(X_{nt})}{\sqrt{n}} \overset{n\to\infty}{\longrightarrow} 0$ et ceci se réduit finalement à *la sous-linéarité* des correcteurs.

La marche aléatoire sur $\mathscr{C}_\infty$ appartient également à l'universalité brownienne. Une explication intuitive est que la géométrie de $\mathscr{C}_\infty$ est très proche de celle de $\mathbb{Z}^d$ à grande échelle. Dans le cas de la percolation, où $\mathbf{a}$ ne peut prendre que les valeurs 0 ou 1, un principe d'invariance moyenné a été prouvé dans [89] par De Masi, Ferrari, Goldstein et Wick. Dans [211], Sidoravicius et Sznitman ont prouvé un principe d'invariance presque sûr pour la marche aléatoire en dimension $d \geqslant 4$. Ce résultat a été étendu à toute dimension $d \geqslant 2$ par Berger et Biskup dans [49] (pour le DTRW) et par Mathieu et Piatnitski dans [180] (pour le CSRW), où leur stratégie consiste à construire les correcteurs sur $\mathscr{C}_\infty$. Les propriétés du noyau de chaleur $p(t, \cdot, y)$ sur l'amas infini ont été étudiées dans la littérature. Dans [181], Mathieu et Remy ont prouvé que, presque sûrement, le noyau de chaleur décroît aussi vite que $t^{-d/2}$. Ces limites ont été étendues dans [39] par Barlow qui a établi des bornes inférieures et supérieures gaussiennes.

Pour le VSRW, un principe d'invariance presque sûr similaire s'applique. Du point de vue de l'homogénéisation, la diffusivité $\bar{\sigma}$ du mouvement brownien limite est liée au coefficient effectif $\bar{\mathbf{a}}$ des problèmes elliptiques par l'identité $\bar{\mathbf{a}} = \frac{1}{2}\theta(\mathfrak{p})\bar{\sigma}^2$.

Dans l'article [41], Barlow et Hambly ont prouvé une inégalité de Harnack parabolique, un théorème central limite local pour le CSRW, et des bornes sur la fonction de Green elliptique sur l'amas infini. Leur résultat principal peut être adapté au cas du VSRW, et se lit comme suit : si nous définissons, pour chaque $t \geqslant 0$ et $x \in \mathbb{R}^d$,

$$\bar{p}(t,x) := \frac{1}{(2\pi\bar{\sigma}^2 t)^{d/2}} \exp\left(-\frac{|x|^2}{2\,\bar{\sigma}^2 t}\right), \tag{42}$$

le noyau de chaleur avec une diffusivité $\bar{\sigma}$, alors, pour chaque temps $T > 0$, la convergence suivante est établie, presque sûrement sur l'événement $\{0 \in \mathscr{C}_\infty\}$,

$$\lim_{n\to\infty} \left| n^{d/2} p(nt, g_n^{\mathbf{a}}(x), 0) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x) \right| = 0, \tag{43}$$

uniformément dans la variable spatiale $x \in \mathbb{R}^d$ et dans la variable temporelle $t \geqslant T$, où la notation $g_n^{\mathbf{a}}(x)$ signifie le point le plus proche de $\sqrt{n}x$ sur l'amas infini sous l'environnement $\mathbf{a}$.

Comme outil important, la théorie des correcteurs sur $\mathscr{C}_\infty$ est également développée. Dans [45], le problème de régularité de Liouville dans une classe générale de graphes aléatoires est étudié par Benjamini, Duminil-Copin, Kozma et Yadin à l'aide de la méthode de l'entropie, qui confirme la dimension des correcteurs du premier ordre et donne la borne pour l'ordre supérieur. La description complète de la régularité de Liouville sur $\mathscr{C}_\infty$ est ensuite donnée par Armstrong et Dario dans [19] par la méthode d'homogénéisation quantitative. Dario donne également l'estimation optimale des correcteurs du même modèle dans [83]. Ces résultats nous fournissent des outils pour l'algorithme AHKM sur les grappes de percolation mentionné dans la section 0.2, et nous aident également à améliorer le TCL local asymptotique dans [41] en un résultat du TCL local quantitatif. Ce sont les principales contributions dans le chapitre 4 et seront résumées dans la section 0.3.1.

Enfin, avant d'énoncer notre contribution, nous remarquons qu'il existe d'autres développements du modèle de conductance aléatoire dans les directions suivantes : la relaxation de la condition i.i.d., le modèle sans condition d'ellipticité uniforme et permettant des queues à la fois près de ∞ et près de 0, la percolation avec conductance corrélée à longue portée, etc. Pour certains de ces modèles, il existe d'autres universalités (*la diffusion anormale*) que le cas gaussien. Nous renvoyons au chapitre 4.1.3 et aux références qui s'y trouvent pour une revue complète.

0.3.1 Résumé du chapitre 4

La principale contribution du chapitre 4 est le taux de convergence du TCL local pour le VSRW défini dans l'eq. (39). Dans les paragraphes suivants, nous présentons d'abord ce résultat, puis nous discutons des techniques développées pour sa preuve. Enfin, nous parlerons également de l'homogénéisation de la fonction de Green elliptique comme son corollaire.

Théorème 0.3.1 (le théorème 4.1.1)**.** *Pour chaque exposant $\delta > 0$, il existe une constante positive $C < \infty$ et un exposant $s > 0$, dépendant uniquement des paramètres $d, \lambda, \mathfrak{p}$ et δ, tels* *que pour chaque* $y \in \mathbb{Z}^d$*, il existe un temps aléatoire positif* $\mathcal{T}_{\mathrm{par},\delta}(y)$ *satisfaisant l'estimation d'intégrabilité stochastique*

$$\forall T \geqslant 0,\ \mathbb{P}\left(\mathcal{T}_{\mathrm{par},\delta}(y) \geqslant T\right) \leqslant C \exp\left(-\frac{T^s}{C}\right), \tag{44}$$

de sorte que, sur l'événement $\{y \in \infty\}$*, pour chaque* $x \in \mathscr{C}_\infty$ *et chaque* $t \geqslant \max\left(\mathcal{T}_{\mathrm{par},\delta}(y), |x-y|\right)$*,*

$$\left|p(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y)\right| \leqslant Ct^{-\frac{d}{2}-\left(\frac{1}{2}-\delta\right)} \exp\left(-\frac{|x-y|^2}{Ct}\right). \tag{45}$$

Nous avons plusieurs remarques à faire sur ce résultat.

- En général, le terme d'erreur a un autre facteur $t^{-\left(\frac{1}{2}-\delta\right)}$ devant la borne gaussienne, il est donc très petit en temps long comparé à la fois à $p(t,x,y)$ et à $\bar{p}(t,x-y)$. L'exposant est presque optimal car δ peut être arbitrairement petit et $t^{-\frac{1}{2}}$ est le taux optimal pour la marche aléatoire simple sur $\mathbb{Z}^d$.

- Dans l'eq. (45), il existe un facteur de normalisation $\theta(\mathfrak{p})^{-1}$. Ceci est nécessaire car le semigroupe $p(t,\cdot,y)$ ne charge que $\mathscr{C}_\infty$, et $\theta(\mathfrak{p})$ est presque la masse totale de $\bar{p}(t,\cdot - y)$ sur $\mathscr{C}_\infty$ par l'argument de densité

$$\int_{\mathscr{C}_\infty} \bar{p}(t,\cdot - y) \simeq \theta(\mathfrak{p}) \int_{\mathbb{Z}^d} \bar{p}(t,\cdot - y) \simeq \theta(\mathfrak{p}) \int_{\mathbb{R}^d} \bar{p}(t,\cdot - y) \simeq \theta(\mathfrak{p}).$$

- Le résultat dans l'eq. (45) est établi pour $t \geqslant \max\left(\mathcal{T}_{\mathrm{par},\delta}(y), |x-y|\right)$. Cette condition peut être décomposée comme suit

$$\left\{t \geqslant \max\left(\mathcal{T}_{\mathrm{par},\delta}(y), |x-y|\right)\right\} = \{t \geqslant |x-y|\} \cap \left\{t \geqslant \mathcal{T}_{\mathrm{par},\delta}(y)\right\}.$$

  Rappelons que $p(t,x,y)$ a une queue exponentielle pour $|x-y| \geqslant t$ au lieu d'une queue gaussienne, la comparaison n'est donc pas vraie dans ce régime. La condition $t \geqslant \mathcal{T}_{\mathrm{par},\delta}(y)$ peut être interprétée comme un temps d'attente aléatoire pour laisser le marcheur aléatoire explorer la percolation. Comme nous le savons, à petite échelle, la configuration de la percolation peut être assez zigzagante et fractale, de sorte que le semigroupe n'a pas convergé assez près de la gaussienne. De plus, le temps d'attente $\mathcal{T}_{\mathrm{par},\delta}(y)$ n'est pas très grand puisque d'après l'eq. (44) sa taille typique est une constante et a une queue sous-exponentielle.

La preuve de ce résultat repose sur la théorie de l'homogénéisation quantitative sur les amas de percolation, et nous développons également quelques nouvelles techniques. Voici une liste des ingrédients principaux.

1. *Une partition de bons cubes* : [19] a développé une partition des cubes de type Calderón-Zygmund, telle que

   (a) Il existe une collection $\mathcal{P}$ de cubes triadiques, $\mathbb{Z}^d = \bigsqcup_{\square \in \mathcal{P}} \square$.

   (b) Dans chaque *partition cube* $\square \in \mathcal{P}$, il existe un *amas maximal* $\mathscr{C}_*(\square)$. L'amas infini $\mathscr{C}_\infty$ a la structure $\mathscr{C}_\infty = \bigsqcup_{\square \in \mathcal{P}} \mathscr{C}_*(\square)$.

   (c) La taille (diamètre) du cube de partition a une estimation $\mathrm{size}(\square) \leqslant \mathcal{O}_1(C)$.

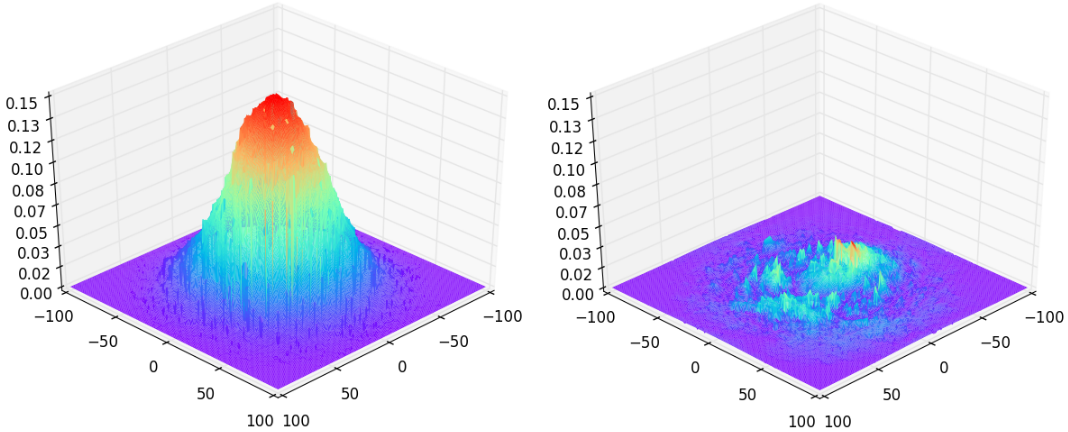

Figure 2: La figure de gauche représente la distribution de densité de la fonction $t^{\frac{d}{2}}p(t,\cdot,0)$ où la carte p est le noyau de chaleur bidimensionnel sur l'amas de percolation infini avec la probabilité $\mathfrak{p} = 0,7$ au temps $t = 1000$; elle est similaire à une distribution gaussienne. La figure de droite représente l'erreur entre la carte $t^{\frac{d}{2}}p(t,\cdot,0)$ et le noyau de chaleur gaussien normalisé $\theta(p)^{-1}t^{\frac{d}{2}}\bar{p}(t,\cdot)$ défini dans l'eq. (42) ; il est petit comparé à la distribution de densité à gauche.

Par conséquent, nous pouvons utiliser cette technique pour effectuer la localisation de $\mathscr{C}_\infty$ à chaque petit amas $\mathscr{C}_*(\square)$, et sa géométrie n'est pas très éloignée du cube $\square$ qui le contient. Ceci nous permet de développer les inégalités fonctionnelles dont l'inégalité de Poincaré et l'inégalité de Meyers sur les amas de percolation. La construction de cette partition de bons cubes est inspirée par le travail de Pisztora [201], et une idée similaire est également utilisée par Barlow dans la preuve de la borne gaussienne dans [39].

2. *L'estimation des correcteurs et l'expansion à deux échelles* : l'estimation optimale des correcteurs sur $\mathscr{C}_\infty$ est prouvée dans [83] et son application à l'expansion à deux échelles sur l'amas de percolation est implémentée dans [134]. (voir le chapitre 3) pour

$$-\nabla \cdot \mathbf{a}_{\mathscr{C}} \nabla u = -\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} \qquad \text{dans } \mathbb{Z}^d, \tag{46}$$

où $\mathbf{a}_{\mathscr{C}}(\{x,y\}) = \mathbf{a}(\{x,y\})\mathbf{1}_{\{x,y\in\mathscr{C}_\infty\}}$. Le développement à deux échelles pour l'eq. (46) est plutôt inhabituel, car son côté gauche est supporté sur $\mathscr{C}_\infty$, alors que le côté droit est supporté sur $\mathbb{Z}^d$. Nous avons donc besoin de l'argument de la partition des bons cubes. Dans la preuve du 0.3.1, nous traitons un cas similaire mais plus général

$$\begin{aligned} (\partial_t - \nabla \cdot \mathbf{a} \nabla) u = 0 \qquad & (0,\infty) \times \mathscr{C}_\infty, \\ \left(\partial_t - \frac{1}{2}\bar{\sigma}^2 \Delta\right) \bar{u} = 0 \qquad & (0,\infty) \times \mathbb{R}^d, \end{aligned} \tag{47}$$

avec une condition limite cohérente appropriée ; voir le théorème 4.3.2 pour plus de détails. Ici, la première équation de l'eq. (47) est définie sur $\mathscr{C}_\infty$ et $-\nabla \cdot \mathbf{a} \nabla$ est un opérateur de différence finie ; la deuxième équation de l'eq. (47) est définie sur $\mathbb{R}^d$ et Δ est l'opérateur de Laplace standard. Contrairement à l'eq. (46), nous n'avons pas de moyen canonique de combiner les deux équations en une seule. Pour cette raison, outre la technique de la partition de bon cube, nous appliquons également *la décomposition de Whitney* de l'analyse harmonique pour surmonter l'obstacle.

3. *L'estimation du flux : le flux centré* $\mathbf{g}_{e_k}$ défini par

$$\mathbf{g}_{e_k} : \mathbb{Z}^d \to \mathbb{R}^d, \qquad \mathbf{g}_{e_k} := \mathbf{a}_{\mathscr{C}}(\mathcal{D}\phi_{e_k} + e_k) - \bar{\mathbf{a}} e_k, \tag{48}$$

est également une quantité importante dans la théorie de l'homogénéisation quantitative. Son rôle et son estimation sont très similaires à ceux de $\nabla \phi_{e_k}$. Dans le chapitre 3, nous développons l'estimation de la norme faible pour $\mathbf{g}_{e_k}$, mais dans la preuve du chapitre 4, nous utilisons une autre quantité similaire $\widetilde{\mathbf{g}}_{e_k}$.

$$\widetilde{\mathbf{g}}_{e_k} : \mathscr{C}_\infty \to \mathbb{R}^d, \qquad \widetilde{\mathbf{g}}_{e_k} := \mathbf{a}(\mathcal{D}\phi_{e_k} + e_k) - \frac{1}{2}\bar{\sigma}^2 e_k, \tag{49}$$

et nous donnons son estimation de H^{-1} dans la proposition 4.B.1. Rappelons l'identité $\bar{\mathbf{a}} = \frac{1}{2}\theta(\mathfrak{p})\bar{\sigma}^2$, la principale différence entre l'eq. (48) et l'eq. (49) est en fait une différence de densité de l'amas. Comme nous avons besoin d'une estimation quantitative de $\widetilde{\mathbf{g}}_{e_k}$, nous prouvons une concentration de la densité d'amas dans la proposition 4.A.1

$$\left| \frac{|\mathscr{C}_\infty \cap \square_m|}{|\square_m|} - \theta(\mathfrak{p}) \right| \leqslant \mathcal{O}_{\frac{2(d-1)}{3d^2+2d-1}}\left(C 3^{-\frac{dm}{2}} \right). \tag{50}$$

Cette estimation est plus explicite que les résultats de grandes déviations qui étaient disponibles dans l'estimation.

Le TCL local 0.3.1 implique également l'homogénéisation quantitative pour la fonction de Green elliptique sur l'amas infini. En dimension $d \geqslant 3$, étant donné un environnement $\{\mathbf{a}(e)\}_{e \in E_d}$ et un point $y \in \mathscr{C}_\infty$, nous définissons la fonction de Green $g(\cdot, y)$ comme la solution de l'équation

$$-\nabla \cdot \mathbf{a} \nabla g(\cdot, y) = \delta_y \text{ dans } \mathscr{C}_\infty \text{ tel que } g(x,y) \overset{x \to \infty}{\longrightarrow} 0.$$

Cette fonction existe, est unique presque sûrement et est liée au semigroupe p par l'identité

$$g(x,y) = \int_0^\infty p(t,x,y)\, dt. \tag{51}$$

En dimension 2, la situation est différente puisque la fonction de Green n'est pas bornée à l'infini, et nous définissons $g(\cdot, y)$ comme l'unique fonction qui satisfasse

$$-\nabla \cdot \mathbf{a} \nabla g(\cdot, y) = \delta_y \text{ dans } \mathscr{C}_\infty, \ \frac{1}{|x|} g(x,y) \overset{x \to \infty}{\longrightarrow} 0 \text{ et } g(y,y) = 0.$$

Cette fonction est liée à la probabilité de transition p par l'identité

$$g(x,y) = \int_0^\infty \left(p(t,x,y) - p(t,y,y) \right) dt.$$

Dans l'énoncé ci-dessous, nous désignons par $\bar{g}$ la fonction de Green homogénéisée définie par la formule, pour chaque point $x \in \mathbb{R}^d \backslash \{0\}$,

$$\bar{g}(x) := \begin{cases} -\frac{1}{\pi \bar{\sigma}^2 \theta(\mathfrak{p})} \ln|x| & \text{if } d = 2, \\ \frac{\Gamma(d/2-1)}{(2\pi^{d/2}\bar{\sigma}^2\theta(\mathfrak{p}))} \frac{1}{|x|^{d-2}} & \text{if } d \geqslant 3, \end{cases} \tag{52}$$

où le symbole Γ désigne la fonction Gamma.

Théorème 0.3.2 (le théorème 4.1.2). *Pour chaque exposant $\delta > 0$, il existe une constante positive $C < \infty$ et un exposant $s > 0$, ne dépendant que des paramètres $d, \lambda, \mathfrak{p}$ et δ, tels que pour chaque $y \in \mathbb{Z}^d$, il existe une variable aléatoire non négative $\mathcal{M}_{\mathrm{ell},\delta}(y)$ satisfaisant à*

$$\forall R \geqslant 0,\ \mathbb{P}\left(\mathcal{M}_{\mathrm{ell},\delta}(y) \geqslant R\right) \leqslant C \exp\left(-\frac{R^s}{C}\right),$$

tel que, sur l'événement $\{y \in \mathscr{C}_\infty\}$:

1. *En dimension $d \geqslant 3$, pour tout point $x \in \mathscr{C}_\infty$ satisfaisant $|x-y| \geqslant \mathcal{M}_{\mathrm{ell},\delta}(y)$,*

$$|g(x,y) - \bar{g}(x-y)| \leqslant \frac{1}{|x-y|^{1-\delta}} \frac{C}{|x-y|^{d-2}}. \tag{53}$$

2. *En dimension 2, la limite*

$$K(y) := \lim_{x\to\infty} \left(g(x,y) - \bar{g}(x-y)\right),$$

existe, est finie presque sûrement et satisfait l'estimation d'intégrabilité stochastique

$$\forall R \geqslant 0,\ \mathbb{P}\left(|K(y)| \geqslant R\right) \leqslant C \exp\left(-\frac{R^s}{C}\right).$$

De plus, pour tout point $x \in \infty$ satisfaisant $|x-y| \geqslant \mathcal{M}_{\mathrm{ell},\delta}(y)$,

$$|g(x,y) - \bar{g}(x-y) - K(y)| \leqslant \frac{C}{|x-y|^{1-\delta}}. \tag{54}$$

0.4 Homogénéisation pour le système de particules en interaction

Un autre sujet étudié dans cette thèse est la théorie de l'homogénéisation pour les systèmes de particules en interaction, qui correspond aux chapitres 5 et 6 et est résumée dans la section 0.4.1. Comme le contexte est un peu différent de celui de l'homogénéisation classique, nous donnons d'abord un bref aperçu de quelques modèles de particules classiques pour rendre nos motivations plus claires.

Dans les modèles précédents, la marche aléatoire dans un environnement aléatoire, qui peut également être considérée comme l'évolution d'une particule, sera proche du mouvement brownien à grande échelle et à long terme. Les systèmes de particules en interaction partagent le même esprit, mais dans ces modèles nous avons une infinité de particules au lieu d'une, et l'environnement aléatoire provient de leur configuration qui est dynamique.

Le modèle le plus étudié est le gaz sur réseau et un modèle de base est *le processus d'exclusion symétrique simple (SSEP)* : soit $\eta : \mathbb{Z}^d \to \{0,1\}$ représente la configuration des particules, où chaque site permet au plus une particule. Dans l'évolution, chaque particule a un taux $\frac{1}{2}$ pour sauter vers un voisin vacant. Ainsi, l'évolution $(\eta_t)_{t\geqslant 0}$ suit le générateur

$$\mathcal{L}f(\eta) = \frac{1}{2}\sum_{x\in\mathbb{Z}^d}\sum_{y\sim x} \eta(x)(1-\eta(y))\left(f(\eta^{x,y}) - f(\eta)\right), \tag{55}$$

avec la notation

$$\eta^{x,y}(z) = \begin{cases} \eta(z) & z \neq x, y; \\ \eta(y) & z = x; \\ \eta(x) & z = y. \end{cases} \tag{56}$$

Ici, la fonction de test dans l'eq. (55) est la fonction locale $f \in C_0$ qui ne dépend que du site fini de η. Dans ce modèle, la mesure de Bernoulli produit $\mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}^d}$ avec $\alpha \in (0,1)$ est une mesure stationnaire et nous la désignons par $\mathbb{P}_\alpha$.

Le comportement à long terme et à grande échelle de la SSEP peut être caractérisé par *la limite hydrodynamique* et *la fluctuation d'équilibre*. Nous désignons par π_t^N la densité empirique de la configuration

$$\pi_t^N := N^{-d} \sum_{x \in \mathbb{Z}^d} \eta_{N^2 t}(x) \delta_{x/N}, \tag{57}$$

La limite hydrodynamique nous dit $(\pi_t^N)_{t \geqslant 0} \xrightarrow{N \to \infty} (\rho_t)_{t \geqslant 0}$, c'est-à-dire que la densité empirique converge vers la solution de l'équation de la chaleur

$$\partial_t \rho_t = \frac{1}{2} \Delta \rho_t, \tag{58}$$

dans la topologie de Skorokhod de la distribution de Schwartz, à condition que la configuration initiale $\pi_0^N \xrightarrow{N \to \infty} \rho_0$ ait un profil limite ρ_0. Si η_0 part de la mesure stationnaire $\mathbb{P}_\alpha$, alors le théorème de fluctuation de l'équilibre dit

$$Y_t^N := N^{-\frac{d}{2}} \sum_{x \in \mathbb{Z}^d} \left(\eta_{N^2 t}(x) - \alpha\right) \delta_{x/N}, \tag{59}$$

converge vers le processus fonctionnel d'Ornstein-Uhlenbeck $(Y_t)_{t \geqslant 0}$ résolvant

$$\mathrm{d}Y_t = \frac{1}{2} \Delta Y_t \, \mathrm{d}t + \sqrt{\alpha(1-\alpha)} \nabla \mathrm{d}B_t, \tag{60}$$

où B_t est le bruit blanc spatio-temporel.

Dans ces résultats, la matrice de coefficient effectif est l'identité, parce que le flux $W_{x,x+e_i}$ de x à $x + e_i$ dans le SSEP est

$$W_{x,x+e_i} = \frac{1}{2} \left(\eta(x) - \eta(x + e_i)\right), \tag{61}$$

et elle peut être écrite comme la différence $W_{x,x+e_i} = \tau_x h(\eta) - \tau_{x+e_i} h(\eta)$ avec $h(\eta) = \eta(0)$, où τ_x est l'opérateur de translation. Cette propriété est *la condition de gradient*. Elle rend le coefficient effectif trivial dans le modèle et elle n'est valable que dans certains systèmes de particules. Une autre explication plus heuristique est que le SSEP peut être traité comme si le saut était toujours autorisé, car dans l'eq. (55)

$$\sum_{x \in \mathbb{Z}^d} \sum_{y \sim x} \eta(x)(1 - \eta(y)) \left(f(\eta^{x,y}) - f(\eta)\right) = \sum_{x \in \mathbb{Z}^d} \sum_{y \sim x} \left(f(\eta^{x,y}) - f(\eta)\right).$$

Pour rendre ce modèle moins spécifique, *le processus d'exclusion symétrique généralisé (GSEP)* est proposé, où chaque site dans $\mathbb{Z}^d$ peut placer au plus κ particules ($\kappa \geqslant 2$), c'est-à-dire $\widetilde{\eta} : \mathbb{Z}^d \to \{0, 1, \cdots, \kappa\}$ et le générateur est

$$\mathcal{L} f(\widetilde{\eta}) = \frac{1}{2} \sum_{x \in \mathbb{Z}^d} \sum_{y \sim x} \mathbf{1}_{\{\widetilde{\eta}(x) > 0, \widetilde{\eta}(y) < \kappa\}} \left(f(\widetilde{\eta}^{x,y}) - f(\widetilde{\eta})\right), \tag{62}$$

avec la notation

$$\widetilde{\eta}^{x,y}(z) = \begin{cases} \widetilde{\eta}(z) & z \neq x, y; \\ \widetilde{\eta}(x) - 1 & z = x; \\ \widetilde{\eta}(y) + 1 & z = y. \end{cases} \tag{63}$$

Dans ce modèle, la mesure stationnaire est $\mathbb{P}_\alpha = \nu_\alpha^{\otimes \mathbb{Z}^d}$ avec

$$\forall n \in \{0, 1, \cdots, \kappa\}, \qquad \nu_\alpha(n) = \frac{\alpha^n}{\sum_{j=0}^{\kappa} \alpha^j}.$$

Ce modèle ne satisfait pas la condition de gradient, car

$$W_{x,x+e_i} = \frac{1}{2}\left(\mathbf{1}_{\{\widetilde{\eta}(x)>0, \widetilde{\eta}(x+e_i)<\kappa\}} - \mathbf{1}_{\{\widetilde{\eta}(x+e_i)>0, \widetilde{\eta}(x)<\kappa\}}\right), \tag{64}$$

ne peut pas être écrite comme la différence $W_{x,x+e_i} = \tau_x h(\widetilde{\eta}) - \tau_{x+e_i} h(\widetilde{\eta})$ pour une certaine fonction locale $h \in C_0$. Dans son résultat de limite hydrodynamique et théorème de fluctuation,

$$\partial_t \rho_t = \nabla \cdot D(\rho_t) \nabla \rho_t, \qquad \mathrm{d}Y_t = \nabla \cdot D(\alpha) \nabla Y_t \, \mathrm{d}t + \sqrt{\bar{\mathbf{a}}(\alpha)} \nabla \mathrm{d}W_t, \tag{65}$$

nous verrons une quantité appelée *le coefficient de diffusion globale* (ou *le coefficient d'auto-diffusion*) $D(\alpha)$ définie par

$$D(\alpha) := \frac{\bar{\mathbf{a}}(\alpha)}{2\chi(\alpha)}, \tag{66}$$

où χ est la quantité appelée *la compressibilité*

$$\chi(\alpha) := \mathbb{V}\mathrm{ar}_\alpha[\widetilde{\eta}(0)], \tag{67}$$

et la quantité $\bar{\mathbf{a}}$ a une description variationnelle. Nous désignons par $\Gamma_f(\widetilde{\eta}) := \sum_{x \in \mathbb{Z}^d} \tau_x f(\widetilde{\eta})$, qui peut être infinie, mais pour toute fonction locale $f \in C_0$, on peut dire que

$$\nabla_{0,e_i} \Gamma_f(\widetilde{\eta}) := \Gamma_f(\widetilde{\eta}^{0,e_i}) - \Gamma_f(\widetilde{\eta}), \tag{68}$$

est bien défini car le saut ne change que la valeur des termes finis dans $\sum_{x \in \mathbb{Z}^d} \tau_x f(\widetilde{\eta})$. Alors $\bar{\mathbf{a}}$ est défini par

$$p \cdot \bar{\mathbf{a}}(\alpha) p = \inf_{f \in C_0} \sum_{i=1}^{d} \mathbb{E}_\alpha\left[\mathbf{1}_{\{\widetilde{\eta}(0)>0, \widetilde{\eta}(e_i)<\kappa\}} (p_i + \nabla_{0,e_i} \Gamma_f(\widetilde{\eta}))^2\right]. \tag{69}$$

Dans la définition de la matrice de diffusion, la quantité $\bar{\mathbf{a}}$ ressemble beaucoup au coefficient effectif dans l'homogénéisation stochastique, où la formule l'eq. (68) est utilisée pour construire un champ de gradient stationnaire. Il est donc très naturel de penser que nous pouvons utiliser une approximation de volume fini dans l'eq. (19) et obtenir son taux de convergence pour $D(\alpha)$. Cela peut nous fournir des résultats quantitatifs dans les systèmes de particules. Cependant, nous devons remarquer que l'eq. (68) est défini pour l'espace de configuration, donc la fonction $\nabla_{0,e_i} \Gamma_f(\widetilde{\eta})$ peut avoir un nombre arbitraire de coordonnées. C'est l'un des principaux défis et nous en discuterons en détail dans la section 0.4.1.

Nous donnons également un bref aperçu des références des résultats mentionnés ci-dessus. Il existe deux approches classiques pour l'identification de la limite hydrodynamique. La première, appelée *la méthode d'entropie*, a été introduite dans [136], et étendue à certains modèles non-gradients dans [221, 204]. La deuxième, appelée *la méthode d'entropie relative*, a été introduite dans [224], et a été étendue à un modèle non-gradient dans [111].

La description asymptotique des fluctuations des systèmes de particules en interaction à l'équilibre a été obtenue dans [66, 214, 91, 69, 71], où l'outil principal est *le théorème de Holley-Strook* [140]. L'extension de ce résultat aux modèles non-gradients a été obtenue dans [174, 70, 110]. Nous n'avons pas connaissance de résultats concernant les fluctuations de non-équilibre d'un modèle non-gradient. Pour les modèles à gradient (ou leurs petites perturbations), nous nous référons en particulier à [202, 90, 106, 71, 144].

Le travail [166] donne une preuve que les approximations en volume fini de la matrice d'auto-diffusion convergent vers la limite correcte. Cependant, aucun taux de convergence n'a pu y être obtenu. Le résultat qualitatif de [166] a été étendu au processus d'exclusion simple à moyenne nulle, et au processus d'exclusion simple asymétrique en dimension $d \geqslant 3$, dans [143]. Enfin, nous nous référons également aux livres [215, 152, 157] pour des expositions beaucoup plus approfondies sur ces sujets, et des revues de la littérature.

0.4.1 Résumé des chapitres 5 et 6

Dans les chapitres 5 et 6, nous visons à développer une théorie d'homogénéisation quantitative pour les systèmes de particules en interaction de type non-gradient. Nos principales contributions sont une décroissance de type gaussien de $t^{-\frac{d}{2}}$ pour le semigroupe, voir le théorème 0.4.1, et un taux de convergence pour l'approximation en volume fini du coefficient de masse, voir le théorème 0.4.2. Notre modèle est construit dans *l'espace de configuration du continuum*, mais les résultats et les preuves peuvent être adaptés dans le modèle classique de gaz sur réseau de type non-gradient, par exemple le GSEP.

Nous présentons d'abord notre système de particules. Soit $\mathcal{M}_\delta(\mathbb{R}^d)$ l'ensemble des mesures σ-finies qui sont des sommes de masses de Dirac sur $\mathbb{R}^d$, que nous considérons comme l'espace des configurations de particules. Nous désignons par $\mathbb{P}_\rho$ la loi sur $\mathcal{M}_\delta(\mathbb{R}^d)$ du processus ponctuel de Poisson de densité $\rho \in (0, \infty)$, avec $\mathbb{E}_\rho, \mathbb{V}\mathrm{ar}_\rho$ l'espérance et la variance associées. On désigne par $\mathcal{F}_U$ le tribu générée par les applications $V \mapsto \mu(V)$, pour tous les ensembles boréliens $V \subseteq U$, complétés par tous les ensembles négligeables, et on fixe $\mathcal{F} := \mathcal{F}_{\mathbb{R}^d}$. Nous nous donnons une fonction $\mathbf{a}_\circ : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, où $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ est l'ensemble des matrices symétriques $d \times d$. Nous supposons que cette correspondance satisfait aux propriétés suivantes :

- *l'ellipticité uniforme* : il existe $\Lambda < \infty$ tel que pour chaque $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$,
$$\forall \xi \in \mathbb{R}^d, \quad |\xi|^2 \leqslant \xi \cdot \mathbf{a}_\circ(\mu)\xi \leqslant \Lambda|\xi|^2 \,; \tag{70}$$
- *une dépendance de portée finie* : en désignant par B_1 la boule euclidienne de rayon 1 centrée à l'origine, on suppose que $\mathbf{a}_\circ$ est $\mathcal{F}_{B_1}$- mesurable.

Nous désignons par $\tau_{-x}\mu$ la translation de la mesure μ par le vecteur $-x \in \mathbb{R}^d$; explicitement, pour tout ensemble borélien U, nous avons $(\tau_{-x}\mu)(U) = \mu(x + U)$. Nous étendons $\mathbf{a}_\circ$ par stationnarité en posant, pour tout $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$ et $x \in \mathbb{R}^d$,

$$\mathbf{a}(\mu, x) := \mathbf{a}_\circ(\tau_{-x}\mu).$$

En désignant par $\mu_t := \sum_{i=1}^{\infty} \delta_{x_{i,t}}$ la configuration au temps $t \geqslant 0$, notre modèle peut être décrit de manière informelle comme un système infiniment dimensionnel avec interaction locale tel que chaque particule $x_{i,t}$ évolue comme une diffusion associée à l'opérateur de forme divergente $-\nabla\cdot\mathbf{a}(\mu_t, x_{i,t})\nabla$. Plus précisément, c'est un processus de Markov $(\Omega, (\mathscr{F}_t)_{t\geqslant 0}, \mathbb{P}_\rho)$ défini par *la forme de Dirichlet*

$$\mathcal{E}^{\mathbf{a}}(f,f) := \mathbb{E}_\rho\left[\int_{\mathbb{R}^d} \nabla f(\mu,x)\cdot\mathbf{a}(\mu,x)\nabla f(\mu,x)\,\mathrm{d}\mu(x)\right], \tag{71}$$

où la dérivée directionnelle

$$e_k\cdot\nabla f(\mu,x) = \lim_{h\to 0}\frac{f(\mu-\delta_x+\delta_{x+he_k})-f(\mu)}{h}, \tag{72}$$

est défini pour une famille de fonctions appropriées et $x\in\operatorname{supp}(\mu)$. La construction de processus de diffusion similaires peut être trouvée dans les travaux précédents d'Albeverio, Kondratiev et Röckner dans [2, 3, 4, 5] ; voir aussi l'étude [206].

Nous avons besoin de quelques explications supplémentaires pour la fonction de test de la forme de Dirichlet dans l'eq. (71). Pour tout ensemble ouvert $U\subseteq\mathbb{R}^d$, nous désignons par l'espace $\mathscr{C}_c^\infty(U)$ les fonctions qui sont $\mathcal{F}_K$-mesurables pour un ensemble compact $K\subseteq U$, et lisses par rapport à toute particule. Par conséquent, l'espace de fonctions $\mathscr{C}_c^\infty(U)$ joue le même rôle que la fonction locale dans le modèle de gaz sur réseau. Ensuite, nous définissons la norme $\mathscr{H}^1(U)$, une analogie de l'espace de Sobolev classique H^1 en dimension infinie

$$\|f\|_{\mathscr{H}^1(U)} = \left(\mathbb{E}_\rho[f^2(\mu)] + \mathbb{E}_\rho\left[\int_U |\nabla f(\mu,x)|^2\,\mathrm{d}\mu(x)\right]\right)^{\frac{1}{2}}. \tag{73}$$

Nous définissons également l'espace $\mathscr{H}_0^1(U)$ comme la fermeture dans $\mathscr{H}^1(U)$ des fonctions $f\in\mathscr{C}_c^\infty(U)$ telles que $\|f\|_{\mathscr{H}^1(U)}$ est fini et c'est l'espace de fonctions pour l'eq. (71).

Le théorème principal du chapitre 5 est une estimation de la décroissance de la variance pour notre système de particules $(\mu_t)_{t\geqslant 0}$. Nous désignons par $\mathscr{L}^p$ l'espace L^p dans $(\Omega,\mathcal{F},\mathbb{P}_\rho)$ pour $p\geqslant 1$. Soit $u:\mathcal{M}_\delta(\mathbb{R}^d)\to\mathbb{R}$ une fonction $\mathcal{F}_{Q_{l_u}}$ mesurable avec $Q_{l_u} := \left(-\frac{l_u}{2},\frac{l_u}{2}\right)^d$, et soit $u_t(\mu_0) := \mathbb{E}_\rho[u(\mu_t)|\mu_0]$.

Théorème 0.4.1 (Le théorème principal dans le chapitre 5)**.** *Il existe deux constantes positives finies $\gamma := \gamma(\rho,d,\Lambda)$, $C := C(\rho,d,\Lambda)$ telles que pour tout $u\in\mathscr{C}_c^\infty(\mathbb{R}^d)\cap\mathscr{L}^\infty$ qui est $\mathcal{F}_{Q_{l_u}}$-mesurable, alors nous avons*

$$\mathbb{V}\mathrm{ar}_\rho[u_t] \leqslant C(1+|\log t|)^\gamma\left(\frac{1+l_u}{\sqrt{t}}\right)^d \|u\|^2_{\mathscr{L}^\infty}. \tag{74}$$

Remark. Dans les travaux [2, 3, 4, 5] d'Albeverio, Kondratiev et Röckner, la forme de Dirichlet $\mathcal{E}^{\mathbf{a}}$ est définie sur l'espace de fonctions

$$\mathcal{F}\mathscr{C}_c^\infty(\mathbb{R}^d) = \left\{G(\mu(g_1),\cdots\mu(g_n)) : n\in\mathbb{N}, G\in C_b^\infty(\mathbb{R}^{nd}), g_i\in C_c^\infty(\mathbb{R}^d)\right\}. \tag{75}$$

C'est un sous-espace concret de $\mathscr{C}_c^\infty(\mathbb{R}^d)$ et on peut prouver que $\mathscr{C}_c^\infty(\mathbb{R}^d)$ et $\mathcal{F}\mathscr{C}_c^\infty(\mathbb{R}^d)$ génèrent le même $\mathscr{H}_0^1(\mathbb{R}^d)$.

La preuve du théorème 0.4.1 s'inspire d'un travail important [142] de Janvresse, Landim, Quastel et Yau, où la décroissance de la variance est prouvée dans *le modèle de zéro-range*. Nous étendons cette preuve au modèle non-gradient dans l'espace de configuration du continuum, et une difficulté technique dans cette généralisation est une estimation de localisation clé : nous désignons par $\overline{Q}_K = \left[-\frac{K}{2},\frac{K}{2}\right]^d$ le cube fermé et rappelons que $\mathcal{F}_{\overline{Q}_K}$ représente l'information de μ dans celui-ci. Nous définissons $\mathsf{A}_K u_t := \mathbb{E}_\rho[u_t|\mathcal{F}_{\overline{Q}_K}]$, alors pour chaque $t \geqslant \max\left\{(l_u)^2, 16\Lambda^2\right\}$ et $K \geqslant \sqrt{t}$ nous avons

$$\mathbb{E}_\rho\left[(u_t - \mathsf{A}_K u_t)^2\right] \leqslant C(\Lambda)\exp\left(-\frac{K}{\sqrt{t}}\right)\mathbb{E}_\rho\left[u^2\right]. \tag{76}$$

Il s'agit d'une estimation clé apparaissant dans [142, la proposition 3.1], et elle est également naturelle puisque $\sqrt{t}$ est l'échelle typique de la diffusion, donc lorsque $K \gg \sqrt{t}$ on obtient une très bonne approximation dans la l'eq. (76). L'idée principale de sa preuve est de définir une fonctionnelle multi-échelle

$$\begin{aligned} S_{k,K,\beta}(f) &:= \alpha_k \mathbb{E}_\rho\left[(\mathsf{A}_k f)^2\right] + \int_k^K \alpha_s \, d\mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right] + \alpha_K \mathbb{E}_\rho\left[(f - \mathsf{A}_K f)^2\right] \\ &= \alpha_K \mathbb{E}_\rho\left[f^2\right] - \int_k^K \alpha'_s \mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right] ds, \end{aligned}$$

avec $\alpha_s = \exp\left(\frac{s}{\beta}\right), \alpha'_s = \frac{\mathrm{d}}{\mathrm{d}s}\alpha_s, \beta > 0$, puis étudie l'évolution de $\frac{\mathrm{d}}{\mathrm{d}t} S_{k,K,\beta}(u_t)$. Dans cette procédure, on peut affirmer que

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right] = 2\mathcal{E}^{\mathbf{a}}(u_t, \mathsf{A}_s u_t),$$

mais $\mathsf{A}_s u_t$ n'est pas dans la fonction de test $\mathscr{H}_0^1(\mathbb{R}^d)$ à cause de la perturbation à ∂Q_s. Plus précisément, cela signifie la discontinuité de $\mathsf{A}_s u_t$ lorsqu'une particule entre ou sort de Q_s. Dans le modèle discret, il existe également une telle perturbation à la frontière, mais $\mathsf{A}_s u_t$ peut toujours être utilisé comme fonction locale grâce à la différence discrète. Pour résoudre ce problème, nous utilisons une version de régularisation de A_s.

$$\mathsf{A}_{s,\varepsilon} f := \frac{1}{\varepsilon}\int_0^\varepsilon \mathsf{A}_{s+r} f \, \mathrm{d}r, \tag{77}$$

pour rendre l'espérance conditionnelle plus lisse. De plus, la dérivée de $\mathsf{A}_{s,\varepsilon} u_t$ près de la frontière est étroitement liée à l'isométrie $\mathscr{L}^2$ de la martingale $(\mathsf{A}_s u_t)_{s \geqslant 0}$.

Enfin, remarquons que [142] obtient également la limite à long terme $\mathbb{V}\mathrm{ar}_\rho[u_t] = Ct^{-\frac{d}{2}} + o(t^{-\frac{d}{2}})$. Cependant, le modèle de zéro-range a la condition de gradient, donc la constante C est plus facile à calculer. Dans notre modèle, nous devons d'abord identifier le coefficient effectif, ce qui nous motive également pour le travail dans le chapitre 6.

Dans le chapitre 6, nous étudions l'approximation par volumes finis du coefficient de diffusion. Pour tout ensemble ouvert borné $U \subseteq \mathbb{R}^d$, nous définissons la matrice $\bar{\mathbf{a}}(U) \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ comme étant telle que, pour tout $p \in \mathbb{R}^d$,

$$\frac{1}{2} p \cdot \bar{\mathbf{a}}(U) p := \inf_{\phi \in \mathscr{H}_0^1(U)} \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}(p + \nabla\phi(\mu, x)) \cdot \mathbf{a}(\mu, x)(p + \nabla\phi(\mu, x)) \, \mathrm{d}\mu(x)\right]. \tag{78}$$

Pour chaque $m \in \mathbb{N}$, nous désignons par $\square_m = Q_{3^m}$ le cube de longueur de côté 3^m. Nous définissons *la matrice de coefficient effectif* comme $\bar{\mathbf{a}} := \lim_{m\to\infty} \bar{\mathbf{a}}(\square_m)$, et le théorème principal consiste à prouver son taux de convergence.

Théorème 0.4.2 (Le théorème principal dans le chapitre 6)**.** *La limite $\bar{\mathbf{a}}$ est bien définie. De plus, il existe un exposant $\alpha(d, \Lambda, \rho) > 0$ et une constante $C(d, \Lambda, \rho) < \infty$ tels que pour tout $m \in \mathbb{N}$,*

$$|\bar{\mathbf{a}}(\square_m) - \bar{\mathbf{a}}| \leqslant C 3^{-\alpha m}. \tag{79}$$

La preuve du théorème 0.4.2 suit l'approche de renormalisation initiée par Armstrong et Smart dans [31], qui est également revue dans la section 0.1. Cependant, remarquons que l'espace des fonctions $\mathscr{H}_0^1(U)$ dans l'eq. (78) est très différent du cas euclidien : la fonction est définie dans la configuration $\mu = \sum_{i=1}^{\infty} \delta_{x_i}$ au lieu de $\mathbb{R}^d$, donc son nombre de coordonnées peut être arbitrairement grand. Dans les paragraphes suivants, nous soulignons quelques nouvelles idées lorsque nous mettons en œuvre l'approche de renormalisation dans les systèmes de particules.

1. *L'espace de fonctions pour les quantités sous-additives* : nous désignons par $\nu(U,p) = \frac{1}{2}p \cdot \bar{\mathbf{a}}(U)p$ et nous pouvons vérifier qu'elle est sous-additive. Nous espérons également construire une quantité duale sous-additive $\nu^*(U,q)$, et nous proposons la formule (voir la discussion dans l'eq. (20))

$$\nu^*(U,q) := \sup_{u \in \mathscr{H}^1(U)} \mathbb{E}_\rho \left[\frac{1}{\rho|U|} \int_U \left(-\frac{1}{2} \nabla u \cdot \mathbf{a} \nabla u + q \cdot \nabla u \right) \mathrm{d}\mu \right]. \tag{80}$$

   Cependant, remarquons que nous n'avons pas défini l'espace de fonctions $\mathscr{H}^1(U)$, bien que la norme soit définie dans l'eq. (73). Une bonne définition devrait être la plus grande classe de fonctions mesurables $\mathcal{F}$ avec une norme finie $\mathscr{H}^1(U)$. De manière informelle, ces fonctions sont différentiables par rapport aux particules dans U, mais la dépendance des particules en dehors de U est juste mesurable. En particulier, contrairement à l'espace de fonctions $\mathscr{H}_0^1(U)$, nous n'avons *pas* besoin de la condition $\mathcal{F}_U$-measurable pour $\mathscr{H}^1(U)$. Pour voir que cette définition est bonne, on peut vérifier :

   (a) Pour tout $V \subseteq U$, on a $\mathscr{H}_0^1(V) \subseteq \mathscr{H}_0^1(U)$ et $\mathscr{H}^1(U) \subseteq \mathscr{H}^1(V)$. C'est la propriété permettant de prouver la sous-additivité de ν et ν^*.

   (b) Soit $B_1(U)$ le voisinage contenant U avec une distance de 1. Alors $\mathbb{E}_\rho[u|\mathcal{F}_{B_1(U)}]$ est un candidat meilleur que u dans la fonctionnelle de $\nu^*(U,q)$. Ceci nous permet de retrouver la condition de mélange du maximiseur de $\nu^*(U,q)$.

2. *L'inégalité de Caccioppoli modifiée* : un autre ingrédient important est l'inégalité de Caccioppoli, car les optimiseurs de ν et ν^* sont des fonctions $\mathbf{a}$-harmoniques. Nous rappelons l'inégalité classique de Caccioppoli : pour chaque $\widetilde{u}$ tel que $\Delta \widetilde{u} = 0$ dans Q_{3r},

$$\int_{Q_r} |\nabla \widetilde{u}|^2 \leqslant \frac{C}{r^2} \int_{Q_{3r}} |\widetilde{u}|^2. \tag{81}$$

   Sa preuve consiste à utiliser une fonction de coupure $\psi \in C_c^\infty(Q_{3r})$ telle que $\psi^2 \widetilde{u} \in H_0^1(Q_{3r})$ et à tester ensuite $\psi^2 \widetilde{u}$ contre $\Delta \widetilde{u}$. Dans notre système de particules, la fonction $\mathbf{a}$-harmonique est

$$\mathcal{A}(U) := \left\{ u \in \mathscr{H}^1(U) : \ \forall \varphi \in \mathscr{H}_0^1(U), \ \mathbb{E}_\rho \left[\int_U \nabla u \cdot \mathbf{a} \nabla \varphi \, \mathrm{d}\mu \right] = 0 \right\},$$

   et nous espérons prouver un résultat similaire à l'eq. (81). Il n'existe pas d' analogie de la fonction de troncature ψ, mais en s'inspirant de l'eq. (77), pour tout $u \in \mathcal{A}(Q_{3r})$, nous pouvons utiliser $\mathsf{A}_{r,\varepsilon} u \in \mathscr{H}_0^1(Q_r)$ comme fonction de test. Cependant, malgré de nombreux efforts, le meilleur que nous puissions prouver est une inégalité de Caccioppoli modifiée dans la proposition 6.3.6 : il existe $\theta(d,\Lambda) \in (0,1)$, $C(d,\Lambda) < \infty$, et

$R_0(d,\Lambda) < \infty$ tels que pour chaque $r \geqslant R_0$ et $u \in \mathcal{A}(Q_{3r})$, nous avons

$$\mathbb{E}_\rho\left[\frac{1}{\rho|Q_r|}\int_{Q_r} \nabla(\mathsf{A}_{r+2}u)\cdot\mathbf{a}\nabla(\mathsf{A}_{r+2}u)\,\mathrm{d}\mu\right] \leqslant \frac{C}{r^2\rho|Q_{3r}|}\mathbb{E}_\rho[u^2] + \theta\mathbb{E}_\rho\left[\frac{1}{\rho|Q_{3r}|}\int_{Q_{3r}} \nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]. \quad (82)$$

L'inégalité dans l'eq. (82) contrôle la norme du gradient d'une fonction **a**-harmonique dans le petit cube Q_r par une somme de termes impliquant la norme du gradient dans le grand cube Q_{3r}. À première vue, cela ne semble pas utile. Cependant, le point essentiel est que le facteur multiplicatif θ est inférieur à un. Cela implique donc également une meilleure régularité à l'intérieur, et l'eq. (82) peut enfin être intégré dans le cadre de l'approche de renormalisation.

3. *L'inégalité de Poincaré dimension-libre* : l'inégalité de Poincaré est un outil nécessaire pour l'analyse, et nous l'établissons également dans $\mathscr{H}^1(U)$ et $\mathscr{H}^1_0(U)$. Pour l'espace de fonctions $\mathscr{H}^1(U)$, sa preuve repose sur l'inégalité d'Efron-Stein, et nous l'améliorons également en l'inégalité de Poincar'e multi-échelle. Pour l'espace de fonctions $\mathscr{H}^1_0(U)$, notre preuve fait implicitement appel au calcul de Malliavin sur l'espace de Poisson. Voir la section 6.3.1 pour plus de détails.

Chapter 1

Introduction

This thesis consists of the research work [133, 134, 85, 135, 115] during my Ph.D. and studies the interplay between quantitative homogenization theory and two stochastic models: *the supercritical percolation model* and *the infinite interacting particle systems*. A basic object of the stochastic homogenization theory is to understand the equation

$$-\nabla \cdot (\mathbf{a} \nabla u) = f \qquad \text{in } B_r, \tag{1.1}$$

with $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^{d\times d}_{sym}$ a symmetric, random $\mathbb{Z}^d$-stationary ergodic coefficient satisfying the uniform ellipticity $|\xi|^2 \leqslant \xi \cdot \mathbf{a}\xi \leqslant \Lambda|\xi|^2$, and where B_r is the Euclidean ball of radius r centered at the origin. For r very large, its solution can be approximated by *the effective solution* $\bar{u}$ satisfying

$$-\nabla \cdot (\bar{\mathbf{a}} \nabla \bar{u}) = f \qquad \text{in } B_r, \tag{1.2}$$

with the same boundary condition. Here $\bar{\mathbf{a}}$ is called *the effective coefficient*, which is a constant matrix. The quantity $\bar{\mathbf{a}}$ not only characterizes the large-scale asymptotic behavior of the elliptic problem, but also captures the long-time, large-scale behavior of the parabolic problem. This is the link between homogenization and various diffusion models in probability, where the quantitative homogenization theory provides tools for estimates. Reciprocally, the two models studied in this thesis bring new techniques back to homogenization theory: the percolation model goes beyond the uniform ellipticity setting, while particle systems require infinite-dimensional PDE analysis and the environment is dynamic.

Despite these applications, we should keep in mind one important motivation of the homogenization at the very beginning: an efficient numerical approximation. In fact, the solution in eq. (1.1) is very costly to compute numerically for a large r if one solves it naively with the finite difference algorithm, while $\bar{u}$ in eq. (1.2) can be computed quickly as the constant coefficient provides very high regularity. However, for a fixed r, there is always a discrepancy between the real solution u and the effective solution $\bar{u}$. Recently, Armstrong, Hannukainen, Kuusi and Mourrat proposed a new iterative algorithm (AHKM) which can approximate u with arbitrary precision in H^1, and the cost is close to that of computing $\bar{u}$. In Chapter 2, we will introduce this algorithm and prove its numerical consistency.

In Chapter 3, the AHKM iterative algorithm is applied to the $\mathbb{Z}^d$-Bernoulli percolation clusters ($d \geqslant 2$), which is a fundamental model of a perforated medium. More precisely, we sample i.i.d. Bernoulli random variables with parameter $\mathfrak{p} \in (\mathfrak{p}_c, 1]$ where $\mathfrak{p}_c$ is the critical point and $\mathfrak{p} > \mathfrak{p}_c$ ensures a unique infinite cluster $\mathscr{C}_\infty$. Then we study the Dirichlet problem eq. (1.1) on the $\mathscr{C}_\infty$-like maximal cluster in a big box. As the uniform ellipticity condition is no longer satisfied and the geometry of the cluster is fractal, the analysis becomes more challenging. Quantitative homogenization theory on percolation is initiated by Armstrong and Dario and one major technique is a Calderón-Zygmund-like decomposition of the space. Based on these results and techniques, we prove a rigorous numerical method to obtain an efficient approximation of both the potential u and the gradient ∇u.

Chapter 4 focuses on the parabolic Green function on the infinite percolation cluster $\mathscr{C}_\infty$, i.e. $p(\cdot,\cdot,y) : \mathbb{R}^+ \times \mathscr{C}_\infty \to [0,1]$ solving

$$\begin{cases} \partial_t p(\cdot,\cdot,y) - \nabla \cdot \mathbf{a} \nabla p(\cdot,\cdot,y) = 0 & \text{in } (0,\infty) \times \mathscr{C}_\infty, \\ p(0,\cdot,y) = \delta_y & \text{in } \mathscr{C}_\infty, \end{cases} \tag{1.3}$$

which is the transition probability of the jump process starting from $y \in \mathscr{C}_\infty$ associated to the generator $\nabla \cdot \mathbf{a} \nabla$. This topic is much studied by many pioneers and the results like the Gaussian bound, the invariance principle, and the asymptotic local central limit theorem have been proved. All these results tell us that $p(t,\cdot,y)$ is close to a Gaussian density for t large. With the collaboration of Dario, we go one step further to prove a near optimal rate of convergence, which can be interpreted as a quantitative central limit theorem. The proof makes use of several results in the previous work of Armstrong and Dario, and also the estimate of flux proved in Chapter 3.

Although the random walk on the infinite percolation cluster is complicated, it can still be seen as the diffusion of a particle in a static random environment. In Chapters 5 and 6, we turn to interacting particle systems, where the environment is dynamic and there are infinitely many particles instead of one. Our model can be thought as *the generalized symmetric exclusion process* in continuum space. It does not satisfy *the gradient condition*, and one has to lift the function space defined on the configuration of particles $\sum_{i=1}^{\infty} \delta_{x_i}$. In Chapter 5, we prove a bound for the relaxation to equilibrium of type $t^{-\frac{d}{2}}$.

In order to describe the long-time asymptotic behavior of this cloud of particles, one needs to identify the *bulk coefficient* $\bar{\mathbf{a}}$, which is the counterpart of the effective coefficient for particle systems. Chapter 6 will present a joint work with Giunti and Mourrat about the finite volume approximation of the bulk coefficient $\bar{\mathbf{a}}$. We remark that for elliptic equations, understanding the convergence of the finite volume approximation of $\bar{\mathbf{a}}$ is the cornerstone of the quantitative homogenization theory if one adopts the renormalization approach by Armstrong, Kuusi, Mourrat and Smart. Our contribution is to generalize this method to infinite-dimensional analysis, with several dimension-free functional inequalities (*the multiscale Poincaré inequality*, *the Caccioppoli inequality* etc.). The function spaces to study particle systems are quite different from the classical function spaces on $\mathbb{R}^d$.

The rest of Chapter 1 is organized as follows. In Section 1.1, we will give an overview of homogenization theory, especially the key quantitative method used throughout the thesis. Then in Section 1.2 we will state the details of the homogenization in numerical algorithms, and the main result in Chapters 2 and 3 concerning the AHKM algorithm. We review the percolation model in Section 1.3 and then present our contribution in Chapter 4. Section 1.4 aims to introduce the results of Chapters 5 and 6, where we will recall some classical results in particle systems. Finally, we end this chapter by discussions about the possible future directions in Section 1.5.

Table of notations

Homogenization on $\mathbb{R}^d$

$\mathbf{a}$:	$\mathbb{R}^{d\times d}_{sym}$-valued coefficient, which is $\mathbb{Z}^d$-periodic or stationary, with the uniform ellipticity $\lvert\xi\rvert^2 \leqslant \xi \cdot \mathbf{a}\xi \leqslant \Lambda\lvert\xi\rvert^2$;
$\bar{\mathbf{a}}$:	the effective coefficient, which is a constant matrix;
$\square_m$:	triadic cube $\left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d$;
Q_r :	general cube with side length r;
$\mathcal{O}_s$:	$\theta, s \in (0,\infty)$, random variable $X \leqslant \mathcal{O}_s(\theta) \iff \mathbb{E}\left[\exp((\theta^{-1}X)_+^s)\right] \leqslant 2$;
$\underline{L}^p(U)$:	weighted norm $\lVert f\rVert_{\underline{L}^p(U)} = \left(\frac{1}{\lvert U\rvert}\int_U \lvert f(x)\rvert^p\,dx\right)^{\frac{1}{p}}$;
$\underline{H}^k(U)$:	weighted norm $\lVert f\rVert_{\underline{H}^k(U)} := \sum_{0\leqslant\lvert\beta\rvert\leqslant k} \lvert U\rvert^{\frac{\lvert\beta\rvert-k}{d}} \lVert \partial^\beta f\rVert_{\underline{L}^2(U)}$.

The infinite percolation cluster

$\mathbf{a}$:	random conductance valued in $\{0\}\cup\{\lambda,1\}$, $(\mathbf{a}(e))_{e\in E_d}$ i.i.d.;
$\mathscr{C}_\infty$:	the infinite cluster of supercritical percolation;
$\square_{\mathcal{P}}(z)$:	the minimal partition cube that contains z;
$\mathscr{C}_*(\square)$:	the maximal cluster in the good cube $\square$;
$\square_m$:	triadic cube in $\mathbb{Z}^d$, i.e. $\square_m = \mathbb{Z}^d \bigcap \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d$;
ϕ_p :	the first-order corrector on $\mathscr{C}_\infty$, i.e. $-\nabla\cdot\mathbf{a}(p+\nabla\phi_p) = 0$;
$\mathbf{g}_p$:	the centered flux on $\mathbb{Z}^d$ i.e. $\mathbf{g}_p = \mathbf{a}_{\mathscr{C}}(\mathcal{D}\phi_p + p) - \bar{\mathbf{a}}p$, where $\mathbf{a}_{\mathscr{C}}(\{x,y\}) = \mathbf{a}(\{x,y\})\mathbf{1}_{\{x,y\in\mathscr{C}_\infty\}}$;
$\widetilde{\mathbf{g}}_p$:	the centered flux on $\mathscr{C}_\infty$ i.e. $\mathbf{g}_p = \mathbf{a}(\mathcal{D}\phi_p + p) - \frac{1}{2}\bar{\sigma}^2 p$;
$\mathcal{D}$:	finite difference $\mathcal{D}_h u(x) := u(x+h) - u(x)$, discrete gradient $\mathcal{D}u(x) := (\mathcal{D}_{e_1}u(x), \mathcal{D}_{e_2}u(x), \cdots, \mathcal{D}_{e_d}u(x))$;
∇ :	$x, y \in \mathbb{Z}^d, y \sim x, \nabla u(x,y) = u(y) - u(x)$;
$\nabla\cdot\mathbf{a}\nabla$:	divergence-form on $\mathbb{Z}^d$, $\nabla\cdot\mathbf{a}\nabla u(x) := \sum_{z\sim x} \mathbf{a}(\{x,z\})\,(u(z)-u(x))$.

Interacting particle systems

$\mathcal{M}_\delta(\mathbb{R}^d)$:	the set of locally finite measures $\mu = \sum_{i=1}^{\infty} \delta_{x_i}$;
$\llcorner$:	restriction operator, for $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$ and a Borel set $U, (\mu \llcorner U)(\cdot) = \mu(U \cap \cdot)$;
$\mathcal{F}_U$:	σ-algebra generated by the mappings $V \mapsto \mu(V)$ for all Borel sets $V \subseteq U$, $\mathcal{F}$ short for $\mathcal{F}_{\mathbb{R}^d}$;
$\mathcal{B}_U$:	σ-algebra generated by the Borel sets $V \subseteq U$;
$\mathbb{P}_\rho$:	Poisson point process on $\mathbb{R}^d$ with density $\rho \in (0, \infty)$;
$\mathbf{a}$:	random conductance $\mathbf{a} : \mathcal{M}_\delta(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^{d \times d}_{sym}$, with the uniform ellipticity $\lvert\xi\rvert^2 \leqslant \xi \cdot \mathbf{a}\xi \leqslant \Lambda\lvert\xi\rvert^2$, $\mathbf{a}(\mu, x)$ only depends on the configuration in $B_1(x)$;
$\mathscr{L}^2$:	$f \in \mathscr{L}^2(\mathcal{M}_\delta(\mathbb{R}^d), \mathcal{F}, \mathbb{P}_\rho)$ i.e. $\mathbb{E}_\rho[f^2] < \infty$;
∇ :	for a function $f : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}, e_k \cdot \nabla f(\mu, x) = \lim_{h \to 0} \frac{f(\mu - \delta_x + \delta_{x+he_k}) - f(\mu)}{h}$;
$\mathscr{H}^1(U)$:	$\nabla f : \mathcal{M}_\delta(\mathbb{R}^d) \times U \to \mathbb{R}$ is $\mathcal{F} \otimes \mathcal{B}_U$-measurable, completion for the norm $\lVert f \rVert_{\mathscr{H}^1(U)} = \left(\mathbb{E}_\rho[f^2] + \mathbb{E}_\rho\left[\int_U \lvert\nabla f\rvert^2 \, \mathrm{d}\mu\right]\right)^2$;
$\mathscr{H}^1_0(U)$:	completion for $f \in \mathscr{H}^1(U)$ and is $\mathcal{F}_K$-measurable for some compact set $K \subseteq U$;
$\mathscr{C}^\infty(U)$:	conditioned on $\mu(U) = n, \mu \llcorner U^c$, $(x_1, \cdots, x_n) \mapsto f(\sum_{i=1}^{n} \delta_{x_i} + \mu \llcorner U^c) \in C^\infty(U^n)$;
$\mathscr{C}^\infty_c(U)$:	$f \in \mathscr{C}^\infty(U)$ and is $\mathcal{F}_K$-measurable for some compact set $K \subseteq U$;
$\mathcal{F}\mathscr{C}^\infty_c(U)$:	$f = G(\mu(g_1), \cdots, \mu(g_n)), n \in \mathbb{N}, G \in C^\infty_b(\mathbb{R}^{nd}), g_i \in C^\infty_c(U)$.

1.1 An overview of homogenization theory

Homogenization theory has a long history and is much studied in various directions. The most classic topic is to study the behavior of the divergence form operator $-\nabla\cdot(\mathbf{a}(\frac{\cdot}{\varepsilon})\nabla)$ when $\varepsilon\to 0$. The two most typical settings are to be assume that $\mathbf{a}$ is either periodic, or stationary and ergodic. From the mathematical viewpoint, there are *qualitative* and *quantitative* results. As is to be expected, the qualitative results were obtained first, and allowed to identify the effective operator $-\nabla\cdot(\bar{\mathbf{a}}\nabla)$, where $\bar{\mathbf{a}}$ is constant, but $\bar{\mathbf{a}}$ is not the average or expectation of $\mathbf{a}$. Quantitative results were obtained significantly later, and aimed to determine rates of convergence. In fact, homogenization is also a useful numerical method, and error estimates are natural questions from the point of view of numerical analysis. Moreover, homogenization provides convenient tools for other topics in PDE and probability. These will be discussed in detail in the other sections of this chapter. Finally, no matter in which setting (periodic, stochastic) and which goals (qualitative, quantitative), most results in homogenization theory are constructed around three key objects: *the effective coefficient matrix*, *the correctors* and *the two-scale expansion*.

In this section, we review some of the general results in homogenization theory. We will talk about the results in the periodic setting at first, then we focus on the stochastic case. Some excellent monographs [47, 219, 145, 199, 210, 25] and expository papers [7, 185] on homogenization theory are good references.

1.1.1 Periodic homogenization

In this paragraph, we suppose that $\mathbf{a}:\mathbb{R}^d\to\mathbb{R}^{d\times d}_{sym}$ is $\mathbb{Z}^d$-periodic, symmetric matrix with uniform ellipticity condition $|\xi|^2\leqslant\xi\cdot\mathbf{a}\xi\leqslant\Lambda|\xi|^2$. We study the Dirichlet problem for $u^\varepsilon\in g+H^1_0(U)$

$$\begin{cases} -\nabla\cdot\left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon\right)=f & \text{in } U,\\ u^\varepsilon=g & \text{on } \partial U,\end{cases}\tag{1.4}$$

with $f\in H^{-1}(U), g\in H^1(U)$, and $U\subseteq\mathbb{R}^d$ with Lipschitz boundary. For $\varepsilon\to0$, the behavior of the solution u^ε can be approximated by the homogenized solution

$$\begin{cases} -\nabla\cdot(\bar{\mathbf{a}}\nabla\bar{u})=f & \text{in } U,\\ \bar{u}=g & \text{on } \partial U.\end{cases}\tag{1.5}$$

We give its precise statement:

Theorem 1.1.1 ([47, 212, 187])**.** *Given $(\mathbf{a}(x))_{x\in\mathbb{R}^d}$ a $\mathbb{Z}^d$-periodic, symmetric matrix field with the uniform ellipticity condition, then there exists a constant effective matrix $\bar{\mathbf{a}}$, such that the solution $(u^\varepsilon)_{\varepsilon>0}$ of the Dirichlet problem eq. (1.4) admits a homogenized solution $\bar{u}$ solving eq. (1.5) and, as ε tends to zero,*

$$u^\varepsilon\xrightarrow{L^2(U)}\bar{u},\qquad \nabla u^\varepsilon\xrightharpoonup{L^2(U)}\nabla\bar{u},\qquad \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon\xrightharpoonup{L^2(U)}\bar{\mathbf{a}}\nabla\bar{u},$$

where $\nabla u^\varepsilon\xrightharpoonup{L^2(U)}\nabla\bar{u}$ and $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon\xrightharpoonup{L^2(U)}\bar{\mathbf{a}}\nabla\bar{u}$ are understood as weak convergence.

Here we give a sketch of its proof. By the energy estimate bound, the weak compactness of $H^1(U)$, and the Rellich theorem, up to a subseqence we have

$$\varepsilon\to0,\qquad u^\varepsilon\xrightarrow{L^2(U)}\bar{u},\qquad \nabla u^\varepsilon\xrightharpoonup{L^2(U)}\nabla\bar{u},\qquad \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon\xrightharpoonup{L^2(U)}q,\tag{1.6}$$

where the quantity $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon$ and its weak limit q are sometimes called *the flux*. The main question is to characterize $\bar{u}, \bar{\mathbf{a}}$ and q. A classical heuristic method for this problem is *the two-scale asymptotic expansion ansatz* (see [47]): we write informally u^ε as

$$u^\varepsilon(x) = u_0\left(x, \frac{x}{\varepsilon}\right) + \varepsilon u_1\left(x, \frac{x}{\varepsilon}\right) + \varepsilon^2 u_1\left(x, \frac{x}{\varepsilon}\right) + \cdots, \tag{1.7}$$

where in each term $u_i : U \times \mathbb{T}^d \to \mathbb{R}$, and $u_i(x, \cdot)$ is $\mathbb{Z}^d$-periodic. The intuition here is to expand the function in different orders of ε like Taylor series, and use the first coordinate x to describe the macroscopic behavior, and the second coordinate $\frac{x}{\varepsilon}$ for the microscopic oscillating behavior. By a comparison of each order of ε, we will see for order zero $u_0 = \bar{u}$; for order ε, it is described by *the first-order correctors* $\{\phi_{e_i}\}_{1 \leqslant i \leqslant d}$ satisfying the equation of *the cell problem*

$$\begin{cases} -\nabla \cdot \mathbf{a}(e_i + \nabla \phi_{e_i}) = 0 & \text{in } \mathbb{T}^d, \\ \int_{\mathbb{T}^d} \phi_{e_i} = 0 & , \end{cases} \tag{1.8}$$

and $u_1\left(x, \frac{x}{\varepsilon}\right) = \sum_{i=1}^d (\partial_{x_i}\bar{u}(x))\phi_{e_i}\left(\frac{x}{\varepsilon}\right)$. Then we calculate the flux

$$\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon = \sum_{i=1}^d \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\left(e_i + \nabla\phi_{e_i}\left(\frac{\cdot}{\varepsilon}\right)\right)\partial_{x_i}\bar{u} + O(\varepsilon), \tag{1.9}$$

which implies the definition of the homogenized coefficient

$$\bar{\mathbf{a}} e_i := \int_{\mathbb{T}^d} \mathbf{a}(e_i + \nabla \phi_{e_i}), \tag{1.10}$$

because it allows us to see the weak limit in eq. (1.9) by passing $\varepsilon \to 0$. Finally, by the same argument, the weak limit of $\nabla \cdot \left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon\right)$ is $\nabla \cdot (\bar{\mathbf{a}}\nabla\bar{u})$ and this gives eq. (1.5).

This ansatz contains many ingredients and inspires many developments in homogenization theory. It helps us derive the definition of the correctors eq. (1.8), the effective coefficient matrix eq. (1.10) and the *two-scale expansion*

$$w^\varepsilon := \bar{u} + \varepsilon \sum_{i=1} (\partial_{x_i}\bar{u})\phi_{e_i}\left(\frac{\cdot}{\varepsilon}\right). \tag{1.11}$$

However, this ansatz is not rigorous, as the error order is $\|w^\varepsilon - u^\varepsilon\|_{H^1(U)} \simeq \sqrt{\varepsilon}$ for the reason of *the boundary layer effect*. See the discussion in [48, 8, 26].

The first rigorous proof for Theorem 1.1.1 is due to De Giorgi and Spanolo [212, 213, 88], where the argument is a compactness style method for differential operator $-\nabla \cdot (\mathbf{a}^\varepsilon \nabla)$. Moreover, this method only supposes the condition of symmetric coefficient matrices $(\mathbf{a}^\varepsilon)_{\varepsilon \geqslant 0}$ with a uniform estimate $|\xi|^2 \leqslant \xi \cdot \mathbf{a}^\varepsilon \xi \leqslant \Lambda|\xi|^2$, so it applies to more general settings than periodic or stationary coefficient. Later this method is extended to asymmetric matrices by Murat and Tartar in [187].

There are also some methods to make the asymptotic expansion ansatz rigorous. One elegant and robust approach is *the oscillating test function method* (also called *energy method*) proposed by Tartar. The main idea is to test eq. (1.4) with the oscillating function like $v^\varepsilon = v + \varepsilon \sum_{i=1}^d (\partial_{x_i} v)\phi_{e_i}\left(\frac{\cdot}{\varepsilon}\right)$ with $v \in C_c^\infty(U)$ and then pass ε to 0. In this procedure, one needs the weak convergence of the product of ∇u^ε and $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla v^\varepsilon$, and this is *the compensated compactness theorem* developed by Tartar in [218] and by Murat in [186]. Another convenient framework for periodic homogenization is *the two-scale convergence* by Nguetseng in [194] and by Allaire in [6], where they define a topology with more information than the classical weak convergence.

Various further results are developed in periodic homogenization. In the celebrated work [33, 35, 34] of Avellaneda and Lin, they prove the regularity results, Liouville theorems and Calderón-Zygmund estimate. In the work [146, 147, 148], Kenig, Lin and Shen develop quantitative homogenization for elliptic systems of periodic coefficient, including rate of convergence for the Dirichlet and Neumann problems and rate of convergence for Green's function. See also [210] for a complete review.

1.1.2 Stochastic homogenization

The theory of qualitative stochastic homogenization is developed in the 80's, with the work of Kozlov [160], Papanicolaou and Varadhan [198] and Yurinskiĭ [225]. The setting for the coefficient $\mathbf{a} : \mathbb{R}^d \to \mathbb{R}^{d\times d}_{sym}$ satisfies the following conditions

1. $\mathbf{a}$ is symmetric matrix with uniform ellipticity condition $|\xi|^2 \leqslant \xi \cdot \mathbf{a}\xi \leqslant \Lambda|\xi|^2$;
2. $\mathbf{a}$ is $\mathbb{Z}^d$-stationary ergodic random field.

Theorem 1.1.2 ([160], [198], [225])**.** *Given a coefficient field $(\mathbf{a}(x))_{x\in\mathbb{R}^d}$ satisfying the conditions above, then there exists a constant effective matrix $\bar{\mathbf{a}}$, such that the solution $(u^\varepsilon)_{\varepsilon>0}$ of the Dirichlet problem eq. (1.4) admits a homogenized solution $\bar{u}$ solving eq. (1.5) and, as ε tends to zero,*

$$u^\varepsilon \xrightarrow{L^2(U)} \bar{u}, \qquad \nabla u^\varepsilon \xrightarrow{L^2(U)} \nabla\bar{u}, \qquad \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon \xrightarrow{L^2(U)} \bar{\mathbf{a}}\nabla\bar{u}.$$

One can repeat the proof of the oscillating test functions, but a major difference is the construction of the corrector, as the corrector is no longer defined by the cell problem eq. (1.8). In fact, as the solution of the cell problem can be seen as a periodic solution in $\mathbb{R}^d$, it is natural to define the corrector ϕ_{e_i} solving

$$-\nabla \cdot \mathbf{a}(e_i + \nabla\phi_{e_i}) = 0 \text{ in } \mathbb{R}^d. \tag{1.12}$$

However, this equation is not well-defined if we do not give the function space. One approach is to add a regularization $\lambda > 0$

$$\lambda\phi^\lambda_{e_i} - \nabla \cdot \mathbf{a}(e_i + \nabla\phi^\lambda_{e_i}) = 0 \text{ in } \mathbb{R}^d. \tag{1.13}$$

We can take $\lambda \searrow 0$ and extract a subsequence of $\nabla\phi^\lambda_{e_i}$ which admits a limit of $\nabla\phi_{e_i}$ as a $\mathbb{Z}^d$-stationary random gradient field solving eq. (1.12). Then it suffices to define

$$\bar{\mathbf{a}}e_i := \mathbb{E}\left[\int_{[0,1]^d} \mathbf{a}(e_i + \nabla\phi_{e_i})\right], \tag{1.14}$$

and justify the weak convergence argument by *the Birkhoff ergodic theorem*

$$\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\left(e_i + \nabla\phi_{e_i}\left(\frac{\cdot}{\varepsilon}\right)\right) \xrightarrow{L^2} \bar{\mathbf{a}}e_i.$$

Qualitative stochastic homogenization then has various applications, while there are some aspects quite non-convenient to use:

- We recall that the solution of eq. (1.12) is solved for $\nabla\phi_{e_i}$ instead of ϕ_{e_i}, so ϕ_{e_i} is defined up to a constant and *a priori* it is not stationary. This is quite different from periodic homogenization, where ϕ_{e_i} itself is periodic.
- To obtain $\nabla\phi_{e_i}$, we have to solve the problem in the whole space $\mathbb{R}^d$, which is impossible in practice. Meanwhile, the cell problem eq. (1.8) only requires to solve the problem in a unit torus.
- To get $\bar{\mathbf{a}}$ in practice also inherits the difficulty from that of $\nabla\phi_{e_i}$.

A practical method to calculate $\bar{\mathbf{a}}$ is to use Theorem 1.1.2 in a unit cube $\square := \left(-\frac{1}{2}, \frac{1}{2}\right)^d$ with affine boundary condition

$$\begin{cases} -\nabla\cdot\left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon\right) = 0 & \text{in } \square, \\ u^\varepsilon(x) = e_i\cdot x & \text{on } \partial\square. \end{cases} \tag{1.15}$$

Since its homogenized solution is $\bar{u}(x) = e_i\cdot x$, the weak convergence of flux is $\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon \xrightarrow{L^2} \bar{\mathbf{a}}e_i$, thus we can use the spatial average to approximate $\bar{\mathbf{a}}$

$$\int_\square \mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla u^\varepsilon \xrightarrow{\varepsilon\to 0} \bar{\mathbf{a}}e_i.$$

After a change of scale, this is equivalent to approximate eq. (1.12) in a big cube $\square_m := \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d$ with $\phi_{e_i,m} \in H^1_0(\square_m)$ solving

$$-\nabla\cdot\mathbf{a}(e_i + \nabla\phi_{e_i,m}) = 0 \text{ in } \square_m, \tag{1.16}$$

and the spatial average in large scale becomes

$$\mathbf{a}(\square_m)e_i := \frac{1}{|\square_m|}\int_{\square_m}\mathbf{a}(e_i + \nabla\phi_{e_i,m}), \qquad \mathbf{a}(\square_m) \xrightarrow{m\to\infty} \bar{\mathbf{a}}. \tag{1.17}$$

This method is called *the representative elementary volume* (or *the volume averaging method*), and is largely used as a numerical method. In the work [63] of Bourgeat and Piatnitski, they prove the consistency of this method for eq. (1.16) with Dirichlet, Neumann, or periodic boundary condition. They also obtain a non-explicit rate of convergence for $\mathbb{E}[\mathbf{a}(\square_m)]$ under certain mixing conditions.

The quantitative theory of stochastic homogenization is developed in the recent years. One approach is using *the Efron-Stein inequality* building upon the ideas of Naddaf and Spencer in [188]. In the work of Gloria and Otto [123, 124], they study the problem defined on lattice graph $(\mathbb{Z}^d, E_d)$ and suppose

$$\{\mathbf{a}(e)\}_{e\in E_d} \text{ i.i.d. and } 0 < \alpha < \mathbf{a}(e) < \beta < \infty.$$

Then for the resolvent problem eq. (1.13), they obtain an uniform estimation for $d \geqslant 3$ [123, Proposition 2.1]

$$\mathbb{E}\left[|\phi^\lambda_{e_i}|^p\right] \leqslant C_p.$$

This ([123, Corollary 2.1]) answers the long time open question: for $d \geqslant 3$, there exists a unique stationary random field ϕ_{e_i} solving eq. (1.12) such that $\mathbb{E}[\phi_{e_i}] = 0$. Later, this method is also generalized to $\mathbb{R}^d$ setting by supposing *the spectral gap condition* for $\mathbf{a}$ in Gloria and Otto [125] and Gloria, Neukamm, Otto [121].

Another approach for quantitative homogenization is *the renormalization approach* initiated by Armstrong and Smart in [31], who extended the techniques of Avellaneda and Lin [33, 35] and the ones of Dal Maso and Modica [80, 81]. These results were then improved in a series of work [30, 23, 24] by Armstrong, Kuusi and Mourrat, and now reformulated in the monograph [25] by the same authors. They work on $\mathbb{R}^d$ setting and suppose

$$(\mathbf{a}(x))_{x\in\mathbb{R}^d} \text{ is of finite range correlation.}$$

A unit range of correlation means that, for any two sets $U, V \subseteq \mathbb{R}^d$ such that $\operatorname{dist}(U,V) \geqslant 1$, the coefficients $(\mathbf{a}(x))_{x\in U}$ and $(\mathbf{a}(x))_{x\in V}$ are independent. In fact, this method is robust and also applies to general coefficient fields with polynomial mixing condition. As this thesis also employs much the renormalization approach, we do a short review in the following paragraphs.

The main idea is similar to eq. (1.17) and we need rate of convergence. Let $\square_m = \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d$ and $\ell_p(x) := p\cdot x$, we define the Dirichlet energy density in the finite volume

$$\nu(\square_m, p) := \inf_{v\in \ell_p + H^1_0(\square_m)} \frac{1}{|\square_m|}\int_{\square_m} \frac{1}{2}\nabla v\cdot \mathbf{a}\nabla v. \tag{1.18}$$

We denote by $v(\square_m, p, \cdot)$ its minimiser, and $\nu(\square_m, p) = \frac{1}{2}p\cdot \mathbf{a}(\square_m)p$ from the definition in eq. (1.16) and eq. (1.17). We observe that $\nu(\square_m, p)$ is a *subadditive quantity*, because for a scale $n < m$

$$\widetilde{v}(x) = \sum_{z\in 3^n\mathbb{Z}^d\cap\square_m} v(z+\square_n, p, x)\mathbf{1}_{\{x\in z+\square_n\}},$$

provides a sub-minimiser for the optimization problem of $\nu(\square_m, p)$. Then we have

$$\nu(\square_m,p) \leqslant \frac{1}{|\square_m|}\int_{\square_m}\frac{1}{2}\nabla\widetilde{v}\cdot\mathbf{a}\nabla\widetilde{v} = 3^{-d(m-n)}\sum_{z\in 3^n\mathbb{Z}^d\cap\square_m}\nu(z+\square_n,p).$$

By the stationarity, we take the expectation and obtain that $\mathbb{E}[\mathbf{a}(\square_m)] \leqslant \mathbb{E}[\mathbf{a}(\square_n)]$, so the decreasing sequence $\{\mathbb{E}[\mathbf{a}(\square_m)]\}_{m\geqslant 1}$ admits a limit. We define

$$\bar{\mathbf{a}} := \lim_{m\to\infty}\mathbb{E}[\mathbf{a}(\square_m)], \tag{1.19}$$

and from eq. (1.17), we know that the definitions in eq. (1.19) and eq. (1.14) coincide.

In order to obtain the convergence rate of $\bar{\mathbf{a}}(\square_m)$ to $\bar{\mathbf{a}}$, we consider *the dual problem*

$$\nu^*(\square_m, q) := \sup_{u\in H^1(\square_m)}\frac{1}{|\square_m|}\int_{\square_m}\left(-\frac{1}{2}\nabla u\cdot\mathbf{a}\nabla u + q\cdot\nabla u\right). \tag{1.20}$$

We denote by $u(\square_m, q, \cdot)$ the maximiser, and $\nu^*(\square_m, q) = \frac{1}{2}q\cdot\mathbf{a}_*^{-1}(\square_m)q$ since one can check that $q\mapsto \nu^*(\square_m, q)$ is also a quadratic form. By a similar argument and notice that $u(\square_m, q, \cdot)$ is a sub-maximiser for every problem $\nu^*(z+\square_n, q), z\in 3^n\mathbb{Z}^d\cap\square_m$, we have

$$\begin{aligned}\nu^*(\square_m,q) &= 3^{-d(m-n)}\sum_{z\in 3^n\mathbb{Z}^d\cap\square_m}\frac{1}{|\square_n|}\int_{z+\square_n}\left(-\frac{1}{2}\nabla u(\square_m,q,\cdot)\cdot\mathbf{a}\nabla u(\square_m,q,\cdot) + q\cdot\nabla u(\square_m,q,\cdot)\right)\\ &\leqslant 3^{-d(m-n)}\sum_{z\in 3^n\mathbb{Z}^d\cap\square_m}\nu^*(z+\square_n,q).\end{aligned}$$

Therefore, $\nu^*(\square_m, q)$ is also a subadditive quantity, and $\{\mathbb{E}[\mathbf{a}_*(\square_m)]\}_{m\geqslant 1}$ is an increasing sequence. The dual quantity helps control rate of convergence because we can test $\nu^*(\square_m, q)$ with the minimiser $v(\square_m, p, \cdot)$ of $\nu(\square_m, p)$ and obtain

$$-\frac{1}{2}p \cdot \mathbf{a}(\square_m)p + p \cdot q \leqslant \frac{1}{2}q \cdot \mathbf{a}_*^{-1}(\square_m)q.$$

By setting $q = \mathbf{a}_*(\square_m)p$, we obtain

$$\mathbf{a}_*(\square_m) \leqslant \mathbf{a}(\square_m). \tag{1.21}$$

The rate of convergence can be bounded by

$$|\mathbb{E}[\mathbf{a}(\square_m)] - \bar{\mathbf{a}}| \leqslant |\mathbb{E}[\mathbf{a}(\square_m)] - \mathbb{E}[\mathbf{a}_*(\square_m)]|. \tag{1.22}$$

In practice, eq. (1.22) can be very useful, because the quantities $\mathbf{a}(\square_m), \mathbf{a}_*(\square_m)$ can always be calculated locally in $\square_m$, and the fluctuation can be estimated by CLT or concentration inequality. Thus, if we observe that $|\mathbf{a}(\square_m) - \mathbf{a}_*(\square_m)|$ is very small, then we can claim the approximation $|\mathbf{a}(\square_m) - \bar{\mathbf{a}}|$ is also very precise.

The theoretical proof that $\lim_{m\to\infty} |\mathbb{E}[\mathbf{a}(\square_m)] - \mathbb{E}[\mathbf{a}_*(\square_m)]| = 0$ requires more work. One can find its original proof in [31], or a simplified proof in [25, Chapter 2] where *the multiscale Poincaré inequality* is used. We can not only prove the convergence of the expectation, but also control the fluctuation: there exist an exponent $\alpha(d,\Lambda) \in \left(0, \frac{1}{2}\right]$ and, for any $s \in (0,d)$, a constant $C(s,d,\Lambda) < \infty$ such that

$$|\mathbf{a}(\square_m) - \bar{\mathbf{a}}| + |\mathbf{a}_*(\square_m) - \bar{\mathbf{a}}| \leqslant C3^{-\alpha(d-s)m} + \mathcal{O}_1(3^{-sm}), \tag{1.23}$$

where the $\mathcal{O}_s$ notation is defined as

$$X \leqslant \mathcal{O}_s(\theta) \iff \mathbb{E}\left[\exp((\theta^{-1}X)_+^s)\right] \leqslant 2. \tag{1.24}$$

Generally, the $\mathcal{O}_s(\theta)$ notation describes a random variable of typical size θ with sub- or super-exponential tail. When we take s close to d to reduce the part of fluctuation, eq. (1.23) allows to control the probability of large deviations of $\mathbf{a}(\square_m)$ very tightly. At the price of reducing the exponent s, one can later improve the exponent α to its optimal value, see [25, Chapter 4]. The rate of convergence for $|\mathbf{a}(\square_m) - \bar{\mathbf{a}}|$ also measures the convergence of the correctors, the flux and the homogenized solution, see [25, Chapter 1].

Finally, the renormalization approach is very robust and now applies to the homogenization of parabolic equations [18], the finite-difference equations on percolation clusters [19, 83, 85], the differential forms [84], the "$\nabla\phi$" interface model [82, 29], the Villain model [86], the Coulomb gases [28], and the interacting particle systems [115].

1.2 Homogenization and numerical algorithms

This part will at first talk about the interest of homogenization theory in the numerical solution of PDE, and then in Section 1.2.1 we introduce the contribution of this thesis (Chapters 2 and 3) in this direction.

The main question that we hope to address is the numerical method for the Dirichlet problem in large scale: let $U \subseteq \mathbb{R}^d$ with Lipschitz boundary and $U_r := rU$

$$\begin{cases} -\nabla \cdot (\mathbf{a}\nabla u) = f & \text{in } U_r, \\ u = g & \text{on } \partial U_r, \end{cases} \tag{1.25}$$

with the coefficient $(\mathbf{a}(x))_{x\in\mathbb{R}^d}$ symmetric matrix satisfying uniform ellipticity condition, $\mathbb{Z}^d$-periodic or $\mathbb{Z}^d$-stationary and ergodic. This problem can be also reformulated in a fixed domain U with the scaling ε like eq. (1.4). The challenge here is the need to refine the mesh when $r \to \infty$ or $\varepsilon \to 0$, thus the numerical cost increases and we hope to find some efficient algorithms. The answer to this question also depends on the concrete setting and here we give a brief review.

Solution at one point

If one only wants to get the solution of eq. (1.25) at one point, for example $u(x_0), x_0 \in U_r$, then the most practical way is the probabilistic method using *the Monte-Carlo Markov Chain (MCMC)* algorithm. Let us see an easy example $f = 0$ and $g \in C^1(U_r)$. It suffices to run a diffusion $(X_t)_{t\geqslant 0}$ associated to the operator $-\nabla \cdot (\mathbf{a}\nabla)$ starting from x_0, and let τ be the hitting time on the boundary ∂U_r, then we have

$$u(x_0) = \mathbb{E}[g(X_\tau)]. \tag{1.26}$$

This probabilistic representation generates a MCMC algorithm which is also dimension free. It even does not use the periodic condition or stationarity of $\mathbf{a}$, and also works for a general large domain U_r with certain regularity of the boundary. (A general sufficient condition is *the cone condition*, see the discussions in [78].) Of course, we also have to do some discrete approximation for the diffusion, see [76, 37, 38, 158, 176] for the approximation error estimate.

Solution at every point

The main challenge is to solve eq. (1.25) for every point in the domain U_r. In this case, the MCMC algorithm also requires many simulations of diffusion issued from different starting points, which increases the complexity. If one solves eq. (1.25) by the classical *finite difference method*, it is equivalent to solve a large linear system. A naive algorithm is *the Jacobi iterative method*: after the discretisation of eq. (1.25) in $(\mathbb{Z}^d, E_d)$, we set

$$P(x,y) := \frac{\mathbf{a}(\{x,y\})}{\sum_{z\sim x}\mathbf{a}(\{x,z\})}, \qquad \widetilde{f}(x) = f(x)/(\sum_{z\sim x}\mathbf{a}(\{x,z\})), \tag{1.27}$$

then we do the iteration

$$u_0 = g, \qquad u_{n+1} = J(u_n, \widetilde{f}), \qquad J(u_n, \widetilde{f}) := Pu_n + \widetilde{f}. \tag{1.28}$$

From the probabilistic viewpoint, this follows the same spirit of the MCMC method, but we do iterations for the semigroup of the Dirichlet problem instead of the simulations of trajectories. The contraction rate in eq. (1.28) depends on *the spectral gap*, and for domain U_r, it can be about $\left(1 - \frac{1}{r^2}\right)$. Therefore, for a precision of ε_0, it requires $O(r^2|\log \varepsilon_0|)$ rounds of iterations. This algorithm can be a little accelerated by *the conjugate gradient method (CGM)*, which achieves a contraction rate $\left(1 - \frac{1}{r}\right)$ thus it suffices $O(r|\log \varepsilon_0|)$ rounds of CGM (see [207, Theorem 6.29, eq.(6.128)]). As the numerical cost for one iteration of CGM is close to that of Jacobi method, all these methods will have a large complexity when r increases.

Solution for constant coefficient equation at every point

The multigrid algorithm is a standard and powerful method for the Dirichlet problem in a large domain with constant coefficient. One can find the comprehensive study of this method in [137, 223, 99, 64] and here we give a version in our context. Suppose we want to solve $-\Delta u = f$ with $u \in g + H_0^1(U_r)$, the algorithm can be stated as follows: set the finest grid of scale $\frac{r}{M}$, and denote by J^M the Jacobi method in eq. (1.28) for this grid.

1. Start by an initial guess $u_0 = g$.

2. Implement a multigrid iteration step with Jacobi method

   (a) $u_1 = J^M(u_0, f)$;

   (b) $f_1 = f - (-\Delta u_1)$, coarsen the grid by 2, and $u_2 = J^{M/2}(0, f_1)$;

   (c) $f_2 = f_1 - (-\Delta u_2)$, coarsen the grid by 2, and $u_3 = J^{M/4}(0, f_2)$.

3. Set $\hat{u} := u_1 + u_2 + u_3$ and put $\hat{u}$ in the place of u_0. Go back to the step 2 and repeat this procedure of iterations.

In practice, one needs to add several intermediate scales in the multigrid iteration step. Notice that the coarsened grid is not precise, but it can recover the macroscopic behavior of the solution with less numerical cost; the fine grid can calculate the solution in microscopic scale, but the value propagates slowly in the Jacobi method and requires many steps of computations. Therefore, we combine different grids and can solve this solution more efficiently. For the precision ε_0, the numerical cost is about $O(|\log \varepsilon_0|)$ rounds of CGM ([64, Chapter 4]) - we can always replace the Jacobi method in the algorithm by the conjugate gradient method, but the former is easier to state. Finally, we remark that some operations are necessary for the passage of functions between the fine grid and the coarsened grid. These are called *coarsening operator* and *projection operator*, which are very important ideas in the multigrid algorithm. In our setting, the coarsening operator and projection operator are just samples of grids and linear interpolations, as Δ gives more regularity to the solution compared to $-\nabla \cdot (\mathbf{a}\nabla)$. This also explains why the classical multigrid algorithm requires the constant coefficient condition.

Homogenized solution

The multigrid algorithm above explains the interest of homogenization theory in the numerical solution of Dirichlet problem. Instead of solving the eq. (1.25) directly, we can solve its homogenized solution $\bar{u}$ with the multigrid algorithm. Then we only have to pay a small error $\|u - \bar{u}\|_{L^2(U_r)}$, and by homogenization theory, this error is quite small compared to $\|u\|_{L^2(U_r)}$ for large r. We refer to the references [42, 36, 101, 128, 197, 178, 159, 196], as well as to [129, 154, 103, 104] for this idea.

Therefore, when we combine the homogenized solution and the multigrid algorithm, it suffices to obtain the effective coefficient $\bar{\mathbf{a}}$. As we have discussed a lot in Section 1.1, this task is more complicated in stochastic homogenization than periodic homogenization.

- For $\mathbf{a}$ a $\mathbb{Z}^d$-periodic coefficient, we can get $\bar{\mathbf{a}}$ by solving the cell problem eq. (1.8) mentioned in Section 1.1.

- For the stochastic coefficient setting, we use the representative elementary volume (REV) mentioned in eq. (1.17). More precisely, we divide the data $(\mathbf{a}(x))_{x \in U_r}$ into subsets of scale l, apply eq. (1.17) in each subset, and then average over the $\left(\frac{r}{l}\right)^d$ copies to reduce the fluctuation.

For the stochastic coefficient case, there are many references [119, 102, 184, 107, 138] discussing about the errors and numerical costs. Among them, [102] studied the model on $(\mathbb{Z}^d, E_d)$ with i.i.d. conductance $\{\mathbf{a}(e)\}_{e\in E_d}$. Its main result says with a choice $l = r^{\frac{1}{2}}$ in the REV, we have a precision $r^{-\frac{d}{2}}$ with complexity $O(r^{\frac{1}{2}})$ rounds of CGM. Later [184] proposed another efficient algorithm which allows us to get the optimal complexity in general stochastic homogenization setting: a precision $r^{-\frac{d}{2}}$ with the $O(\log r)$ rounds of CGM. No matter which method we take, to get $\bar{\mathbf{a}}$ with good precision is not very costly.

Beyond the homogenized solution

Although the homogenized solution $\bar{u}$ is a good approximation to eq. (1.25), it is too smooth to recover the microscopic details. To go one step further, one strategy is the two-scale expansion method. As mentioned in eq. (1.11), $\bar{u}$ only converges to u in L^2, but $w = \bar{u} + \sum_{i=1}^d (\partial_{x_i}\bar{u})\phi_{e_i}$ gives an approximation of the solution of eq. (1.25) in the sense H^1. So in the periodic coefficient setting, we can attack at first the cell problem eq. (1.8) to get both $\bar{\mathbf{a}}$ and all the first-order correctors $\{\phi_{e_i}\}_{1\leqslant i\leqslant d}$, then we solve the homogenized solution $\bar{u}$. Combining the correctors and homogenized solution, w gives us a better approximation. This method can also be improved a little by using a modified correctors that

$$\widetilde{w} = \bar{u} + \sum_{i=1}^d (\partial_{x_i}\bar{u})\phi_{e_i}\eta, \tag{1.29}$$

with $\eta \in C_c^\infty(U_r)$ a smooth cut-off function in order to remove the main error sources - *the boundary layer effect*. See the discussion of this topic in [48, 8, 26]. Unfortunately, this idea can hardly be used in stochastic homogenization because the complexity to compute the correctors is the same as to calculate the original solution u.

Notice that for a fixed r, there is always a limit of precision between the approximated solutions $\bar{u}, w, \widetilde{w}$ and the real solution u, so it is natural to look for an efficient algorithm with a resolution beyond this limit. From the discussion above, it seems that there is a trade-off between precision and numerical cost. In fact, there is a third dimension: probability. We will see that we can pay some probability of consistency to gain both precision and numerical efficiency. One example is the AHKM iterative algorithm [22], which will be discussed in details in Section 1.2.1.

1.2.1 Summary of Chapters 2 and 3

The AHKM iterative algorithm is the main topic studied in Chapters 2 and 3. It is invented by Armstrong, Hannukainen, Kuusi and Mourrat in [22], which aims to get an approximation of u beyond the precision of the homogenized solution $\bar{u}$ with reasonable numerical costs. It follows the spirit of the multigrid algorithm and also makes use of homogenization theory.

Let us present at first the structure of the AHKM algorithm for eq. (1.25).

1. We start by an initial guess $u_0 = g$, and choose a parameter of regularization $\lambda \in \left(\frac{1}{r}, \frac{1}{2}\right)$.

2. We solve the following systems

$$\begin{cases} (\lambda^2 - \nabla\cdot\mathbf{a}\nabla)u_1 &= f + \nabla\cdot\mathbf{a}\nabla u_0 & \text{in } U_r, \\ -\nabla\cdot\bar{\mathbf{a}}\nabla\bar{u} &= \lambda^2 u_1 & \text{in } U_r, \\ (\lambda^2 - \nabla\cdot\mathbf{a}\nabla)u_2 &= (\lambda^2 - \nabla\cdot\bar{\mathbf{a}}\nabla)\bar{u} & \text{in } U_r. \end{cases} \tag{1.30}$$

3. We set $\hat{u} := u_0 + u_1 + u_2$ and put it back to the place of u_0 to repeat the iterations in step 2.

This looks very similar to the multigrid algorithm: in the first equation of eq. (1.30), we solve the Dirichlet problem in fine grid. But we add some regularization to reduce the rounds of CGM. Since $(u_0 + u_1)$ cannot recover all the solution, we put the residual

$$\lambda^2 u_1 = f - (-\nabla \cdot \mathbf{a} \nabla (u_0 + u_1)),$$

as the source in the right-hand side of the second equation of eq. (1.30). In the second equation of eq. (1.30), we just solve the problem in coarsened grid with the homogenized solution. However, the homogenized solution is too smooth for the fine grid. Thus in the third equation of eq. (1.30), we do some post-treatment and one can think u_2 as the projection of $\bar{u}$ in the find grid for operator $\lambda^2 - \nabla \cdot \mathbf{a} \nabla$.

To prove the consistency of the AHKM algorithm, the main ingredient is the two-scale expansion $w := \bar{u} + \sum_{k=1}^{d} (\partial_{x_k} \bar{u}) \phi_{e_k}$. Combining the first equation, the second equation of eq. (1.30) and eq. (1.25), we can obtain that

$$-\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} = -\nabla \cdot \mathbf{a} \nabla (u - u_0 - u_1) \qquad \text{in } U_r,$$

which is an equation of homogenization, so we have $(u - u_0 - u_1) \simeq w$. Moreover, the third equation in eq. (1.30) also follows the form of homogenization. Thus, we have

$$(u - u_0 - u_1) \simeq w \simeq u_2,$$

up to a small error in the sense $H^1(U_r)$, so we can estimate $\|\hat{u} - u\|_{H^1(U_r)}$ by studying

$$\|\hat{u} - u\|_{H^1(U_r)} = \|u - (u_0 + u_1 + u_2)\|_{H^1(U_r)} \leqslant \|(u - u_0 - u_1) - w\|_{H^1(U_r)} + \|w - u_2\|_{H^1(U_r)}.$$

Most of the idea above has already been included in the paper [22], but as the environment is random, the contraction rate is also a random variable. In [22], the authors obtain a bound for this contraction rate of one step, but this estimate cannot be iterated. The contribution in Chapter 2 is a *uniform bound* for the contraction rate. This uniform bound can then be iterated to justify the validity of the algorithm. In the following statement, the $\mathcal{O}_s$ notation is defined in eq. (1.24) and $\ell(\lambda)$ is defined as

$$\ell(\lambda) := \begin{cases} (\log(1 + \lambda^{-1}))^{\frac{1}{2}} & d = 2, \\ 1 & d > 2. \end{cases}$$

Theorem 1.2.1 (Main theorem in Chapter 2, Theorem 2.1.1)**.** *For every bounded domain $U \subseteq \mathbb{R}^d$ with $C^{1,1}$ boundary and every $s \in (0,2)$, there exists a positive finite constant $C(U, \Lambda, s, d)$ and, for every $r \geqslant 2$ and $\lambda \in \left(\frac{1}{r}, \frac{1}{2}\right)$, a random variable $\mathcal{Z}$ satisfying*

$$\mathcal{Z} \leqslant \mathcal{O}_s \left(C \ell(\lambda)^{\frac{1}{2}} \lambda^{\frac{1}{2}} (\log r)^{\frac{1}{s}} \right), \tag{1.31}$$

such that the following holds. Denote $U_r := rU$, let $f \in H^{-1}(U_r)$, $g \in H^1(U_r)$, $u_0 \in g + H^1_0(U_r)$, let $u \in g + H^1_0(U_r)$ be the solution of eq. (1.25), and let $u_1, \bar{u}, u_2 \in H^1_0(U_r)$ solve eq. (1.30) with null Dirichlet boundary condition. Then for $\hat{u} := u_0 + u_1 + u_2$, we have the contraction estimate

$$\|\nabla(\hat{u} - u)\|_{L^2(U_r)} \leqslant \mathcal{Z} \|\nabla(u_0 - u)\|_{L^2(U_r)}. \tag{1.32}$$

Therefore, the contraction rate of the AHKM algorithm can be bounded by a random variable $\mathcal{Z}$ of the order of $\lambda^{\frac{1}{2}}$, and more precisely,

$$\mathbb{P}[\mathcal{Z} \geqslant x] \leqslant 2\exp\left(-\left(\frac{x}{C\ell(\lambda)^{\frac{1}{2}}\lambda^{\frac{1}{2}}(\log r)^{\frac{1}{s}}}\right)^s\right).$$

By a reasonable choice $\lambda \simeq (\log r)^{-1}$, for a precision ε_0 the complexity of AHKM algorithm is $O(\log r|\log \varepsilon_0|^2)$. In conclusion, the AHKM algorithm achieves both high precision and low numerical cost, with the price of the exclusion of an event of very small probability.

The AHKM algorithm is a quite robust method and it also applies to other Dirichlet problems in degenerate random environment. The main contribution in Chapter 3 is an example for its application on *percolation clusters*, which can be used to simulate the model in porous medium of two types of composites with high contrast. See [220] for a comprehensive introduction and [95, 163, 175] for some examples of its applications in nanomaterials.

We give a brief introduction of the percolation model here and more details can be found in Section 1.3. On the lattice graph $(\mathbb{Z}^d, E_d)$, let $\mathbf{a} : E_d \to \{0\} \cup [\Lambda^{-1}, 1]$ such that the random variables $\{\mathbf{a}(e)\}_{e \in E_d}$ are independent and identically distributed. The Bernoulli percolation is defined by the random conductance $\{\mathbf{a}(e)\}_{e \in E_d}$: for every bond $e \in E_d$, we say that e is an *open bond* if $\mathbf{a}(e) > 0$, and that e is a *closed bond* otherwise. The connected components on $(\mathbb{Z}^d, E_d)$ generated by the open bonds are called *clusters*. For $d \geqslant 2$, there exists a parameter $\mathfrak{p}_c(d)$ such that for $\mathfrak{p} := \mathbb{P}[\mathbf{a}(e) > 0] > \mathfrak{p}_c$, there exists a unique infinite percolation cluster $\mathscr{C}_\infty$ [149]. This case is called *supercritical percolation*, and under this setting in a finite cube $\square_m := \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d \cap \mathbb{Z}^d$, typically we will see a giant cluster $\mathscr{C}_*(\square_m)$. This is a counterpart of $\mathscr{C}_\infty$ (see Figure 1.1 for an illustration) and we call this case "$\square_m$ *is a good cube*". The rigorous definitions of "$\square_m$ is a good cube" and of "the maximal cluster $\mathscr{C}_*(\square_m)$" will be given in Section 3.2, and they are typical since there exists a positive constant $C(d, \mathfrak{p})$ such that

$$\mathbb{P}[\square_m \text{ is a good cube}] \geqslant 1 - C(d, \mathfrak{p})\exp(-C(d, \mathfrak{p})^{-1}3^m).$$

Informally, one can just treat $\mathscr{C}_*(\square_m)$ as $\mathscr{C}_\infty \cap \square_m$. Our goal is to find an algorithm for solving the Dirichlet problems on $\mathscr{C}_*(\square_m)$

$$\begin{cases} -\nabla \cdot \mathbf{a}\nabla u = f & \text{in } \mathscr{C}_*(\square_m), \\ u = g & \text{on } \mathscr{C}_*(\square_m) \cap \partial\square_m, \end{cases} \tag{1.33}$$

where the divergence-form operator is defined as

$$-\nabla \cdot \mathbf{a}\nabla u(x) := \sum_{y \sim x} \mathbf{a}(\{x, y\})\,(u(x) - u(y))\,. \tag{1.34}$$

The AHKM algorithm on percolation clusters is as follows: we denote by $C_0(\square_m)$ the functions with null boundary condition on $\square_m$ and $\lambda_{\mathscr{C},m} := \lambda \mathbf{1}_{\{\mathscr{C}_*(\square_m)\}}$.

Theorem 1.2.2 (Main theorem in Chapter 3, Theorem 3.1.1)**.** *There exist two finite positive constants $s := s(d, \mathfrak{p}, \Lambda), C := C(d, \mathfrak{p}, \Lambda, s)$, and for every integer $m > 1$ and $\lambda \in \left(\frac{1}{3^m}, \frac{1}{2}\right)$, an $\mathcal{F}$-measurable random variable $\mathcal{Z}$ satisfying*

$$\mathcal{Z} \leqslant \mathcal{O}_s\left(C\ell(\lambda)^{\frac{1}{2}}\lambda^{\frac{1}{2}}m^{\frac{1}{s}+d}\right),$$

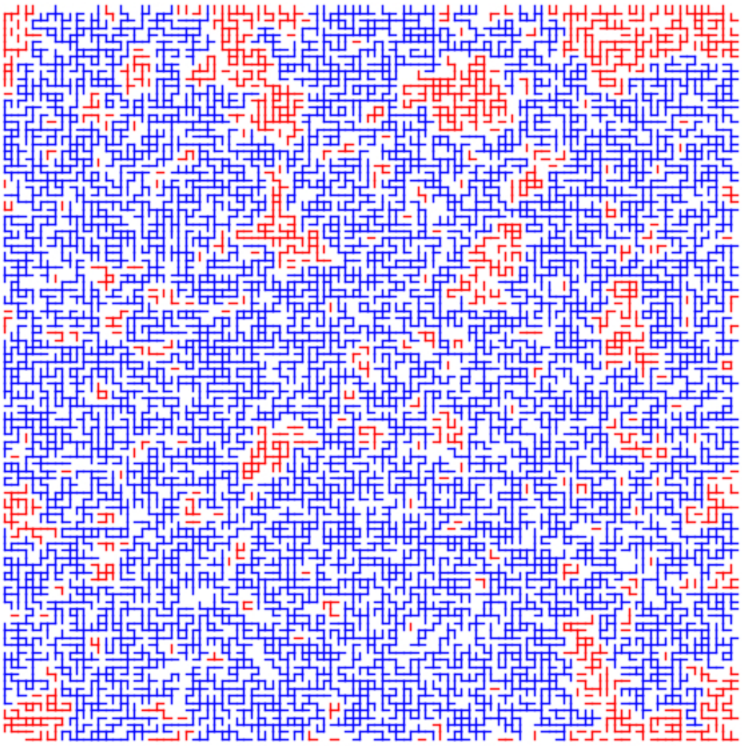

Figure 1.1: A simulation of $2D$ Bernoulli bond percolation with $\mathfrak{p} = 0.51$ in a cube $\square$ of size 100×100. The cluster in blue is the maximal cluster $\mathscr{C}_*(\square)$ while the clusters in red are the other small ones.

such that the following holds. Let $f, g : \square_m \to \mathbb{R}, u_0 \in g + C_0(\square_m)$ *and* $u \in g + C_0(\square_m)$ *be the solution of eq.* (1.33)*. On the event that* $\square_m$ *is a good cube, for* $u_1, \bar{u}, u_2 \in C_0(\square_m)$ *solving (with null Dirichlet boundary condition)*

$$\begin{cases} (\lambda^2 - \nabla \cdot \mathbf{a} \nabla) u_1 & = f + \nabla \cdot \mathbf{a} \nabla u_0 & \text{in } \mathscr{C}_*(\square_m) \backslash \partial \square_m, \\ -\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} & = \lambda^2_{\mathscr{C},m} u_1 & \text{in } \operatorname{int}(\square_m), \\ (\lambda^2 - \nabla \cdot \mathbf{a} \nabla) u_2 & = (\lambda^2 - \nabla \cdot \bar{\mathbf{a}} \nabla) \bar{u} & \text{in } \mathscr{C}_*(\square_m) \backslash \partial \square_m, \end{cases} \tag{1.35}$$

and for $\hat{u} := u_0 + u_1 + u_2$*, we have the contraction estimate*

$$\left\| \nabla (\hat{u} - u) \mathbf{1}_{\{\mathbf{a} \neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))} \leqslant \mathcal{Z} \left\| \nabla (u_0 - u) \mathbf{1}_{\{\mathbf{a} \neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))}. \tag{1.36}$$

The novelty of this application on the percolation cluster is to define a multigrid algorithm on the singular random environment because $\mathbf{a}$ does not have uniform ellipticity. Therefore, in the first and third equation of eq. (1.35), the fine grid is defined on the percolation cluster, while the coarsened grid of the second equation of eq. (1.35) is defined on $\square_m$. This implies that not only the random coefficient, but also the random geometry is homogenized. To see more precisely eq. (1.35) defines the proper coarseing and projection operator, we observe that u_1, u_2 also solves the following equivalent iterations with *any arbitrary extension* of value on $\square_m \backslash \mathscr{C}_*(\square_m)$

$$\begin{cases} (\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla) u_1 & = f_{\mathscr{C},m} + \nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla u_0 & \text{in } \operatorname{int}(\square_m), \\ -\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} & = \lambda^2_{\mathscr{C},m} u_1 & \text{in } \operatorname{int}(\square_m), \\ (\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla) u_2 & = (\lambda^2_{\mathscr{C},m} - \nabla \cdot \bar{\mathbf{a}} \nabla) \bar{u} & \text{in } \operatorname{int}(\square_m), \end{cases} \tag{1.37}$$

where $\mathbf{a}_{\mathscr{C},m}(\{x,y\}) = \mathbf{a}(\{x,y\})\mathbf{1}_{\{x,y\in\mathscr{C}_*(\square_m)\}}$ and $f_{\mathscr{C},m} = f\mathbf{1}_{\{\mathscr{C}_*(\square_m)\}}$. See more details in Proposition 3.1.1. A second challenge in this application is the consistency analysis, because quantitative homogenization theory on the percolation clusters is missing for long time until the recent work of [19, 83].

1.3 Homogenization on percolation clusters

In this part we will introduce at first *the* $\mathbb{Z}^d$*-Bernoulli percolation model*, and then review the results of the random walk on it. We will also point out its links with homogenization theory, and present our contribution of Chapter 4 in Section 1.3.

The $\mathbb{Z}^d$-Bernoulli percolation model is introduced at first by Broadbent and Hammersley in order to study the porous media. Here we give its definition in our context: let $(\mathbb{Z}^d, E_d)$ be lattice graph, *the random conductance* $\mathbf{a} : E_d \to \{0\} \cup [\lambda, 1]$ and $\{\mathbf{a}(e)\}_{e\in E_d}$ i.i.d. random variables. We say an edge e is *open* if $\mathbf{a}(e) > 0$ and e is *closed* if $\mathbf{a}(e) = 0$. The connected components defined by the open edges are called *the clusters*, and we denote by $x \longleftrightarrow y$ if x and y are in the same cluster. A specific case $x \longleftrightarrow \infty$ implies an *infinite cluster* $\mathscr{C}_\infty$ containing x. The behavior of the clusters depends on the parameter $\mathfrak{p} := \mathbb{P}[\mathbf{a}(e) > 0]$ and we denote by $\theta(\mathfrak{p}) := \mathbb{P}[0 \longleftrightarrow \infty]$ *the parameter of connectivity.* For $d = 1$ the behavior of clusters is trivial, and for $d \geqslant 2$ there is a *phase transition* in this model: there exists a *critical point* $\mathfrak{p}_c \in (0,1)$ such that

1. *Subcritical phase*: $\mathfrak{p} \in [0, \mathfrak{p}_c)$, there is no infinite cluster and $\theta(\mathfrak{p}) = 0$.
2. *Supercritical phase*: $\mathfrak{p} \in (\mathfrak{p}_c, 1]$, there is a unique infinite cluster $\mathscr{C}_\infty$ and $\theta(\mathfrak{p}) > 0$.
3. *Critical phase*: $\mathfrak{p} = \mathfrak{p}_c$, it is known that $\theta(\mathfrak{p}_c) = 0$ for $d = 2$ and $d \geqslant 11$, but for $3 \leqslant d \leqslant 10$ it is still a conjecture that $\theta(\mathfrak{p}_c) = 0$.

See the monograph [131] and the recent survey [97] for more background about percolation. We are interested in the random walk on the supercritical percolation model. This model can be used to describe the diffusion in porous media, or in two-composite material with high contrast. More precisely, let $\mathfrak{p} > \mathfrak{p}_c(d)$ and we consider *the variable speed random walk* (VSRW), which is a continuous-time Markov jump process $(X_t)_{t\geqslant 0}$ starting from some $y \in \mathscr{C}_\infty$, and associated to *the generator*

$$\mathcal{L}u(x) = \nabla \cdot \mathbf{a}\nabla u(x) := \sum_{z\sim x} \mathbf{a}(\{x,z\})\,(u(z) - u(x))\,. \tag{1.38}$$

We denote *the semigroup* (or *the transition probability*) of the random walk by

$$p\,(t,x,y) = p^{\mathbf{a}}\,(t,x,y) := \mathbb{P}^{\mathbf{a}}_y\,(X_t = x)\,,$$

which is defined as the solution of the parabolic equation

$$\begin{cases} \partial_t p(\cdot,\cdot,y) - \nabla\cdot\mathbf{a}\nabla p(\cdot,\cdot,y) = 0 & \text{in } (0,\infty)\times\mathscr{C}_\infty, \\ p(0,\cdot,y) = \delta_y & \text{in } \mathscr{C}_\infty. \end{cases} \tag{1.39}$$

Due to this characterization, we often refer to the semigroup $p(t,\cdot,y)$ as *the heat kernel* or *the parabolic Green's function.*

We remark that the VSRW defined above is just one possible way to construct the random walk on $\mathscr{C}_\infty$ and there exist other related models. Two of the most common ones are:

1. *The constant speed random walk* (CSRW): it is a continuous-time Markov jump process starting $y \in \mathscr{C}_\infty$, with jump rate 1 and the transition probability

$$P(x,z) = \frac{\mathbf{a}(\{x,z\})}{\sum_{w\sim x}\mathbf{a}(\{x,w\})}. \tag{1.40}$$

   In other words, its associated generator is

$$\mathcal{L}u(x) = \frac{1}{\sum_{w\sim x}\mathbf{a}(\{x,w\})}\sum_{z\sim x}\mathbf{a}(\{x,z\})(u(z)-u(x)).$$

2. *The discrete time random walk* (DTRW): the random walk $(X_n)_{n\in\mathbb{N}}$ is indexed on the integers. It starts from a point $y \in \mathscr{C}_\infty$, and when $X_n = x$, the value of X_{n+1} is chosen randomly among all the neighbors of x following the transition probability (1.40).

These processes have similar, although not identical, properties and have been the subject of interest in the literature.

The random walk on the percolation cluster is one topic among the more general *random conductance models*, where many models belongs to *the Brownian universality*. For example, for the VSRW on $(\mathbb{Z}^d, E_d)$ with $\{\mathbf{a}(e)\}_{e\in E_d}$ i.i.d. satisfying the uniform ellipticity condition $0 < \lambda \leqslant \mathbf{a} \leqslant 1$, then its semigroup $p(t,\cdot,y)$ has a *Gaussian bound*

$$\forall |x-y|\leqslant t, \qquad \frac{C_1}{(2\pi t)^{\frac{d}{2}}}\exp\left(-\frac{|x-y|^2}{2C_1 t}\right) \leqslant p(t,x,y) \leqslant \frac{C_2}{(2\pi t)^{\frac{d}{2}}}\exp\left(-\frac{|x-y|^2}{2C_2 t}\right). \tag{1.41}$$

Moreover, the process almost surely has a scaling limit of Brownian motion in the Skorokhod topology

$$\left(\frac{1}{\sqrt{n}}X_{nt}\right)_{t\geqslant 0} \xRightarrow{n\to\infty} (\bar{\sigma}B_t)_{t\geqslant 0}.$$

This is the *the quenched invariance principle* for $(X_t)_{t\geqslant 0}$, where the environment $\{\mathbf{a}(e)\}_{e\in E_d}$ is fixed in the statement. It is generally easier to establish *the annealed invariance principle* by averaging on the environment.

Let us give some more remarks on these two results. The Gaussian bound for the divergence type operator is initiated by the work of De Giorgi, Moser and Nash on $\mathbb{R}^d$, then it is generalized to manifold by Grigor'yan in [130] and by Saloff-Coste in [208]. For the CSRW on $\mathbb{Z}^d$, its proof can be found in the work of Delmotte [92], where the theorem is known as *"the volume doubling condition and the Poincaré inequality imply the Gaussian bound"*. The condition $|x-y| \leqslant t$ in eq. (1.41) is necessary for the Gaussian bound on $(\mathbb{Z}^d, E_d)$, because the generator is a finite difference operator rather than a differential operator. For the regime $|x-y| \geqslant t$, the tail is exponential rather than Gaussian; see the work of Davies [87]. The quenched invariance principle has a very close link with homogenization theory. One powerful tool to prove it is *the corrector method* initiated by Kozlov in [160]: let $\{\phi_{e_i}\}_{1\leqslant i\leqslant d}$ be the first-order correctors associated to the canonical basis $\{e_i\}_{1\leqslant i\leqslant d}$ and the generator $\nabla\cdot\mathbf{a}\nabla$, then

$$M_t = (X_t\cdot e_1 + \phi_{e_1}(X_t), \cdots, X_t\cdot e_d + \phi_{e_d}(X_t)),$$

is a martingale. Now the martingale convergence theorem [139] applies

$$\left(\frac{1}{\sqrt{n}}M_{nt}\right)_{t\geqslant 0} \overset{n\to\infty}{\Longrightarrow} (\bar{\sigma}B_t)_{t\geqslant 0}.$$

It suffices to prove that the part of corrector vanish almost surely $\frac{\phi_{e_1}(X_{nt})}{\sqrt{n}} \xrightarrow{n\to\infty} 0$ and this is finally reduced to *the sublinearity* of correctors.

The random walk on $\mathscr{C}_\infty$ also belongs to the Brownian universality. One intuitive explanation is that the geometry of $\mathscr{C}_\infty$ is very close to that of $\mathbb{Z}^d$ in large scale. In the case of percolation, in which $\mathbf{a}$ is only allowed to take the values 0 or 1, an annealed invariance principle was proved in [89] by De Masi, Ferrari, Goldstein and Wick. In [211], Sidoravicius and Sznitman proved a quenched invariance principle for the simple random walk in dimension $d \geqslant 4$. This result was extended to every dimension $d \geqslant 2$ by Berger and Biskup in [49] (for the DTRW) and by Mathieu and Piatnitski in [180] (for the CSRW), where their strategy are to construct the correctors on $\mathscr{C}_\infty$. The properties of the heat kernel $p(t,\cdot,y)$ on the infinite cluster have been investigated in the literature. In [181], Mathieu and Remy proved that, almost surely, the heat kernel decays as fast as $t^{-d/2}$. These bounds were extended in [39] by Barlow who established Gaussian lower and upper bounds.

For the VSRW, a similar quenched invariance principle holds. From a homogenization perspective, the diffusivity $\bar{\sigma}$ of the limit Brownian motion is related to the effective coefficient $\bar{\mathbf{a}}$ of the elliptic problems by the identity $\bar{\mathbf{a}} = \frac{1}{2}\theta(\mathfrak{p})\bar{\sigma}^2$.

In the article [41], Barlow and Hambly proved a parabolic Harnack inequality, a local central limit theorem for the CSRW, and bounds on the elliptic Green's function on the infinite cluster. Their main result can be adapted to the case of the VSRW, and reads as follows: if we define, for each $t \geqslant 0$ and $x \in \mathbb{R}^d$,

$$\bar{p}(t,x) := \frac{1}{(2\pi\bar{\sigma}^2 t)^{d/2}} \exp\left(-\frac{|x|^2}{2\,\bar{\sigma}^2 t}\right), \tag{1.42}$$

the heat kernel with diffusivity $\bar{\sigma}$, then, for each time $T > 0$, the following convergence holds, $\mathbb{P}$-almost surely on the event $\{0 \in \mathscr{C}_\infty\}$,

$$\lim_{n\to\infty} \left| n^{d/2} p(nt, g_n^{\mathbf{a}}(x), 0) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x) \right| = 0, \tag{1.43}$$

uniformly in the spatial variable $x \in \mathbb{R}^d$ and in the time variable $t \geqslant T$, where the notation $g_n^{\mathbf{a}}(x)$ means the closest point to $\sqrt{n}x$ in the infinite cluster under the environment $\mathbf{a}$.

As an important tool, the theory of the correctors on $\mathscr{C}_\infty$ is also further developed. In [45], the Liouville regularity problem in a general class of random graphs is studied by Benjamini, Duminil-Copin, Kozma, and Yadin using the entropy method, which confirms the dimension of the first-order correctors and gives the bound for higher order. The complete description of the Liouville regularity on $\mathscr{C}_\infty$ is then given by Armstrong and Dario in [19] by quantitative homogenization method. Dario also gives the optimal estimate of the correctors of the same model in [83]. These results provide us with tools for the AHKM algorithm on percolation clusters mentioned in Section 1.2, and also help us improve the asymptotic local CLT in [41] to a quantitative local CLT result. This is the main contributions in Chapter 4 and will be summarized in Section 1.3.1.

Finally, before stating our contribution, we remark that there are other developments in the random conductance model in the following directions: the relaxation of the i.i.d. condition, the model without uniform ellipticity condition and allowing tails both near ∞ and near 0, the percolation with long-range correlated conductance, etc. For some models among them, there are other universalities (*the anomalous diffusion*) rather than the Gaussian case. We refer Section 4.1.3 and the references there for a complete review.

1.3.1 Summary of Chapter 4

The main contribution of Chapter 4 is convergence rate of the local CLT for the VSRW defined in eq. (1.39). In the following paragraphs, we present this result at first, and then discuss the techniques developed in its proof. Finally, we will also talk about the homogenization of the elliptic Green's function as its corollary.

Theorem 1.3.1 (Main theorem in Chapter 4, Theorem 4.1.1)**.** *For each exponent $\delta > 0$, there exist a positive constant $C < \infty$ and an exponent $s > 0$, depending only on the parameters $d, \lambda, \mathfrak{p}$ and δ, such that for every $y \in \mathbb{Z}^d$, there exists a non-negative random time $\mathcal{T}_{\mathrm{par},\delta}(y)$ satisfying the stochastic integrability estimate*

$$\forall T \geqslant 0, \ \mathbb{P}\left(\mathcal{T}_{\mathrm{par},\delta}(y) \geqslant T\right) \leqslant C \exp\left(-\frac{T^s}{C}\right), \tag{1.44}$$

such that, on the event $\{y \in \mathscr{C}_\infty\}$, for every $x \in \mathscr{C}_\infty$ and every $t \geqslant \max\left(\mathcal{T}_{\mathrm{par},\delta}(y), |x-y|\right)$,

$$\left|p(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y)\right| \leqslant Ct^{-\frac{d}{2}-\left(\frac{1}{2}-\delta\right)} \exp\left(-\frac{|x-y|^2}{Ct}\right). \tag{1.45}$$

We have several remarks on this result.

- In general, the error term has another factor $t^{-\left(\frac{1}{2}-\delta\right)}$ in front of the Gaussian bound, so it is very small in long time compared to both $p(t,x,y)$ and $\bar{p}(t,x-y)$. The exponent is nearly optimal as δ can be arbitrarily small and $t^{-\frac{1}{2}}$ is the optimal rate for the simple random walk on $\mathbb{Z}^d$.

- In eq. (1.45), there is a normalization $\theta(\mathfrak{p})^{-1}$ factor. This is necessary because the semigroup $p(t,\cdot,y)$ only charges $\mathscr{C}_\infty$, and $\theta(\mathfrak{p})$ is nearly the total mass of $\bar{p}(t,\cdot-y)$ on $\mathscr{C}_\infty$ by the density argument

$$\int_{\mathscr{C}_\infty} \bar{p}(t,\cdot-y) \simeq \theta(\mathfrak{p}) \int_{\mathbb{Z}^d} \bar{p}(t,\cdot-y) \simeq \theta(\mathfrak{p}) \int_{\mathbb{R}^d} \bar{p}(t,\cdot-y) \simeq \theta(\mathfrak{p}).$$

- The result eq. (1.45) only holds for $t \geqslant \max\left(\mathcal{T}_{\mathrm{par},\delta}(y), |x-y|\right)$, here we give its reason. This condition can be decomposed as

$$\left\{t \geqslant \max\left(\mathcal{T}_{\mathrm{par},\delta}(y), |x-y|\right)\right\} = \left\{t \geqslant |x-y|\right\} \cap \left\{t \geqslant \mathcal{T}_{\mathrm{par},\delta}(y)\right\}.$$

  Recall that $p(t,x,y)$ has an exponential tail for $|x-y| \geqslant t$ instead of a Gaussian tail, so the comparison is not true in that regime. The condition $t \geqslant \mathcal{T}_{\mathrm{par},\delta}(y)$ can be interpreted as a random waiting time to let the random walker explore the percolation. As we know, in a small scale the configuration of the percolation can be quite zigzag and fractal, so the semigroup has not converged close enough to the Gaussian. Moreover, the waiting time $\mathcal{T}_{\mathrm{par},\delta}(y)$ is not very large since from eq. (1.44) its typical size is a constant and has a sub-exponential tail.

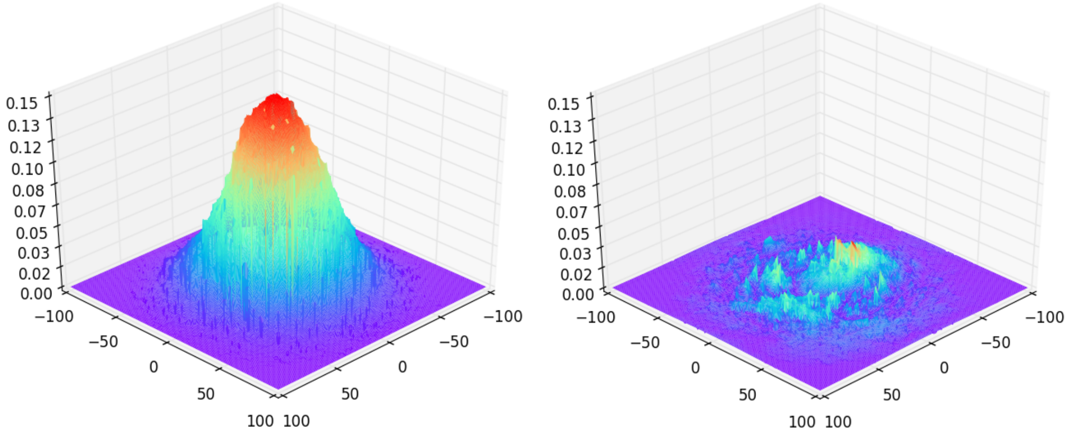

Figure 1.2: The figure on the left represents the density distribution of the function $t^{\frac{d}{2}}p(t,\cdot,0)$ where the map p is the 2-dimensional heat kernel on the infinite percolation cluster with probability $\mathfrak{p} = 0.7$ at time $t = 1000$; it is similar to a Gaussian distribution. The figure on the right is the error between the map $t^{\frac{d}{2}}p(t,\cdot,0)$ and the normalized Gaussian heat kernel $\theta(p)^{-1}t^{\frac{d}{2}}\bar{p}(t,\cdot)$ defined in eq. (1.42); it is small compared to density distribution on the left.

The proof of this result relies on quantitative homogenization theory on percolation clusters, and we also develop some new techniques. The following is a list of the main ingredients.

1. *A partition of good cubes*: [19] developed a Calderón-Zygmund type partition of the cubes, such that

   (a) There exists a collection $\mathcal{P}$ of triadic cubes, $\mathbb{Z}^d = \bigsqcup_{\square\in\mathcal{P}} \square$.

   (b) In every *partition cube* $\square \in \mathcal{P}$, there exists a *maximal cluster* $\mathscr{C}_*(\square)$. The infinite cluster $\mathscr{C}_\infty$ has the structure $\mathscr{C}_\infty = \bigsqcup_{\square\in\mathcal{P}} \mathscr{C}_*(\square)$.

   (c) The size (diameter) of the partition cube has an estimate $\mathrm{size}(\square) \leqslant \mathcal{O}_1(C)$.

   Therefore, we can use this technique to do the localization from $\mathscr{C}_\infty$ to every small cluster $\mathscr{C}_*(\square)$, and its geometry is not very far from the cube $\square$ containing it. These allow us to develop the functional inequalities including the Poincaré inequality and the Meyers inequality on percolation clusters. The construction of this partition of good cubes is inspired by the work of Pisztora [201], and some similar idea is also used by Barlow in the proof of the Gaussian bound in [39].

2. *The estimate of correctors and two-scale expansion*: the optimal estimate of the correctors on $\mathscr{C}_\infty$ is proved in [83] and its application to the two-scale expansion on percolation clusters is implemented in [134] (see Chapter 3) for

   $$-\nabla\cdot\mathbf{a}_{\mathscr{C}}\nabla u = -\nabla\cdot\bar{\mathbf{a}}\nabla\bar{u} \qquad \text{in } \mathbb{Z}^d, \tag{1.46}$$

   where $\mathbf{a}_{\mathscr{C}}(\{x,y\}) = \mathbf{a}(\{x,y\})\mathbf{1}_{\{x,y\in\mathscr{C}_\infty\}}$. The two-scale expansion for eq. (1.46) is rather unusual, because its left-hand side is supported on $\mathscr{C}_\infty$, while the right-hand side is supported on $\mathbb{Z}^d$. Thus we need the argument of the partition of good cubes. In the proof of Theorem 1.3.1, we treat a similar but more general case

   $$\begin{aligned}(\partial_t - \nabla\cdot\mathbf{a}\nabla)u = 0 \qquad &(0,\infty)\times\mathscr{C}_\infty,\\ \left(\partial_t - \frac{1}{2}\bar{\sigma}^2\Delta\right)\bar{u} = 0 \qquad &(0,\infty)\times\mathbb{R}^d,\end{aligned} \tag{1.47}$$

with suitable coherent boundary condition; see Theorem 4.3.2 for details. Here the first equation in eq. (1.47) is defined on $\mathscr{C}_\infty$ and $-\nabla\cdot\mathbf{a}\nabla$ is a finite difference operator; the second equation in eq. (1.47) is defined on $\mathbb{R}^d$ and Δ is the standard Laplace operator. Unlike eq. (1.46), we have no canonical way to combine the two equations into one. For this reason, besides the technique of the partition of good cubes, we also apply *the Whitney decomposition* from harmonic analysis to overcome the obstacle.

3. *The estimate of flux*: the *centered flux* $\mathbf{g}_{e_k}$ defined by

$$\mathbf{g}_{e_k}:\mathbb{Z}^d\to\mathbb{R}^d,\qquad \mathbf{g}_{e_k}:=\mathbf{a}_{\mathscr{C}}(\mathcal{D}\phi_{e_k}+e_k)-\bar{\mathbf{a}}e_k, \tag{1.48}$$

is also an important quantity in quantitative homogenization theory. Its role and estimate are very similar to $\nabla\phi_{e_k}$. In Chapter 3 we develop the weak norm estimate for $\mathbf{g}_{e_k}$, but in the proof of Chapter 4 we use another similar quantity $\widetilde{\mathbf{g}}_{e_k}$

$$\widetilde{\mathbf{g}}_{e_k}:\mathscr{C}_\infty\to\mathbb{R}^d,\qquad \widetilde{\mathbf{g}}_{e_k}:=\mathbf{a}(\mathcal{D}\phi_{e_k}+e_k)-\frac{1}{2}\bar{\sigma}^2e_k, \tag{1.49}$$

and we give its H^{-1} estimate in Proposition 4.B.1. Recall the identity $\bar{\mathbf{a}}=\frac{1}{2}\theta(\mathfrak{p})\bar{\sigma}^2$, the main difference between eq. (1.48) and eq. (1.49) is in fact a difference of cluster density. As we need a quantitative estimate for $\widetilde{\mathbf{g}}_{e_k}$, we prove a concentration of the cluster density in Proposition 4.A.1

$$\left|\frac{|\mathscr{C}_\infty\cap\square_m|}{|\square_m|}-\theta(\mathfrak{p})\right|\leqslant\mathcal{O}_{\frac{2(d-1)}{3d^2+2d-1}}\left(C3^{-\frac{dm}{2}}\right). \tag{1.50}$$

This estimate is more explicit than the large deviation results that were available in the estimate.

The local CLT Theorem 1.3.1 also implies the quantitative homogenization for the elliptic Green's function on the infinite cluster. In dimension $d\geqslant 3$, given an environment $\{\mathbf{a}(e)\}_{e\in E_d}$ and a point $y\in\mathscr{C}_\infty$, we define the Green's function $g(\cdot,y)$ as the solution of the equation

$$-\nabla\cdot\mathbf{a}\nabla g(\cdot,y)=\delta_y \text{ in } \mathscr{C}_\infty \text{ such that } g(x,y)\overset{x\to\infty}{\longrightarrow}0.$$

This function exists, is unique almost surely and is related to the semigroup p through the identity

$$g(x,y)=\int_0^\infty p(t,x,y)\,dt. \tag{1.51}$$

In dimension 2, the situation is different since the Green's function is not bounded at infinity, and we define $g(\cdot,y)$ as the unique function which satisfies

$$-\nabla\cdot\mathbf{a}\nabla g(\cdot,y)=\delta_y \text{ in } \mathscr{C}_\infty,\ \frac{1}{|x|}g(x,y)\overset{x\to\infty}{\longrightarrow}0 \text{ and } g(y,y)=0.$$

This function is related to the transition probability p through the identity

$$g(x,y)=\int_0^\infty\left(p(t,x,y)-p(t,y,y)\right)dt.$$

In the statement below, we denote by $\bar{g}$ the homogenized Green's function defined by the formula, for each point $x\in\mathbb{R}^d\setminus\{0\}$,

$$\bar{g}(x):=\begin{cases}-\frac{1}{\pi\bar{\sigma}^2\theta(\mathfrak{p})}\ln|x| & \text{if } d=2,\\ \frac{\Gamma(d/2-1)}{\left(2\pi^{d/2}\bar{\sigma}^2\theta(\mathfrak{p})\right)}\frac{1}{|x|^{d-2}} & \text{if } d\geqslant 3,\end{cases} \tag{1.52}$$

where the symbol Γ denotes the standard Gamma function.

Theorem 1.3.2 (Main theorem in Chapter 4, Theorem 4.1.2)**.** *For each exponent $\delta > 0$, there exist a positive constant $C < \infty$ and an exponent $s > 0$, depending only on the parameters $d, \lambda, \mathfrak{p}$ and δ, such that for every $y \in \mathbb{Z}^d$, there exists a non-negative random variable $\mathcal{M}_{\mathrm{ell},\delta}(y)$ satisfying*

$$\forall R \geqslant 0, \ \mathbb{P}\left(\mathcal{M}_{\mathrm{ell},\delta}(y) \geqslant R\right) \leqslant C \exp\left(-\frac{R^s}{C}\right),$$

such that, on the event $\{y \in \mathscr{C}_\infty\}$:

1. *In dimension $d \geqslant 3$, for every point $x \in \mathscr{C}_\infty$ satisfying $|x-y| \geqslant \mathcal{M}_{\mathrm{ell},\delta}(y)$,*

$$|g(x,y) - \bar{g}(x-y)| \leqslant \frac{1}{|x-y|^{1-\delta}} \frac{C}{|x-y|^{d-2}}. \tag{1.53}$$

2. *In dimension 2, the limit*

$$K(y) := \lim_{x \to \infty} \left(g(x,y) - \bar{g}(x-y)\right),$$

exists, is finite almost surely and satisfies the stochastic integrability estimate

$$\forall R \geqslant 0, \ \mathbb{P}\left(|K(y)| \geqslant R\right) \leqslant C \exp\left(-\frac{R^s}{C}\right).$$

Moreover, for every point $x \in \mathscr{C}_\infty$ satisfying $|x-y| \geqslant \mathcal{M}_{\mathrm{ell},\delta}(y)$,

$$|g(x,y) - \bar{g}(x-y) - K(y)| \leqslant \frac{C}{|x-y|^{1-\delta}}. \tag{1.54}$$

1.4 Homogenization for interacting particle systems

Another topic studied in this thesis is the homogenization theory for interacting particle systems, which corresponds to Chapters 5 and 6 and is summarized in Section 1.4.1. Since the context is a little different from the classical homogenization, we give at first a brief review of some classical particle models to make our motivations clearer.

In the previous models, the random walker in random environment, which can also be seen as the evolution of one particle, will be close to the Brownian motion in large scale and long time. The interacting particle systems share the similar spirit, while in these models we have infinitely many particles instead of one, and the random environment comes from their configuration which is dynamic.

The most studied model is *the lattice gas* and a basic model is the *simple symmetric exclusion process (SSEP)*: let $\eta : \mathbb{Z}^d \to \{0,1\}$ stand the configuration of particles, where every site allows at most one particle. In the evolution, every particle has rate $\frac{1}{2}$ to jump to a vacant neighbor. Thus the evolution $(\eta_t)_{t \geqslant 0}$ follows the generator

$$\mathcal{L} f(\eta) = \frac{1}{2} \sum_{x \in \mathbb{Z}^d} \sum_{y \sim x} \eta(x)(1-\eta(y)) \left(f(\eta^{x,y}) - f(\eta)\right), \tag{1.55}$$

with the notation

$$\eta^{x,y}(z) = \begin{cases} \eta(z) & z \neq x, y; \\ \eta(y) & z = x; \\ \eta(x) & z = y. \end{cases} \tag{1.56}$$

Here the test function in eq. (1.55) is the local function $f \in C_0$ which only depends on finite site of η. In this model, the product Bernoulli measure $\mathrm{Ber}(\alpha)^{\otimes \mathbb{Z}^d}$ with $\alpha \in (0,1)$ is a stationary measure and we denote it by $\mathbb{P}_\alpha$.

The long-time and large-scale behavior of the SSEP can be characterized by *the hydrodynamic limit* and *the equilibrium fluctuation*. We denote by π_t^N the empirical density of the configuration

$$\pi_t^N := N^{-d} \sum_{x \in \mathbb{Z}^d} \eta_{N^2 t}(x) \delta_{x/N}, \tag{1.57}$$

The hydrodynamic limit tells us $(\pi_t^N)_{t \geqslant 0} \xrightarrow{N \to \infty} (\rho_t)_{t \geqslant 0}$, i.e. the empirical density converges to the solution of the heat equation

$$\partial_t \rho_t = \frac{1}{2} \Delta \rho_t, \tag{1.58}$$

in the Skorokhod topology of the Schwartz distribution, provided the initial configuration $\pi_0^N \xrightarrow{N \to \infty} \rho_0$ has a limit profile ρ_0. If η_0 starts from the stationary measure $\mathbb{P}_\alpha$, then the equilibrium fluctuation theorem says

$$Y_t^N := N^{-\frac{d}{2}} \sum_{x \in \mathbb{Z}^d} \left(\eta_{N^2 t}(x) - \alpha\right) \delta_{x/N}, \tag{1.59}$$

converges to the functional Ornstein–Uhlenbeck process $(Y_t)_{t \geqslant 0}$ solving

$$\mathrm{d}Y_t = \frac{1}{2} \Delta Y_t \, \mathrm{d}t + \sqrt{\alpha(1-\alpha)} \nabla \mathrm{d}B_t, \tag{1.60}$$

where B_t is the space-time white noise.

In these results, the effective coefficient matrix is the identity, because the flux $W_{x,x+e_i}$ from x to $x + e_i$ in SSEP is

$$W_{x,x+e_i} = \frac{1}{2} \left(\eta(x) - \eta(x+e_i)\right), \tag{1.61}$$

and it can be written as the difference $W_{x,x+e_i} = \tau_x h(\eta) - \tau_{x+e_i} h(\eta)$ with $h(\eta) = \eta(0)$, where τ_x is the translation operator. This property is *the gradient condition*. It makes the effective coefficient trivial in the model and it is only valid in some particle systems. Another more heuristic explanation is that the SSEP can be treated as if the jump is always permitted because in eq. (1.55)

$$\sum_{x \in \mathbb{Z}^d} \sum_{y \sim x} \eta(x)(1-\eta(y)) \left(f(\eta^{x,y}) - f(\eta)\right) = \sum_{x \in \mathbb{Z}^d} \sum_{y \sim x} \left(f(\eta^{x,y}) - f(\eta)\right).$$

To make this model less specific, a *generalized symmetric exclusion process (GSEP)* is proposed, where every site in $\mathbb{Z}^d$ can place at most κ particles ($\kappa \geqslant 2$), i.e. $\widetilde{\eta} : \mathbb{Z}^d \to \{0, 1, \cdots, \kappa\}$ and the generator is

$$\mathcal{L} f(\widetilde{\eta}) = \frac{1}{2} \sum_{x \in \mathbb{Z}^d} \sum_{y \sim x} \mathbf{1}_{\{\widetilde{\eta}(x) > 0, \widetilde{\eta}(y) < \kappa\}} \left(f(\widetilde{\eta}^{x,y}) - f(\widetilde{\eta})\right), \tag{1.62}$$

with the notation

$$\widetilde{\eta}^{x,y}(z) = \begin{cases} \widetilde{\eta}(z) & z \neq x, y; \\ \widetilde{\eta}(x) - 1 & z = x; \\ \widetilde{\eta}(y) + 1 & z = y. \end{cases} \tag{1.63}$$

In this model, the stationary measure is $\mathbb{P}_\alpha = \nu_\alpha^{\otimes \mathbb{Z}^d}$ with

$$\forall n \in \{0, 1, \cdots, \kappa\}, \qquad \nu_\alpha(n) = \frac{\alpha^n}{\sum_{j=0}^{\kappa} \alpha^j}.$$

This model does not satisfy the gradient condition, since

$$W_{x,x+e_i} = \frac{1}{2}\left(\mathbf{1}_{\{\widetilde{\eta}(x)>0, \widetilde{\eta}(x+e_i)<\kappa\}} - \mathbf{1}_{\{\widetilde{\eta}(x+e_i)>0, \widetilde{\eta}(x)<\kappa\}}\right), \tag{1.64}$$

cannot be written as the difference $W_{x,x+e_i} = \tau_x h(\widetilde{\eta}) - \tau_{x+e_i} h(\widetilde{\eta})$ for some local function $h \in C_0$. In its result of hydrodynamic limit and fluctuation theorem,

$$\partial_t \rho_t = \nabla \cdot D(\rho_t) \nabla \rho_t, \qquad \mathrm{d}Y_t = \nabla \cdot D(\alpha) \nabla Y_t \,\mathrm{d}t + \sqrt{\bar{\mathbf{a}}(\alpha)} \nabla \mathrm{d}W_t, \tag{1.65}$$

we will see a quantity called *the bulk diffusion coefficient* (or *the self-diffusion coefficient*) $D(\alpha)$ defined by

$$D(\alpha) := \frac{\bar{\mathbf{a}}(\alpha)}{2\chi(\alpha)}, \tag{1.66}$$

where χ is the quantity called *the compressibility*

$$\chi(\alpha) := \mathbb{V}\mathrm{ar}_\alpha[\widetilde{\eta}(0)], \tag{1.67}$$

and the quantity $\bar{\mathbf{a}}$ has a variational description. We denote by $\Gamma_f(\widetilde{\eta}) := \sum_{x \in \mathbb{Z}^d} \tau_x f(\widetilde{\eta})$, which may be infinite but for any local function $f \in C_0$

$$\nabla_{0,e_i} \Gamma_f(\widetilde{\eta}) := \Gamma_f(\widetilde{\eta}^{0,e_i}) - \Gamma_f(\widetilde{\eta}), \tag{1.68}$$

is well-defined as the jump only changes the value of finite terms in $\sum_{x \in \mathbb{Z}^d} \tau_x f(\widetilde{\eta})$. Then $\bar{\mathbf{a}}$ is defined by

$$p \cdot \bar{\mathbf{a}}(\alpha) p = \inf_{f \in C_0} \sum_{i=1}^{d} \mathbb{E}_\alpha \left[\mathbf{1}_{\{\widetilde{\eta}(0)>0, \widetilde{\eta}(e_i)<\kappa\}} (p_i + \nabla_{0,e_i} \Gamma_f(\widetilde{\eta}))^2\right]. \tag{1.69}$$

In the definition of the bulk diffusion matrix, the quantity $\bar{\mathbf{a}}$ looks very similar as the effective coefficient in stochastic homogenization theory, where the formula eq. (1.68) is used to construct some stationary gradient field. Thus it is very natural to think if we can use some finite volume approximation in eq. (1.19) and get its rate of convergence for $D(\alpha)$. This may provide us quantitative results in particle systems. However, we should notice that eq. (1.68) is defined for the configuration space, thus the function $\nabla_{0,e_i} \Gamma_f(\widetilde{\eta})$ can have arbitrarily many coordinates. This is one of the major challenges and we will discuss it in detail in Section 1.4.1.

We also give a short review of the references for the results mentioned above. There are two classical approaches to the identification of the hydrodynamic limit. The first, called *the entropy method*, was introduced in [136], and extended to certain non-gradient models in [221, 204]. The second, called *the relative entropy method*, was introduced in [224], and was extended to a non-gradient model in [111].

The asymptotic description of the fluctuations of interacting particle systems at equilibrium has been obtained in [66, 214, 91, 69, 71], where the main tool is *the Holley-Strook theorem* [140]. The extension of this result to non-gradient models was obtained in [174, 70, 110]. We are not aware of any results concerning the non-equilibrium fluctuations of a non-gradient model. For gradient models (or small perturbations thereof), we refer in particular to [202, 90, 106, 71, 144].

The work [166] gives a proof that finite-volume approximations of the self-diffusion matrix converge to the correct limit. However, no rate of convergence could be obtained there. The qualitative result of [166] was extended to the mean-zero simple exclusion process, and to the asymmetric simple exclusion process in dimension $d \geqslant 3$, in [143]. Finally, we also refer to the books [215, 152, 157] for much more thorough expositions on these topics, and reviews of the literature.

1.4.1 Summary of Chapters 5 and 6

In Chapters 5 and 6, we aim to develop a quantitative homogenization theory for interacting particle systems of non-gradient type. Our main contributions are a $t^{-\frac{d}{2}}$ Gaussian type decay for the semigroup, see Theorem 1.4.1, and a convergence rate for the finite volume approximation of the bulk coefficient, see Theorem 1.4.2. Our model is constructed in *the continuum configuration space*, but the results and proofs can be adapted in the classical non-gradient type lattice gas model, for example the GSEP.

We introduce at first our particle system. Let $\mathcal{M}_\delta(\mathbb{R}^d)$ be the set of σ-finite measures that are sums of Dirac masses on $\mathbb{R}^d$, which we think of as the space of configurations of particles. We denote by $\mathbb{P}_\rho$ the law on $\mathcal{M}_\delta(\mathbb{R}^d)$ of the Poisson point process of density $\rho \in (0,\infty)$, with $\mathbb{E}_\rho, \mathbb{V}\mathrm{ar}_\rho$ the associated expectation and variance. We denote by $\mathcal{F}_U$ the σ-algebra generated by the mappings $V \mapsto \mu(V)$, for all Borel sets $V \subseteq U$, completed with all the $\mathbb{P}_\rho$-null sets, and we set $\mathcal{F} := \mathcal{F}_{\mathbb{R}^d}$. We give ourselves a function $\mathbf{a}_\circ : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, where $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ is the set of d-by-d symmetric matrices. We assume that this mapping satisfies the following properties:

- *uniform ellipticity*: there exists $\Lambda < \infty$ such that for every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$,

$$\forall \xi \in \mathbb{R}^d, \quad |\xi|^2 \leqslant \xi \cdot \mathbf{a}_\circ(\mu)\xi \leqslant \Lambda|\xi|^2 \, ; \tag{1.70}$$

- *finite range of dependence*: denoting by B_1 the Euclidean ball of radius 1 centered at the origin, we assume that $\mathbf{a}_\circ$ is $\mathcal{F}_{B_1}$-measurable.

We denote by $\tau_{-x}\mu$ the translation of the measure μ by the vector $-x \in \mathbb{R}^d$; explicitly, for every Borel set U, we have $(\tau_{-x}\mu)(U) = \mu(x+U)$. We extend $\mathbf{a}_\circ$ by stationarity by setting, for every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$,

$$\mathbf{a}(\mu, x) := \mathbf{a}_\circ(\tau_{-x}\mu).$$

Denoting by $\mu_t := \sum_{i=1}^\infty \delta_{x_{i,t}}$ the configuration at time $t \geqslant 0$, our model can be informally described as an infinite-dimensional system with local interaction such that every particle $x_{i,t}$ evolves as a diffusion associated to the divergence-form operator $-\nabla \cdot \mathbf{a}(\mu_t, x_{i,t})\nabla$. More precisely, it is a Markov process $(\Omega, (\mathscr{F}_t)_{t\geqslant 0}, \mathbb{P}_\rho)$ defined by *the Dirichlet form*

$$\mathcal{E}^{\mathbf{a}}(f,f) := \mathbb{E}_\rho\left[\int_{\mathbb{R}^d} \nabla f(\mu,x) \cdot \mathbf{a}(\mu,x) \nabla f(\mu,x) \, \mathrm{d}\mu(x)\right], \tag{1.71}$$

where the directional derivative

$$e_k \cdot \nabla f(\mu, x) = \lim_{h\to 0} \frac{f(\mu - \delta_x + \delta_{x+he_k}) - f(\mu)}{h}, \tag{1.72}$$

is defined for a family of suitable functions and $x \in \mathrm{supp}(\mu)$. The construction of similar diffusion processes can be found in the the previous work by Albeverio, Kondratiev and Röckner in [2, 3, 4, 5]; see also the survey [206].

We need some more explanations for the test function of the Dirichlet form in eq. (1.71). For any open set $U \subseteq \mathbb{R}^d$, we denote by the space $\mathscr{C}_c^\infty(U)$ the functions that are $\mathcal{F}_K$-measurable for some compact set $K \subseteq U$, and smooth with respect to every particle. Therefore, the function space $\mathscr{C}_c^\infty(U)$ plays the same role as the local function in the lattice gas model. Then we define the norm $\mathscr{H}^1(U)$, an infinite-dimensional analogue of the classical Sobolev space H^1 that

$$\|f\|_{\mathscr{H}^1(U)} = \left(\mathbb{E}_\rho[f^2(\mu)] + \mathbb{E}_\rho\left[\int_U |\nabla f(\mu,x)|^2 \,\mathrm{d}\mu(x)\right]\right)^{\frac{1}{2}}. \tag{1.73}$$

We also define the space $\mathscr{H}_0^1(U)$ as the closure in $\mathscr{H}^1(U)$ of functions $f \in \mathscr{C}_c^\infty(U)$ such that $\|f\|_{\mathscr{H}^1(U)}$ is finite and this is the suitable function space for eq. (1.71).

The main theorem in Chapter 5 is an estimate for the variance decay for our particle system $(\mu_t)_{t\geqslant 0}$. We denote by $\mathscr{L}^p$ the L^p space in $(\Omega, \mathcal{F}, \mathbb{P}_\rho)$ for $p \geqslant 1$. Let $u : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}$ be an $\mathcal{F}_{Q_{l_u}}$-measurable function with $Q_{l_u} := \left(-\frac{l_u}{2}, \frac{l_u}{2}\right)^d$, and let $u_t(\mu_0) := \mathbb{E}_\rho[u(\mu_t)|\mu_0]$.

Theorem 1.4.1 (Main theorem in Chapter 5, Theorem 5.1.1)**.** *There exist two finite positive constants* $\gamma := \gamma(\rho, d, \Lambda)$, $C := C(\rho, d, \Lambda)$ *such that for any* $u \in \mathscr{C}_c^\infty(\mathbb{R}^d) \cap \mathscr{L}^\infty$ *which is* $\mathcal{F}_{Q_{l_u}}$*-measurable, then we have*

$$\mathbb{V}\mathrm{ar}_\rho[u_t] \leqslant C(1 + |\log t|)^\gamma \left(\frac{1 + l_u}{\sqrt{t}}\right)^d \|u\|^2_{\mathscr{L}^\infty}. \tag{1.74}$$

Remark. In the work [2, 3, 4, 5] of Albeverio, Kondratiev and Röckner, the Dirichlet form $\mathcal{E}^{\mathbf{a}}$ is defined on the function space

$$\mathcal{F}\mathscr{C}_c^\infty(\mathbb{R}^d) = \left\{G(\mu(g_1), \cdots \mu(g_n)) : n \in \mathbb{N}, G \in C_b^\infty(\mathbb{R}^{nd}), g_i \in C_c^\infty(\mathbb{R}^d)\right\}. \tag{1.75}$$

This is a concrete subspace of $\mathscr{C}_c^\infty(\mathbb{R}^d)$ and one can prove that $\mathscr{C}_c^\infty(\mathbb{R}^d)$ and $\mathcal{F}\mathscr{C}_c^\infty(\mathbb{R}^d)$ generate the same $\mathscr{H}_0^1(\mathbb{R}^d)$.

The proof of Theorem 1.4.1 takes inspiration from an important work [142] by Janvresse, Landim, Quastel and Yau, where the decay of variance is proved in *the* $\mathbb{Z}^d$ *zero range model*. We extend this proof to non-gradient model in the continuum configuration space, and a technical difficulty in this generalization is a key localization estimate: we denote by $\overline{Q}_K = \left[-\frac{K}{2}, \frac{K}{2}\right]^d$ the closed cube and recall that $\mathcal{F}_{\overline{Q}_K}$ represents the information of μ in it. We define $\mathsf{A}_K u_t := \mathbb{E}_\rho[u_t|\mathcal{F}_{\overline{Q}_K}]$, then for every $t \geqslant \max\left\{(l_u)^2, 16\Lambda^2\right\}$ and $K \geqslant \sqrt{t}$ we have

$$\mathbb{E}_\rho\left[(u_t - \mathsf{A}_K u_t)^2\right] \leqslant C(\Lambda) \exp\left(-\frac{K}{\sqrt{t}}\right) \mathbb{E}_\rho\left[u^2\right]. \tag{1.76}$$

This is a key estimate appearing in [142, Proposition 3.1], and is also natural as $\sqrt{t}$ is the typical scale of diffusion, thus when $K \gg \sqrt{t}$ one gets very good approximation in eq. (1.76). The main idea of its proof is to define a multiscale functional

$$\begin{aligned} S_{k,K,\beta}(f) &:= \alpha_k \mathbb{E}_\rho\left[(\mathsf{A}_k f)^2\right] + \int_k^K \alpha_s \, d\mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right] + \alpha_K \mathbb{E}_\rho\left[(f - \mathsf{A}_K f)^2\right] \\ &= \alpha_K \mathbb{E}_\rho\left[f^2\right] - \int_k^K \alpha_s' \mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right] ds, \end{aligned}$$

with $\alpha_s = \exp\left(\frac{s}{\beta}\right), \alpha'_s = \frac{\mathrm{d}}{\mathrm{d}s}\alpha_s, \beta > 0$, and then studies the evolution of $\frac{\mathrm{d}}{\mathrm{d}t}S_{k,K,\beta}(u_t)$. In this procedure, one may argue that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right] = 2\mathcal{E}^{\mathbf{a}}(u_t, \mathsf{A}_s u_t),$$

but $\mathsf{A}_s u_t$ is not in the test function $\mathscr{H}_0^1(\mathbb{R}^d)$ because of the perturbation at ∂Q_s. More precisely, it means the discontinuity of $\mathsf{A}_s u_t$ when a particle enters or exits Q_s. In the lattice model, there is also such perturbation at the boundary, but $\mathsf{A}_s u_t$ can still be used as the local function thanks to the discrete difference. To solve this problem, we use a regularization version of A_s

$$\mathsf{A}_{s,\varepsilon} f := \frac{1}{\varepsilon}\int_0^\varepsilon \mathsf{A}_{s+r} f \,\mathrm{d}r, \tag{1.77}$$

to make the conditional expectation more smooth. Moreover, the derivative of $\mathsf{A}_{s,\varepsilon} u_t$ near the boundary has a close link with the $\mathscr{L}^2$ isometry of the martingale $(\mathsf{A}_s u_t)_{s\geqslant 0}$.

Finally, let us remark that [142] also obtains the long-time limit $\mathbb{V}\mathrm{ar}_\rho[u_t] = Ct^{-\frac{d}{2}} + o(t^{-\frac{d}{2}})$. However, the zero range model has the gradient condition, so the constant C is easier to calculate. In our model, we have to at first identify the bulk coefficient and this also motivates us for the work in Chapter 6.

In Chapter 6, we study the finite volume approximation of the bulk diffusion coefficient. For every bounded open set $U \subseteq \mathbb{R}^d$, we define the matrix $\bar{\mathbf{a}}(U) \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ to be such that, for every $p \in \mathbb{R}^d$,

$$\frac{1}{2}p\cdot\bar{\mathbf{a}}(U)p := \inf_{\phi\in\mathscr{H}_0^1(U)} \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}(p+\nabla\phi(\mu,x))\cdot\mathbf{a}(\mu,x)(p+\nabla\phi(\mu,x))\,\mathrm{d}\mu(x)\right]. \tag{1.78}$$

For every $m \in \mathbb{N}$, we let $\square_m = Q_{3^m}$ denote the cube of side length 3^m. We define *the effective coefficient matrix* as $\bar{\mathbf{a}} := \lim_{m\to\infty}\bar{\mathbf{a}}(\square_m)$, and the main theorem is to prove its convergence rate.

Theorem 1.4.2 (Main theorem in Chapter 6, Theorem 6.2.1). *The limit $\bar{\mathbf{a}}$ is well-defined. Moreover, there exist an exponent $\alpha(d,\Lambda,\rho) > 0$ and a constant $C(d,\Lambda,\rho) < \infty$ such that for every $m \in \mathbb{N}$,*

$$|\bar{\mathbf{a}}(\square_m) - \bar{\mathbf{a}}| \leqslant C3^{-\alpha m}. \tag{1.79}$$

The proof of Theorem 1.4.2 follows the renormalization approach initiated by Armstrong and Smart in [31], which is also reviewed in Section 1.1. However, notice the function space $\mathscr{H}_0^1(U)$ in eq. (1.78) is quite different from the Euclidean case: the function is defined in the configuration $\mu = \sum_{i=1}^\infty \delta_{x_i}$ instead of $\mathbb{R}^d$, so its number of coordinates can be arbitrarily large. In the following paragraphs, we point out some new ideas when we implement the renormalization approach in particle systems.

1. *Good function spaces for subadditive quantities*: we denote by $\nu(U,p) = \frac{1}{2}p\cdot\bar{\mathbf{a}}(U)p$ and can check that it is subadditive. We also hope to construct a subadditive dual quantity $\nu^*(U,q)$, and we propose the formula (see the discussion in eq. (1.20))

$$\nu^*(U,q) := \sup_{u\in\mathscr{H}^1(U)} \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(-\frac{1}{2}\nabla u\cdot\mathbf{a}\nabla u + q\cdot\nabla u\right)\mathrm{d}\mu\right]. \tag{1.80}$$

However, notice that we have not defined the function space $\mathscr{H}^1(U)$, although the norm is defined in eq. (1.73). A good definition should be the largest class of $\mathcal{F}$-measurable functions with finite $\mathscr{H}^1(U)$ norm. Informally speaking, these functions are differentiable with respect to the particles in U, but the dependence of the particles outside U is just measurable. In particular, unlike the function space $\mathscr{H}^1_0(U)$, we do *not* require the $\mathcal{F}_U$-measurable condition for $\mathscr{H}^1(U)$. To see this definition is good, one can check:

(a) For any $V \subseteq U$, we have $\mathscr{H}^1_0(V) \subseteq \mathscr{H}^1_0(U)$ and $\mathscr{H}^1(U) \subseteq \mathscr{H}^1(V)$. That is the property to prove the subadditivity of ν and ν^*.

(b) Let $B_1(U)$ be the neighborhood containing U with distance 1. Then $\mathbb{E}_\rho[u|\mathcal{F}_{B_1(U)}]$ is a better candidate than u in the functional of $\nu^*(U,q)$. This helps us recover the mixing condition of the maximiser of $\nu^*(U,q)$.

2. *The modified Caccioppoli inequality*: another important ingredient is the Caccioppoli inequality as the optimisers of ν and ν^* are $\mathbf{a}$-harmonic functions. We recall the classical Caccioppoli inequality: for every $\widetilde{u}$ such that $\Delta\widetilde{u} = 0$ in Q_{3r},

$$\int_{Q_r} |\nabla \widetilde{u}|^2 \leqslant \frac{C}{r^2} \int_{Q_{3r}} |\widetilde{u}|^2. \tag{1.81}$$

Its proof is to use a cut-off function $\psi \in C_c^\infty(Q_{3r})$ such that $\psi^2\widetilde{u} \in H^1_0(Q_{3r})$ and then test $\psi^2\widetilde{u}$ against $\Delta\widetilde{u}$. In our particle system, the analogue $\mathbf{a}$-harmonic function is

$$\mathcal{A}(U) := \left\{ u \in \mathscr{H}^1(U) : \ \forall \varphi \in \mathscr{H}^1_0(U), \ \mathbb{E}_\rho\left[\int_U \nabla u \cdot \mathbf{a}\nabla\varphi \, \mathrm{d}\mu\right] = 0 \right\},$$

and we hope to prove a similar result as eq. (1.81). There is no direct counterpart of the cut-off function ψ, but inspired from eq. (1.77), for any $u \in \mathcal{A}(Q_{3r})$, we can use $\mathsf{A}_{r,\varepsilon}u \in \mathscr{H}^1_0(Q_r)$ as a test function. However, in spite of many efforts, the best we can prove is a modified Caccioppoli inequality Proposition 6.3.6: there exist $\theta(d,\Lambda) \in (0,1)$, $C(d,\Lambda) < \infty$, and $R_0(d,\Lambda) < \infty$ such that for every $r \geqslant R_0$ and $u \in \mathcal{A}(Q_{3r})$, we have

$$\mathbb{E}_\rho\left[\frac{1}{\rho|Q_r|}\int_{Q_r} \nabla(\mathsf{A}_{r+2}u)\cdot\mathbf{a}\nabla(\mathsf{A}_{r+2}u)\,\mathrm{d}\mu\right] \leqslant \frac{C}{r^2\rho|Q_{3r}|}\mathbb{E}_\rho[u^2] + \theta\mathbb{E}_\rho\left[\frac{1}{\rho|Q_{3r}|}\int_{Q_{3r}} \nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]. \tag{1.82}$$

Inequality eq. (1.82) controls the norm of the gradient of a $\mathbf{a}$-harmonic function in the small cube Q_r by a sum of terms involving the norm of the gradient in the larger cube Q_{3r}. This does not seem to be useful at first glance. However, the key point is that the multiplicative factor θ is smaller than one. Thus it also implies a better regularity in the interior, and eq. (1.82) can finally be integrated into the framework of the renormalization approach.

3. *The dimension-free Poincaré inequality*: the Poincaré inequality is a necessary tool for analysis, and we also establish it in $\mathscr{H}^1(U)$ and $\mathscr{H}^1_0(U)$. For the function space $\mathscr{H}^1(U)$, its proof relies on the Efron-Stein inequality, and we also improve it to the multiscale Poincaré inequality. For the function space $\mathscr{H}^1_0(U)$, our proof implicitly makes use of the Malliavin calculus on the Poisson space. See Section 6.3.1 for details.

1.5 Perspectives

Various numerical experiments in other models

There exists a large class of Dirichlet problems in singular random environments: the inhomogeneous Bernoulli percolation model, the Boolean model, the Voronoi percolation model, etc. Their qualitative homogenization result is proved in [226]. Although we do not establish all the quantitative results for these models, most of them share the same spirits with the percolation model, and the AHKM algorithm should also work. Thus, various numerical experiments can be tested and we hope to try them in practice for other applied problems.

KMT coupling for random walk on percolation cluster

A natural question is how the quantitative estimate for the semigroup can be interpreted in the trajectory of the stochastic processes. For the simple random walk on $\mathbb{Z}^d$, this is known as *the Skorokhod embedding problem* and the optimal result is proved by Komlós, Major and Tusnády (KMT) [155]: one can construct a common probability space for the simple random walk $(S_n)_{n\geqslant 0}$ and the Brownian motion $(B_t)_{t\geqslant 0}$ such that the error estimate is

$$\max_{k\leqslant n}|S_k - B_k| \simeq O(\log n). \tag{1.83}$$

The main ingredient for KMT theorem is the semigroup error estimate and a dyadic endpoint coupling strategy, and the former is not hard in the simple random walk setting by invoking directly the local central limit theorem. Viewing the estimate eq. (1.45), hopefully we can also prove a similar result as eq. (1.83) in percolation cluster setting. As a consequence, this coupling will answer the mixing time estimate on the percolation cluster posed in [46].

Optimal concentration inequality of the cluster density

One byproduct in Chapter 4 is a concentration inequality of the cluster density Proposition 4.A.1: there exists a finite positive exponent $s = \frac{2(d-1)}{3d^2+2d-1}$ and a finite positive constant $C(d,\mathfrak{p},s)$ such that in any cube Q_r of diameter r, we have

$$\mathbb{P}\left[\left|\frac{|\mathscr{C}_\infty \cap Q_r|}{|Q_r|} - \theta(\mathfrak{p})\right| > \varepsilon\right] \leqslant 2\exp\left(-\left(\frac{\varepsilon r^{\frac{d}{2}}}{C}\right)^s\right). \tag{1.84}$$

To the best of our knowledge, the result in previous work is a large deviation estimate [201, Theorem 1.2] that

$$\mathbb{P}\left[\left|\frac{|\mathscr{C}_\infty \cap Q_r|}{|Q_r|} - \theta(\mathfrak{p})\right| > \varepsilon\right] \leqslant C_1(d,\mathfrak{p},\varepsilon)\exp\left(-C_2(d,\mathfrak{p},\varepsilon)r^{d-1}\right), \tag{1.85}$$

where the constant C_1, C_2 depends on ε. Thus it is interesting to ask the optimal concentration inequality that recovers both eq. (1.84) and eq. (1.85), and a natural guess is to improve the s in eq. (1.84) to $s = \frac{2(d-1)}{d}$. It may require more analysis on the geometry for the infinite percolation clusters.

Heat kernel estimate for stable type random walk / long-range percolation

Recently a series of work [73, 75, 74] look at homogenization associated to the stable type operators. Sometimes we are asked if our strategy is robust enough to get the quantitative estimate for these models. We believe the answer is yes, otherwise it will be interesting to see what is missing.

On the other hand, if we also hope to generalize the result for other long-range correlation percolation models, then the renormalization step for the geometry should be used more carefully. In fact, for the Bernoulli bond percolation, the good cube exists with exponentially high probability. But if this event is only valid with polynomially high probability, every implement of the renormalization step will reduce the stochastic integrability. We mention the work [209] of Sapozhnikov for some explorations in this direction.

Optimal rate for hydrodynamic limit and fluctuation

We hope to obtain a quantitative version of hydrodynamic limit and fluctuation theorem in general particle systems without gradient condition, because it will provide us with a more flexible framework of convergence. See also the recent progress [144] in this direction for the model with gradient condition. The ultimate object is clearly the optimal rate, but we can start from some preliminary version.

A related question is the long-time variance estimate for particle systems without gradient condition mentioned in Chapter 5

$$\mathbb{V}\mathrm{ar}_\rho[u_t] = Ct^{-\frac{d}{2}} + o(t^{-\frac{d}{2}}). \tag{1.86}$$

We also recall that a similar result is proved in [142] for the zero range model, but the lack of gradient condition in our model is the main challenge. The constant C in eq. (1.86) is related to the bulk diffusion coefficient, and we hope to use its finite volume approximation in Chapter 6 to conclude eq. (1.86). Finally, we remark a minor difference between the definitions of bulk diffusion coefficient in eq. (1.69) and eq. (1.78), which pushes us to make the structure of the corrector clearer in particle systems.

Particles with more singular interactions

We can also consider a particle system with more singular interactions: every particle has a radius 1 and evolves as an independent Brownian motion before collisions; once two particles touch each other, there will be a reflecting boundary condition for the diffusion. This may recall the work of Bodineau, Gallagher and Saint-Raymond [57, 56, 58] on hard-spheres, but our model is a soft version and closer to the model in Chapters 5 and 6. From another viewpoint, the reflecting boundary condition is also similar to the percolation model, but now the environment is dynamic. Therefore, we hope to develop homogenization theory on this particle system to see the generalization on singular dynamic random environment.

Chapter 2

Uniform bound of the AHKM iterative algorithm

We study the iterative algorithm proposed by S. Armstrong, A. Hannukainen, T. Kuusi, J.-C. Mourrat in [22] to solve elliptic equations in divergence form with stochastic stationary coefficients. Such equations display rapidly oscillating coefficients and thus usually require very expensive numerical calculations, while this iterative method is comparatively easy to compute. In this chapter, we strengthen the estimate for the contraction factor achieved by one iteration of the algorithm. We obtain an estimate that holds uniformly over the initial function in the iteration, and which grows only logarithmically with the size of the domain.

This chapter corresponds to the article [133] and is published in Stochastics and Partial Differential Equations: Analysis and Computations.

2.1 Introduction

2.1.1 Main theorem

The problem of homogenization is a subject widely studied in mathematics and other disciplines for its applications and interesting properties. Let $(\mathbf{a}(x), x \in \mathbb{R}^d)$ be a random coefficient field, which takes values in the set of $\mathbb{R}^{d\times d}$ symmetric matrices, and which we assume to be $\mathbb{Z}^d$-stationary, with a unit range of dependence and uniformly elliptic that $\Lambda^{-1}|\xi|^2 \leqslant \xi \cdot \mathbf{a}(x)\xi \leqslant \Lambda|\xi|^2$ for any $x, \xi \in \mathbb{R}^d$. We give ourselves a bounded domain $U \subseteq \mathbb{R}^d$ with boundary $C^{1,1}$, a scale parameter $0 < \varepsilon < 1$, and for given $f \in H^{-1}(U)$ and $g \in H^1(U)$, we consider the elliptic Dirichlet problem

$$\begin{cases} -\nabla \cdot \left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right) \nabla u_\varepsilon\right) = f & \text{in } U, \\ u_\varepsilon = g & \text{on } \partial U. \end{cases} \tag{2.1}$$

For the scale $0 < \varepsilon \ll 1$, a naive numerical algorithm for this problem is generally very expensive, due to the rapid oscillations of the coefficients (comparatively to the size of the domain) and we have to refine the mesh of the numerical schema. Thus, different methods have been proposed to approximate the solution and one of them is to replace the conductance matrix $\mathbf{a}$ by a constant *effective conductance matrix* $\bar{\mathbf{a}}$ in eq. (2.1) and use its solution $\bar{u}$ as an approximation, which can be solved quickly thanks to the multi-grid algorithm. However, $\bar{u}$ is close to u_ε in the sense $L^2(U)$ or $H^{-1}(U)$, but not in some stronger topology, for example $H^1(U)$. Furthermore, the approximation only becomes accurate in the limit $\varepsilon \to 0$, but for a small finite scale ε, one can not expect a precision much smaller than ε with $\bar{u}$.

Recently, [22] proposed an iterative algorithm to solve the problem eq. (2.2) efficiently for a given ε-scale and with a better precision. We recap at first their algorithm here with the same formulation in large scale: Instead of considering eq. (2.1) with a small scale ε, we treat the Dirichlet problem with a dilation parameter $r \geqslant 1$, and set $U_r := rU$. Then given $f \in H^{-1}(U_r)$ and $g \in H^1(U_r)$, we consider the elliptic equation given by

$$\begin{cases} -\nabla \cdot \mathbf{a}\nabla u = f & \text{in } U_r, \\ u = g & \text{on } \partial U_r, \end{cases} \tag{2.2}$$

and [22] proposes to start by an initial guess of solution $v \in g + H_0^1(U_r)$, and solve $u_0, \bar{u}, \widetilde{u} \in H_0^1(U)$ satisfying (with null Dirichlet boundary condition)

$$\begin{cases} (\lambda^2 - \nabla \cdot \mathbf{a}\nabla)u_0 & = f + \nabla \cdot \mathbf{a}\nabla v & \text{in } U_r, \\ -\nabla \cdot \bar{\mathbf{a}}\nabla \bar{u} & = \lambda^2 u_0 & \text{in } U_r, \\ (\lambda^2 - \nabla \cdot \mathbf{a}\nabla)\widetilde{u} & = (\lambda^2 - \nabla \cdot \bar{\mathbf{a}}\nabla)\bar{u} & \text{in } U_r. \end{cases} \tag{2.3}$$

Then, the iteration $\hat{v} := v + u_0 + \widetilde{u}$ is a contraction. Since this rate of contraction is random, to estimate its size, [22] introduces the notation $\mathcal{O}_s$ for random variable X that for any $s, \theta > 0$,

$$X \leqslant \mathcal{O}_s(\theta) \Longleftrightarrow \mathbb{E}\left[\exp(\max(\theta^{-1}X, 0)^s)\right] \leqslant 2. \tag{2.4}$$

Informally speaking, the statement $X \leqslant \mathcal{O}_s(1)$ tells us that X has a tail lighter than $\exp(-x^s)$. We also introduce the shorthand notation for every $\lambda \in (0, 1]$,

$$\ell(\lambda) := \begin{cases} (\log(1 + \lambda^{-1}))^{\frac{1}{2}} & d = 2, \\ 1 & d > 2. \end{cases} \tag{2.5}$$

Then, [22, Thoerem 1.1] states that under a supplementary condition that the coefficient field $(\mathbf{a}(x), x \in \mathbb{R}^d)$ is α-Hölder, for any $s \in (0,2)$, we have a positive finite constant $C(s, U, \Lambda, \alpha, d)$ such that for the algorithm eq. (2.3) in a domain U_r with $r \geqslant 1$, we have

$$\|\nabla(\hat{v} - u)\|_{L^2(U_r)} \leqslant \mathcal{O}_s\left(C\ell^{\frac{1}{2}}(\lambda)\lambda^{\frac{1}{2}}\|\nabla(v-u)\|_{L^2(U_r)}\right). \tag{2.6}$$

However, in eq. (2.6) the contraction factor is proved for a given initialisation, but cannot be iterated to guarantee the convergence of the whole procedure. More precisely, after one iteration, the initial data becomes random and then we cannot make use of the noation $\mathcal{O}_s$ eq. (2.4) with a random θ directly. In order to go past this obstacle, the authors of [22] mention the possibility to get a uniform bound in [22, eq.(1.10)], as is necessary to guarantee the convergence of the iterated method. The goal of this article is to confirm this idea and provide a proof of this uniform bound.

Here is our main theorem. The assumption about the random field $(\mathbf{a}(x), x \in \mathbb{R}^d)$ is the one given at the beginning of the introduction, and its precise definition can be found in Section 2.2.1, where Λ is the constant of the uniform ellipticity condition.

Theorem 2.1.1 (Uniform H^1 contraction). *For every bounded domain $U \subseteq \mathbb{R}^d$ with $C^{1,1}$ boundary and every $s \in (0,2)$, there exists a positive finite constant $C(U, \Lambda, s, d)$ and, for every $r \geqslant 2$ and $\lambda \in \left(\frac{1}{r}, \frac{1}{2}\right)$, a random variable $\mathcal{Z}$ satisfying*

$$\mathcal{Z} \leqslant \mathcal{O}_s\left(C\ell(\lambda)^{\frac{1}{2}}\lambda^{\frac{1}{2}}(\log r)^{\frac{1}{s}}\right), \tag{2.7}$$

such that the following holds. Denote $U_r := rU$, let $f \in H^{-1}(U_r)$, $g \in H^1(U_r)$, $v \in g + H^1_0(U_r)$, let $u \in g + H^1_0(U_r)$ be the solution of eq. (2.2), and let $u_0, \bar{u}, \widetilde{u} \in H^1_0(U_r)$ solve eq. (2.3) with null Dirichlet boundary condition. Then for $\hat{v} := v + u_0 + \widetilde{u}$, we have the contraction estimate

$$\|\nabla(\hat{v} - u)\|_{L^2(U_r)} \leqslant \mathcal{Z}\|\nabla(v-u)\|_{L^2(U_r)}. \tag{2.8}$$

Compared with the result of [22], we have two fundamental improvements: The first one is the explicit random variable $\mathcal{Z}$, which is an upper bound for the contraction factor in the iteration, *does not depend on the function* v. Hence, the algorithm can be iterated, replacing v by $\hat{v}$, etc. In the event that $\mathcal{Z} \leqslant \frac{1}{2}$, say, we thus obtain exponential convergence of the iterative method to the solution u, a conclusion that cannot be inferred from the results of [22]. By the estimate eq. (2.7), in order to guarantee that $\mathcal{Z} \leqslant \frac{1}{2}$ with high probability, it suffices to take λ sufficiently small that $\ell(\lambda)^{\frac{1}{2}}\lambda^{\frac{1}{2}}(\log r)^{\frac{1}{s}}$ is below a certain positive constant. The second improvement is that we do not make any assumption on the regularity of the coefficient field $x \mapsto \mathbf{a}(x)$, while [22] assumed it to be Hölder continuous.

The price is that the bound eq. (2.7) has another factor $(\log r)^{\frac{1}{s}}$ than eq. (2.6) and this point is also conjectured in [22, eq.(1.10)]. This factor does not weaken our algorithm viewing the range of λ and more detailed study will be discussed in Section 2.1.3. In [22, Section 4], one can find some examples for the practical choice of λ and numerical experiences. The author has also repeated the algorithm in a domain 128×128 with $\lambda = 0.1$, for $\mathbf{a}$ of type chessboard with law Bernoulli$(\frac{1}{2})$ taking value in $\left\{\frac{1}{\sqrt{2}}, \sqrt{2}\right\}$, and it takes several seconds to obtain a precision 10^{-5} on a laptop. In [134], this algorithm is applied to the *supercritical percolation* setting and one can find the numerical results in Section 6. Hopefully, this algorithm also works on other stochastic homogenization models with stationary ergodic random coefficient field.

Remark. One can also state the algorithm eq. (2.3) and Theorem 2.1.1 in ε-scale for eq. (2.1): In fact, by a simple change of variable that $\varepsilon = \frac{1}{r}$, and $u_\varepsilon(\cdot) = u(\frac{\cdot}{\varepsilon})$ for u in eq. (2.2), then the same estimates eq. (2.8) and eq. (2.7) hold in the domain U for eq. (2.1) with $0 < \varepsilon \leqslant \frac{1}{2}$, when applying the following iteration with $\lambda \in (\varepsilon, 1)$

$$\begin{cases} \left(\left(\frac{\lambda}{\varepsilon}\right)^2 - \nabla \cdot \left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla\right)\right) u_0 & = f + \nabla \cdot \left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla\right) v \quad \text{in } U, \\ -\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} & = \left(\frac{\lambda}{\varepsilon}\right)^2 u_0 \quad \text{in } U, \\ \left(\left(\frac{\lambda}{\varepsilon}\right)^2 - \nabla \cdot \left(\mathbf{a}\left(\frac{\cdot}{\varepsilon}\right)\nabla\right)\right) \widetilde{u} & = \left(\left(\frac{\lambda}{\varepsilon}\right)^2 - \nabla \cdot \bar{\mathbf{a}} \nabla\right) \bar{u} \quad \text{in } U. \end{cases} \tag{2.9}$$

There exists a large amount of references about the homogenization theory and how we apply them in numerical solution. For example, see [47, 161, 219, 145, 6, 225, 188] for the classical homogenization theory and see [36, 101, 128, 197, 178, 159, 196, 129, 154, 103, 104, 100] for the multi-grid algorithm in homogenization problem. To analyze a numerical algorithm for stochastic homogenization problem, it requires quantitative description and it was open for long time. Thanks to the recent progress in a series of works of Armstrong, Kuusi, Mourrat and Smart [30, 23, 31, 24], and also the works of Gloria, Neukamm and Otto [123, 124, 121, 125, 122], we get a further understanding in this direction; see also the [25] a monograph and [185] as a brief introduction. In both this work and [22], the analysis depends on two-scale expansion theorem, which is introduced in [6] in periodic case and [8, Theorem 2.2, Theorem 2.3] gives its rate of convergence; the quantitative analysis for this problem with random coefficient is studied in [120] and [25, Chapter 6]. Finally, we remark that in all the context, we suppose that the effective conductance $\bar{\mathbf{a}}$ is known, because this part is now well understood and there exist many efficient methods to do it quickly, see for example [119, 102, 184, 107, 138].

The rest of the paper is organised as follows: At the end of the introduction, we focus on the numerical part to study why this algorithm is more efficient in Section 2.1.2, and explain heuristically how this algorithm converges to the solution in Section 2.1.3. Section 2.2 introduces some notations and then we turn to the proof of Theorem 2.1.1. The main improvements compared to [25] are the two technical lemmas in Section 2.3, then we put our technique in the proof of [22], which is a quantitative two-scale expansion theorem, and we reformulate it in Section 2.4. Finally, in Section 2.5, we combine all the results and obtain the main theorem.

2.1.2 Complexity analysis in numeric

In this part we give a numerical consequence of Theorem 2.1.1. We start by recalling why solving for $(\lambda^2 - \nabla \cdot \mathbf{a} \nabla)$ is computationally less difficult than solving for $-\nabla \cdot \mathbf{a} \nabla$. In our context, after discretization, the elliptic equations is transformed to a symmetric linear system

$$\mathbf{A}\mathbf{u} = \mathbf{f}, \tag{2.10}$$

where $\mathbf{A} \in \mathbb{R}^{N \times N}$ is positive definite, $\mathbf{u}, \mathbf{f} \in \mathbb{R}^N$ and N stands for the number of elements which is fixed during all this subsection. To capture all the information of coefficients, the minimal numerical resolution requires that $N = O(r^d)$. Then the problem becomes solving a large sparse linear system.

One basic method for this problem is the *conjugate gradient method* (CGM), whose rate of convergence is

$$\tau(\mathbf{A}) = \frac{\sqrt{\rho(\mathbf{A})} - 1}{\sqrt{\rho(\mathbf{A})} + 1},$$

where ρ is the *spectral condition number* defined as

$$\rho(\mathbf{A}) = \frac{\kappa_{max}(\mathbf{A})}{\kappa_{min}(\mathbf{A})},$$

and $\kappa_{max}, \kappa_{min}$ stands for the maximum and minimum eigenvalues ([207, Theorem 6.29, eq.(6.128)]). In practice, $\kappa_{max}(\mathbf{A}) \approx$ constant while $\kappa_{min}(\mathbf{A}) \approx r^{-2}$. Thus, when r grows bigger, the ratio of convergence $\tau(\mathbf{A}) \approx 1 - \frac{1}{r}$. It is still a geometric convergence but the rate is very small and to solve eq. (2.10) with a resolution ε_0, it requires $O\left(r|\log(\varepsilon_0)|\right)$ rounds of CGM.

Now we focus on the complexity to solve eq. (2.10) with eq. (2.3). Since in every iteration we solve two regularised equations and a homogenized equation, we investigate their complexity at first:

- For the homogenized equation, since the matrix is constant, which alows us to apply the multi-grid algorithm and for a resolution ε_1, the complexity is $O(|\log(\varepsilon_1)|)$ rounds CGM [65, Chapter 4].
- For the regularised equation $(\lambda^2 \mathsf{Id} + \mathbf{A})\mathbf{u}^\lambda = \mathbf{f}$, we use CGM and the spectral condition number for $\frac{1}{r} \ll \lambda \ll 1$ is

$$\rho(\lambda^2 \mathsf{Id} + \mathbf{A}) = \frac{\lambda^2 + \kappa_{max}(\mathbf{A})}{\lambda^2 + \kappa_{min}(\mathbf{A})} \approx \frac{C}{\lambda^2},$$

and this operation also changes the typical size of the rate of convergence

$$\tau(\lambda^2 \mathsf{Id} + \mathbf{A}) = \frac{\sqrt{\rho(\lambda^2 \mathsf{Id} + \mathbf{A})} - 1}{\sqrt{\rho(\lambda^2 \mathsf{Id} + \mathbf{A})} + 1} \approx 1 - \frac{\lambda}{C}.$$

Then for a resolution ε_1, it requires $O(\lambda^{-1}|\log(\varepsilon_1)|)$.

When we implement the algorithm eq. (2.3), generally speaking, the eq. (2.7) tells us with a large probability, after every iteration the precision will be multiplied $\lambda^{\frac{1}{2}}$. Thus, totally it demands $O\left(|\log(\lambda)|^{-1}|\log(\varepsilon_0)|\right)$ iterations. Moreover, in the k-th iteration, it suffices to obtain a resolution $\varepsilon_1 = \lambda^{\frac{k}{2}}$ for the regularised equation and the homogenized equation, so the complexity is

$$\sum_{k=1}^{O\left(|\log(\lambda)|^{-1}|\log(\varepsilon_0)|\right)} \lambda^{-1}|\log(\lambda^{\frac{k}{2}})| \approx |\log(\lambda)|^{-1}\lambda^{-1}\log^2(\varepsilon_0).$$

When we take a typical choice of $\lambda = (\log r)^{-1}$, the complexity of the iterative algorithm is $O(\log r|\log \varepsilon_0|^2)$ rounds of CGM. This quanity is much smaller than the complexity of solving eq. (2.10) directly for a big r, provided that the precision is reasonable, for example $\varepsilon_0 = r^{-n}$ for n fixed.

2.1.3 Heuristic analysis of algorithm

In this part, we explain heuristicly why the iterative algorthm converges to the true solution and how we figure out this algorithm. We keep in mind the main idea: To solve eq. (2.2), we divide it into sub-problems of regularized equations and homogenized equations.

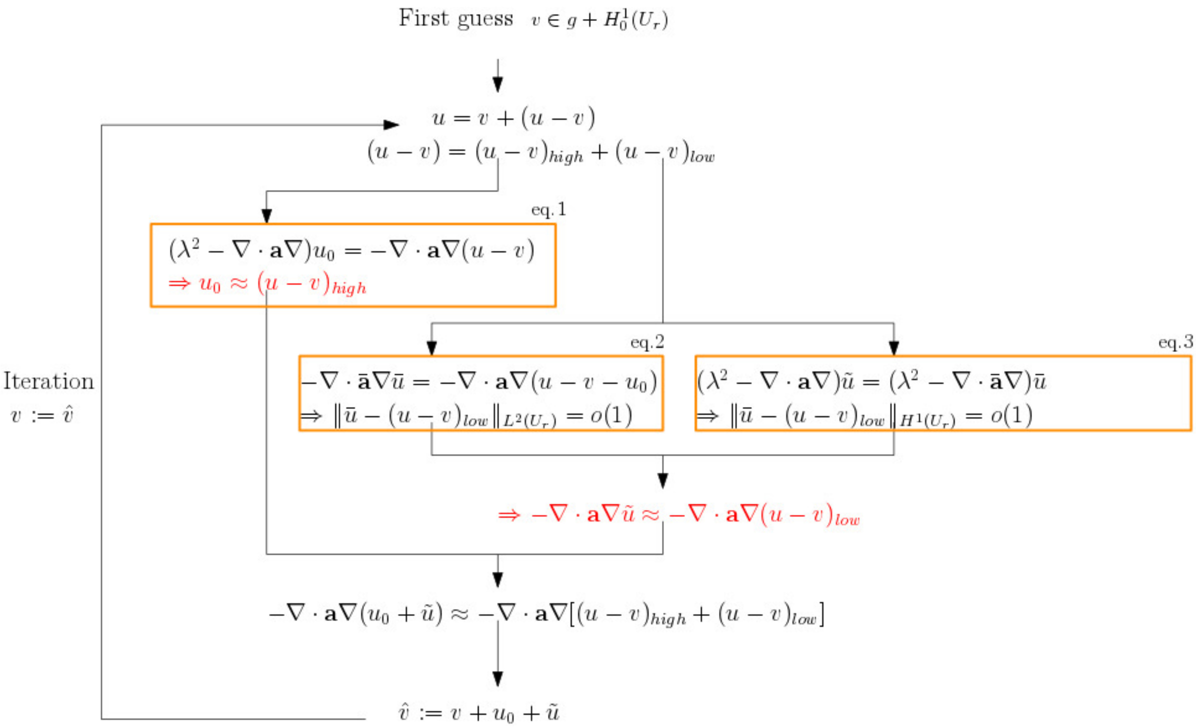

Figure 2.1: A flowchart shows the mechanic of the algorithm.

We start with an initial guess of the solution v, then we write $u = v + (u - v)$ and we want to recover the part $(u - v)$. Since the divergence form is linear, we have

$$-\nabla \cdot \mathbf{a}\nabla(u - v) = -\nabla \cdot \mathbf{a}\nabla u + \nabla \cdot \mathbf{a}\nabla v = f + \nabla \cdot \mathbf{a}\nabla v.$$

In the first step of our algorithm, we solve the problem with regularization

$$(\lambda^2 - \nabla \cdot \mathbf{a}\nabla)u_0 = f + \nabla \cdot \mathbf{a}\nabla v,$$

and u_0 gives the high frequency part of $(u - v)$ associated to the operator $-\nabla \cdot \mathbf{a}\nabla$ on U_r. To see it, we apply the theorem of spectral decomposition [205, Chapter 5]

$$(u - v) = \sum_i^\infty \psi_i,$$

where ψ_i is the projection on the subspace of the eigenvalue κ_i associated to $-\nabla \cdot \mathbf{a}\nabla$ on U_r, with $0 < \kappa_1 < \kappa_2 \cdots$. Then u_0 has an expression

$$u_0 = \sum_{i=1}^\infty \frac{\kappa_i}{\lambda^2 + \kappa_i}\psi_i.$$

We see the projections associated to large eigenvalues have a small perturbation after the regularization. Thus, we consider the solution u_0 of the high frequency projection of $(u - v)$. Informally, we write

$$\begin{aligned} (u - v) &\approx (u - v)_{high} + (u - v)_{low}, \\ u_0 &= (u - v)_{high}. \end{aligned}$$

Therefore, after the first step, we do not get all the information of $(u-v)$ but $(u-v)_{high}$ and the second and the third equation serve to recover $(u-v)_{low}$ of $(u-v)$. Using the superposition, a direct idea is to solve

$$-\nabla\cdot\mathbf{a}\nabla\widetilde{u}\approx-\nabla\cdot\mathbf{a}\nabla(u-v-u_0)=\lambda^2u_0. \tag{2.11}$$

But for the reason of less numerical cost, we choose to solve at first a homogenized problem

$$-\nabla\cdot\bar{\mathbf{a}}\nabla\bar{u}=-\nabla\cdot\mathbf{a}\nabla(u-v-u_0)=\lambda^2u_0, \tag{2.12}$$

and if we believe that $-\nabla\cdot\mathbf{a}\nabla\widetilde{u}\approx-\nabla\cdot\bar{\mathbf{a}}\nabla\bar{u}$, we can also solve the one by adding regularization

$$(\lambda^2-\nabla\cdot\mathbf{a}\nabla)\widetilde{u}=(\lambda^2-\nabla\cdot\bar{\mathbf{a}}\nabla)\bar{u}. \tag{2.13}$$

Then, we hope that this $\widetilde{u}$ gives us $(u-v)_{low}$. Perhaps this "≈" is not very precise, so we do $\hat{v}:=v+u_0+\widetilde{u}$ and put $\hat{v}$ in the role of v for several iterations in order to get a further approach to the solution of eq. (2.2).

2.2 Notation

In this section, we state our assumptions about the coefficient field precisely and introduce some notations.

2.2.1 Assumptions on the coefficient field

We denote by $((\mathbf{a}(x))_{x\in\mathbb{R}^d},\mathcal{F},\mathbb{P})$ the probability space, and by $\mathcal{F}_V$ the σ-algebra generated by

$$\mathbf{a}\mapsto\int_{\mathbb{R}^d}\chi\mathbf{a}_{i,j},\ \text{where } i,j\in\{1,2,3\cdots d\},\chi\in C_c^\infty(V).$$

$\mathcal{F}$ is short for $\mathcal{F}_{\mathbb{R}^d}$. T_y denotes the operator of translation i.e. for any function f, $T_y(f)(x):=f(x+y)$ and for any set A, $T_y(A):=\{x+y|x\in A\}$.

The precise assumptions for the coefficient field are as follows.

1. $\mathbb{Z}^d$-*stationarity*: For each $A\in\mathcal{F}$ and each $y\in\mathbb{Z}^d$, we have $\mathbb{P}[T_y(A)]=\mathbb{P}[A]$.
2. *Unit range correlation*:
$$\forall W,V\in\mathcal{B}(\mathbb{R}^d),\quad d_H(W,V)>1\implies\mathcal{F}_W,\mathcal{F}_V\text{ are independent.}$$
Here d_H is the Hausdorff distance in $\mathbb{R}^d$.
3. *Uniform ellipticity*: There exists $0<\Lambda<\infty$ such that with probability one and for every $x,\xi\in\mathbb{R}^d$, we have $\Lambda^{-1}|\xi|^2\leqslant\xi\cdot\mathbf{a}(x)\xi\leqslant\Lambda|\xi|^2$.

2.2.2 Notation $\mathcal{O}_s$

We recall the definition of $\mathcal{O}_s$

$$X\leqslant\mathcal{O}_s(\theta)\iff\mathbb{E}\left[\exp((\theta^{-1}X)_+^s)\right]\leqslant2, \tag{2.14}$$

where $(\theta^{-1}X)_+$ means $\max\{\theta^{-1}X,0\}$. This notation gives the tail estimate of random variables: One can use Markov's inequality to obtain that

$$X\leqslant\mathcal{O}_s(\theta)\implies\forall x>0,\mathbb{P}[X\geqslant\theta x]\leqslant2\exp(-x^s).$$

Moreover, we can also obtain the estimate of the sum of a series of random variables without its joint distribution: For a measure space $(E, \mathcal{S}, m)$ and $\{X(z)\}_{z\in E}$ a family of random variables, we have

$$\forall z \in \mathcal{O}_s(E), X(z) \leqslant \mathcal{O}_s(\theta(z)) \Longrightarrow \int_E X(z)m(dz) \leqslant \mathcal{O}_s\left(C_s \int_E \theta(z)m(dz)\right), \tag{2.15}$$

where $0 < C_s < \infty$ is a constant and $C_s = 1$ for $s \geqslant 1$. See Appendix of [25, Appendix A] for proofs and other operations on $\mathcal{O}_s$.

2.2.3 Convolution

For $f \in L^p(\mathbb{R}^d), g \in L^q(\mathbb{R}^d)$ where $\frac{1}{p} + \frac{1}{q} = 1$, we denote by $f \star g$ the convolution of the function f, g

$$f \star g(x) = \int_{\mathbb{R}^d} f(y)g(x-y)\, dy.$$

In this article, two mollifiers used are the heat kernel $\Phi_r(x)$, defined for $r > 0$ and $x \in \mathbb{R}^d$ by

$$\Phi_r(x) := \frac{1}{(4\pi r^2)^{d/2}} \exp\left(-\frac{x^2}{4r^2}\right),$$

and the bump function $\zeta \in C_c^\infty(\mathbb{R}^d)$

$$\zeta(x) := c_d \exp\left(-\left(\frac{1}{4} - |x|^2\right)^{-1}\right) \mathbf{1}_{\{\{|x|<\frac{1}{2}\}\}},$$

where c_d is the constant of normalization such that $\int_{\mathbb{R}^d} \zeta(x)dx = 1$. Finally, we use the notation

$$\zeta_\varepsilon(x) = \frac{1}{\varepsilon^d}\zeta\left(\frac{x}{\varepsilon}\right),$$

as a mollifier in scale $\varepsilon > 0$ and we have $\operatorname{supp}(\zeta_\varepsilon) \subseteq \overline{B_{\varepsilon/2}}$.

2.2.4 Function spaces

We use $\{e_1, e_2, \cdots e_d\}$ as the canonical basis of $\mathbb{R}^d$. For every Borel set $U \subseteq \mathbb{R}^d$, let $|U|$ be its Lebesgue measure. For each $p \in [1, +\infty]$, we denote by $L^p(U)$ the classical Lebesgue space and $L^p_{loc}(\mathbb{R}^d)$ the function space for the functions with finite $L^p(V)$ norm for any compact set V. The weighted norm $\underline{L}^p(U)$ is defined for a bounded Borel set U as

$$\|f\|_{\underline{L}^p(U)} = \left(\frac{1}{|U|}\int_U |f(x)|^p\, dx\right)^{\frac{1}{p}} = |U|^{-\frac{1}{p}} \|f\|_{L^p(U)}.$$

For each $k \in \mathbb{N}$, we denote by $H^k(U)$ the classical Sobolev space on U equipped with the norm

$$\|f\|_{H^k(U)} := \sum_{0\leqslant|\beta|\leqslant k} \|\partial^\beta f\|_{L^2(U)},$$

where $\beta \in \mathbb{N}^d$ represents a multi-index weak derivative,

$$|\beta| := \sum_{i=1}^d \beta_i \quad \text{and} \quad \partial^\beta f = \partial_{x_1}^{\beta_1} \cdots \partial_{x_d}^{\beta_d} f.$$

We also use $|\nabla^k f|$ to indicate $\sum_{|\beta|=k} |\partial^\beta f|$. When $|U| < \infty$, we define the weighted norm that

$$\|f\|_{\underline{H}^k(U)} := \sum_{0\leqslant|\beta|\leqslant k} |U|^{\frac{|\beta|-k}{d}} \|\partial^\beta f\|_{\underline{L}^2(U)}.$$

$H_0^k(U)$ denotes the closure of $C_c^\infty(U)$ in $H^k(U)$ and $H^{-1}(U)$ for the dual of $H^1(U)$. The weighted norm $\underline{H}^{-1}(U)$ is

$$\|f\|_{\underline{H}^{-1}(U)} := \sup\left\{\frac{1}{|U|}\int_U f(x)g(x)\,dx, g \in H_0^1(U), \|g\|_{\underline{H}^1(U)} \leqslant 1\right\}.$$

Here, we abuse the use of the integration $\int_U f(x)g(x)\,dx$, since the function space $H^{-1}(U)$ also contains linear functional, which is not necessarily a function.

Finally, we remark that one advantage of the definition of $\underline{H}^k$ is that it is consistent with the scale constant of Poincaré's inequality [114, eq.(7.44)] and Sobolev extension theorem [114, Theorem 7.25]. That is, under the condition that the Borel set U has $C^{1,1}$ boundary, for any function $f \in H_0^1(U_r)$

$$\|f\|_{\underline{H}^1(U_r)} \leqslant C(U,d)\|\nabla f\|_{\underline{L}^2(U_r)}, \tag{2.16}$$

and for any $f \in H^2(U_r)$, there exists an extension $\mathsf{Ext}(f) \in H_0^2(\mathbb{R}^d)$ such that $\mathsf{Ext}(f) \equiv f$ in U_r

$$\sum_{0\leqslant|\beta|\leqslant 2} |U_r|^{-\frac{d}{2}-\frac{|\beta|-2}{d}} \|\partial^\beta \mathsf{Ext}(f)\|_{L^2(\mathbb{R}^d)} \leqslant C(U,d)\|f\|_{\underline{H}^2(U_r)}. \tag{2.17}$$

The proof depends on the scaling argument: For eq. (2.16), we prove at first the result in domain U and then apply to $x \mapsto f(rx)$. For eq. (2.17), we apply [114, Theorem 7.25] to the domain U and obtain an extension Ext_U satisfying eq. (2.17) for $r = 1$. Then, for the extension Ext_{U_r} on the domain U_r, we define $\mathsf{Ext}_{U_r}(f)(x) := \mathsf{Ext}_U(f(r\cdot))(x/r)$ and this satisfies eq. (2.17) with a constant depending only on U, d. In the following paragraphs we neglect the index of domain and still note the extension Ext.

2.2.5 Cubes

We use the notation $\square$ to refer the open unit cube $\square := \left(-\frac{1}{2},\frac{1}{2}\right)^d$. For any $z \in \mathbb{R}^d$, the translation of $\square$ in the direction z writes $z + \square := z + \left(-\frac{1}{2},\frac{1}{2}\right)^d$. The sum of a cube and a Borel set U is defined as

$$\square + U := \left\{z \in \mathbb{R}^d | z = x + y, x \in \square, y \in U\right\}.$$

We also denote define scaling of the cube by size $\varepsilon > 0$ that $\square_\varepsilon := \left(-\frac{\varepsilon}{2},\frac{\varepsilon}{2}\right)^d$.

2.3 Two technical lemmas

We prove two useful lemmas that improves the estimate of the iterative algorithm in this section. A formula similar to Lemma 2.3.1 can be found in [25, Lemme 6.7]. Here we introduce a variant version and it works well together with Lemma 2.3.2.

2.3.1 An inequality of localization

Lemma 2.3.1 (Mixed norm)**.** *There exists a constant* $0 < C(d) < \infty$ *such that for every* $f \in L^2_{loc}(\mathbb{R}^d)$*,* $g \in L^2(\mathbb{R}^d)$ *and every* $\varepsilon > 0, r \geqslant 2$*, we have the following inequality*

$$\|f(g \star \zeta_\varepsilon)\|_{L^2(U_r)} \leqslant C(d) \left(\max_{z \in \varepsilon\mathbb{Z}^d \cap (U_r + \square_\varepsilon)} \|f\|_{\underline{L}^2(z+\square_\varepsilon)} \right) \|g\|_{L^2(U_r + \square_{3\varepsilon})}. \tag{2.18}$$

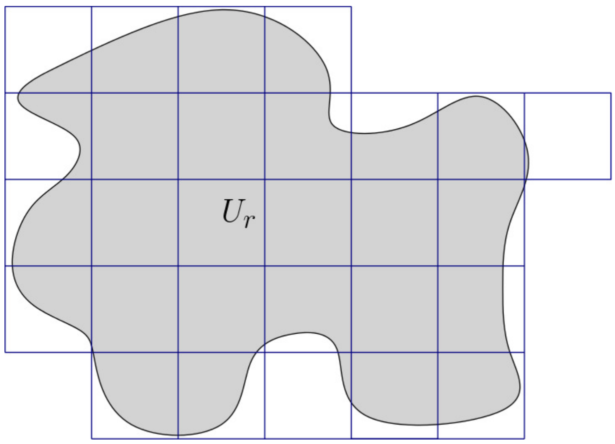

Figure 2.2: We calculate the L^2 norm by the sum of all norm in small cubes of size ε, so we counts all cubes in the domain $(U_r + \square_\varepsilon)$.

Proof. We decompose the L^2 norm as the sum of that in small cubes $\square_\varepsilon$

$$\begin{aligned}
\|f(g \star \zeta_\varepsilon)\|^2_{L^2(U_r)} &\leqslant \sum_{z \in \varepsilon\mathbb{Z}^d \cap (U_r + \square_\varepsilon)} \|f(g \star \zeta_\varepsilon)\|^2_{L^2(z+\square_\varepsilon)} \\
\text{(Hölder's inequality)} &\leqslant \sum_{z \in \varepsilon\mathbb{Z}^d \cap (U_r + \square_\varepsilon)} \left(\|f\|^2_{L^2(z+\square_\varepsilon)} \|g \star \zeta_\varepsilon\|^2_{L^\infty(z+\square_\varepsilon)} \right) \\
&\leqslant \left(\max_{z \in \varepsilon\mathbb{Z}^d \cap (U_r + \square_\varepsilon)} \|f\|^2_{L^2(z+\square_\varepsilon)} \right) \left(\sum_{z \in \varepsilon\mathbb{Z}^d \cap (U_r + \square_\varepsilon)} \|g \star \zeta_\varepsilon\|^2_{L^\infty(z+\square_\varepsilon)} \right).
\end{aligned}$$

Noticing that for any $x \in z + \square_\varepsilon, y \in B_{\varepsilon/2}$ then $(x - y) \in z + \square_{2\varepsilon}$ and we have

$$\begin{aligned}
|g \star \zeta_\varepsilon(x)| &= \left| \int_{\square_\varepsilon} g(x-y) \frac{1}{\varepsilon^d} \zeta(\frac{y}{\varepsilon}) \, dy \right| \\
&\leqslant \frac{C(d)}{\varepsilon^d} \int_{z+\square_{2\varepsilon}} |g(y)| \, dy \\
&\leqslant \frac{C(d)}{\varepsilon^d} \left(\int_{z+\square_{2\varepsilon}} |g(y)|^2 \, dy \right)^{\frac{1}{2}} |\square_{2\varepsilon}|^{\frac{1}{2}} \\
&\leqslant \frac{C(d)}{\varepsilon^{\frac{d}{2}}} \|g\|_{L^2(z+\square_{2\varepsilon})}.
\end{aligned}$$

So we get

$$\|g \star \zeta_\varepsilon\|_{L^\infty(z+\square_\varepsilon)} \leqslant \frac{C(d)}{\varepsilon^{\frac{d}{2}}} \|g\|_{L^2(z+\square_{2\varepsilon})},$$

and we add this analysis in the former inequality and obtain that

$$\begin{aligned}\|f(g\star\zeta_\varepsilon)\|^2_{L^2(U_r)} &\leqslant C(d)\left(\frac{1}{\varepsilon^d}\max_{z\in\varepsilon\mathbb{Z}^d\cap(U_r+\square_\varepsilon)}\|f\|^2_{L^2(z+\square_\varepsilon)}\right)\left(\sum_{z\in\varepsilon\mathbb{Z}^d\cap(U_r+\square_\varepsilon)}\|g\|^2_{L^2(z+\square_{2\varepsilon})}\right)\\ &\leqslant C(d)\left(\max_{z\in\varepsilon\mathbb{Z}^d\cap(U_r+\square_\varepsilon)}\|f\|^2_{\underline{L}^2(z+\square_\varepsilon)}\right)\|g\|^2_{L^2(U_r+\square_{3\varepsilon})}.\end{aligned}$$

This is the desired inequality. □

2.3.2 Maximum of finite number of random variables of type $\mathcal{O}_s(1)$

Since $\left(\max_{z\in\varepsilon\mathbb{Z}^d\cap(U_r+\square_\varepsilon)}\|f\|_{\underline{L}^2(z+\square_\varepsilon)}\right)$ often appears in the context as the maximum of a family of random variables, we prove the following lemma to analyze the maximum of a finite number of random variables of type $\mathcal{O}_s(1)$. Note that we do not make any assumptions on the joint law of the random variables.

Lemma 2.3.2. *For all $N \geqslant 1$ and a family of random variables $\{X_i\}_{1\leqslant i\leqslant N}$ satisfying that $X_i \leqslant \mathcal{O}_s(1)$, we have*

$$\left(\max_{1\leqslant i\leqslant N} X_i\right) \leqslant \mathcal{O}_s\left(\left(\frac{\log(2N)}{\log(4/3)}\right)^{\frac{1}{s}}\right). \tag{2.19}$$

Proof. For the case $N = 1$, we could check that eq. (2.19) is established since

$$M = X_1 \leqslant \mathcal{O}_s(1) \Longrightarrow M \leqslant \mathcal{O}_s\left(\frac{\log(4)}{\log(4/3)}\right),$$

so we focus on the case $N \geqslant 2$. By Markov's inequality, $X_i \leqslant \mathcal{O}_s(1)$ gives us $\mathbb{P}[X_i > x] \leqslant 2e^{-x^s}$. Then we use the union bound to get

$$\mathbb{P}\left[\max_{1\leqslant i\leqslant N} X_i > x\right] = \mathbb{P}\left[\bigcup_{i=1}^N \{X_i > x\}\right] \leqslant \left(1\wedge\sum_{i=1}^N \mathbb{P}[X_i > x]\right) \leqslant 1\wedge 2Ne^{-x^s}. \tag{2.20}$$

We denote by x_0 the critical point such that $e^{x_0^s} = 2N$ and $a > 0$ such that $a^s > 3$ and its value will be chosen carefully later. We also set $M = \max_{1\leqslant i\leqslant N} X_i$ and use Fubini's theorem: For an increasing positive function $g \in C^1(\mathbb{R})$ such that $g(0) = 0$, we have

$$\mathbb{E}[g(M)] = \int_0^\infty g(t)\mathbb{P}[M > t]\,dt. \tag{2.21}$$

We apply eq. (2.21) the function $x \mapsto \exp\left(\left(\frac{x}{a}\right)_+^s\right) - 1$,

$$\begin{aligned}\mathbb{E}\left[\exp\left(\left(\frac{M}{a}\right)_+^s\right)\right] &= \int_0^\infty \frac{s}{a}\left(\frac{x}{a}\right)^{s-1} e^{\left(\frac{x}{a}\right)^s}\mathbb{P}[M > x]\,dx + 1\\ &= \int_0^{x_0} \frac{s}{a}\left(\frac{x}{a}\right)^{s-1} e^{\left(\frac{x}{a}\right)^s}\mathbb{P}[M > x]\,dx\\ &\quad + \int_{x_0}^\infty \frac{s}{a}\left(\frac{x}{a}\right)^{s-1} e^{\left(\frac{x}{a}\right)^s}\mathbb{P}[M > x]\,dx + 1.\end{aligned} \tag{2.22}$$

For the integration on the interval $(0, x_0]$, we use the estimate eq. (2.20) that $\mathbb{P}[M > x] \leqslant 1$ and $e^{x_0^s} = 2N$ to bound it

$$\begin{aligned} \int_0^{x_0} \frac{s}{a}\left(\frac{x}{a}\right)^{s-1} e^{\left(\frac{x}{a}\right)^s} \mathbb{P}[M > x]\, dx &\leqslant \int_0^{x_0} \frac{s}{a}\left(\frac{x}{a}\right)^{s-1} e^{\left(\frac{x}{a}\right)^s}\, dx \\ &= \int_0^{x_0} e^{\left(\frac{x}{a}\right)^s}\, d\left(\frac{x}{a}\right)^s \\ &= (2N)^{\frac{1}{a^s}} - 1. \end{aligned} \tag{2.23}$$

Similarly, for the integration on the interval (x_0, ∞), we use eq. (2.20) to control the probability that $\mathbb{P}[M > x] \leqslant 2Ne^{-x^s}$ and give an estimate

$$\begin{aligned} \int_{x_0}^{\infty} \frac{s}{a}\left(\frac{x}{a}\right)^{s-1} e^{\left(\frac{x}{a}\right)^s} \mathbb{P}[M > x]\, dx &\leqslant 2N \int_{x_0}^{\infty} \frac{s}{a}\left(\frac{x}{a}\right)^{s-1} e^{\left(\frac{x}{a}\right)^s - x^s}\, dx \\ &= 2N \int_{x_0}^{\infty} e^{\left(\frac{x}{a}\right)^s - x^s}\, d\left(\frac{x}{a}\right)^s \\ &= \underbrace{\frac{1}{a^s - 1}}_{\text{Using } a^s \geqslant 3} (2N)^{\frac{1}{a^s}} \\ &\leqslant \frac{1}{2}(2N)^{\frac{1}{a^s}}. \end{aligned} \tag{2.24}$$

Now we fix $a = \left(\frac{\log(2N)}{\log(4/3)}\right)^{\frac{1}{s}}$. For the case $N \geqslant 2$, we can check that

$$a^s = \left(\frac{\log(2N)}{\log(4/3)}\right) \geqslant \left(\frac{\log(4)}{\log(4/3)}\right) > 3,$$

so $a^s \geqslant 3$ is satisfied. Finally, we put back the estimate eq. (2.23) and eq. (2.24) back to eq. (2.22) and verify the definition of $\mathcal{O}_s$

$$\mathbb{E}\left[\exp\left(\left(\frac{M}{a}\right)_+^s\right)\right] \leqslant \frac{3}{2}(2N)^{\frac{1}{a^s}} = \frac{3}{2}e^{\log(4/3)} = 2.$$

This finishes the proof. □

2.4 Two-scale expansion estimate

This section reformulates [22, Theorem 3.1] a quantitative two-scale expansion theorem with the improvements from the lemmas in Section 2.3.

2.4.1 Main structure

The two-scale expansion allows us to approximate the solution of eq. (2.2) by the solution of the homogenized equation and the *first order corrector* $\{\phi_{e_k}\}_{1 \leqslant k \leqslant d}$. We recall at first the definition of the first order corrector.

Definition 2.4.1 (First order corrector, Lemma 3.16 and Theorem 4.1 of [25])**.** For each $p \in \mathbb{R}^d$, the corrector ϕ_p is the sublinear function satisfying that $p \cdot x + \phi_p$ is $\mathbf{a}$-harmonic in whole space $\mathbb{R}^d$ i.e.

$$-\nabla \cdot \mathbf{a}(p + \nabla \phi_p) = 0, \text{ in } \mathbb{R}^d. \tag{2.25}$$

$\nabla \phi_p$ is $\mathbb{Z}^d$-stationary and ϕ_p is well-defined up to a constant. For $d \geqslant 3$, we can choose a constant such that $\mathbb{E}\left[\int_{\square} \phi_p\right] = 0$, and in this case ϕ_p is also $\mathbb{Z}^d$-stationary.

We remark that proof of the property ϕ_p is $\mathbb{Z}^d$-stationary for $d \geqslant 3$ requires a detailed quantitative study of the following modified corrector

$$\phi_p^{(\lambda)} := \phi_p - \phi_p \star \Phi_{\lambda^{-1}}, \tag{2.26}$$

which is $\mathbb{Z}^d$-stationary and is always well-defined. Then the quantitative study of $\phi_p^{(\lambda)}$ allows us to extract a limit in the subsequence when $\lambda \to 0$ and this gives us a candidate for the choice of constant. For its complete proof, see [25, Chapter 4.6, Page 175]. In the following paragraphs, we specify the constant by $\mathbb{E}\left[\int_{\square} \phi_p\right] = 0$. Therefore, we can use the property that ϕ_p is $\mathbb{Z}^d$-stationary for $d \geqslant 3$.

Once we define the first-order corrctor, as well-known from [6], for v the heterogeneous and $\bar{v}$ the homogenized solution, the two-scale expansion $\bar{v} + \sum_{k=1}^d (\partial_{x_k}\bar{v})\phi_{e_k}$ approximates v in the sense H^1. We follow the similar idea in [25, Chapter 6] and apply a modified two-scale expansion to $\bar{v} \in H_0^1(U_r) \cap H^2(U_r)$ with $\{\phi_{e_k}^{(\lambda)}\}_{1 \leqslant k \leqslant d}$ defind in eq. (2.26)

$$w := \bar{v} + \sum_{k=1}^{d} \partial_{x_k}(\mathsf{Ext}(\bar{v}) \star \zeta)\phi_{e_k}^{(\lambda)}, \tag{2.27}$$

where $\mathsf{Ext}(\bar{v})$ is the Sobolev extension such that eq. (2.17) is established. In the following we abuse a little the notation to identify

$$\bar{v} \star \zeta := \mathsf{Ext}(\bar{v}) \star \zeta, \tag{2.28}$$

in order to shorter the equation. We want to prove a H^1 convergence theorem for the operator $(\mu^2 - \nabla \cdot \mathbf{a}\nabla)$. We also recall the definition of $\ell(\lambda)$

$$\ell(\lambda) = \begin{cases} (\log(1+\lambda^{-1}))^{\frac{1}{2}} & d = 2, \\ 1 & d > 2. \end{cases}$$

Theorem 2.4.1 (Two-scale estimate)**.** *For every $r \geqslant 2, \lambda \in (\frac{1}{r}, \frac{1}{2})$, there exists three $\mathcal{F}$-measurable random variables $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1$ satisfying for each $s \in (0,2)$, there exists a constant $C'(U,s,d)$ such that the following holds: $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1$ have an estimate*

$$\mathcal{X}_1 \leqslant \mathcal{O}_s\left(C'(U,s,d)\ell(\lambda)(\log r)^{\frac{1}{s}}\right), \qquad \mathcal{X}_2 \leqslant \mathcal{O}_s\left(C'(U,s,d)\lambda^{\frac{d}{2}}(\log r)^{\frac{1}{s}}\right), \tag{2.29}$$

$$\mathcal{Y}_1 \leqslant \mathcal{O}_s\left(C'(U,s,d)\ell(\lambda)(\log r)^{\frac{1}{s}}\right), \tag{2.30}$$

and for every $\bar{v} \in H_0^1(U_r) \cap H^2(U_r), \mu \in [0,\lambda]$ and $v \in H_0^1(U_r)$ satisfying

$$(\mu^2 - \nabla \cdot a\nabla)v = (\mu^2 - \nabla \cdot \bar{\mathbf{a}}\nabla)\bar{v} \qquad \textit{in } U_r, \tag{2.31}$$

we have an H^1 estimate for the two-scale expansion w associated to $\bar{v}$ defined in eq. (2.27)

$$\begin{aligned} \|v - w\|_{\underline{H}^1(U_r)} \leqslant C(U,\Lambda,d)\Bigg[\|\bar{v}\|_{\underline{H}^2(U_r)} &+ \left(\|\bar{v}\|_{\underline{H}^2(U_r)} + \mu\|\bar{v}\|_{\underline{H}^1(U_r)}\right)\mathcal{X}_1 \\ &+ \left(\ell(\lambda)^{\frac{1}{2}}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^2(U_r)}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^1(U_r)} + \|\bar{v}\|_{\underline{H}^1(U_r)}\right)\mathcal{X}_2 \\ &+ \left(\ell(\lambda)^{\frac{1}{2}}\left(\mu + \frac{1}{r} + \frac{1}{\ell(\lambda)}\right)\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^2(U_r)}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^1(U_r)} + \|\bar{v}\|_{\underline{H}^2(U_r)}\right)\mathcal{Y}_1\Bigg]. \end{aligned} \tag{2.32}$$

Remark. The explicit expression of $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1$ can be checked in fig. 2.3. They are the maximum of local spatial average of gradient and the flux of the first order corrector.

Proof. We give at first the proof of the deterministic part. We will see that the errors can finally be reduced to the estimates of two norms: the interior error term and a boundary layer term. The latter boundary term comes from the fact that v and w do not have the same boundary condition. So we introduce b the solution of the equation

$$\begin{cases} (\mu^2 - \nabla\cdot\mathbf{a}\nabla)b = 0 & \text{in } U_r, \\ b = \sum_{k=1}^d \partial_{x_k}(\bar{v}\star\zeta)\phi_{e_k}^{(\lambda)} & \text{on } \partial U_r. \end{cases} \tag{2.33}$$

Then $(w-b)$ shares the same boundary condition as v. So, we have

$$\|v-w\|_{\underline{H}^1(U_r)} \leqslant \|v+b-w\|_{\underline{H}^1(U_r)} + \|b\|_{\underline{H}^1(U_r)}, \tag{2.34}$$

and we do estimates of the two parts respectively.

- **Estimate for $(v+b-w)$.** We denote by $z := v+b-w \in H_0^1(U_r)$ and test it in eq. (2.31) and eq. (2.33)

$$\begin{aligned} \mu^2\int_{U_r} zv + \int_{U_r}\nabla z\cdot\mathbf{a}\nabla v &= \mu^2\int_{U_r} z\bar{v} + \int_{U_r}\nabla z\cdot\bar{\mathbf{a}}\nabla\bar{v}, \\ \mu^2\int_{U_r} zb + \int_{U_r}\nabla z\cdot\mathbf{a}\nabla b &= 0. \end{aligned}$$

We do the sum to obtain that

$$\mu^2\int_{U_r} z(v+b) + \int_{U_r}\nabla z\cdot\mathbf{a}\nabla(v+b) = \mu^2\int_{U_r} z\bar{v} + \int_{U_r}\nabla z\cdot\bar{\mathbf{a}}\nabla\bar{v}.$$

Using the fact $v+b = z+w$, we obtain

$$\mu^2\int_{U_r}|z|^2 + \int_{U_r}\nabla z\cdot\mathbf{a}\nabla z = \mu^2\int_{U_r} z(\bar{v}-w) + \int_{U_r}\nabla z\cdot(\bar{\mathbf{a}}\nabla\bar{v} - \mathbf{a}\nabla w),$$

and we apply the uniform ellipticity condition to obtain

$$\begin{aligned} \mu^2\|z\|^2_{\underline{L}^2(U_r)} + \Lambda^{-1}\|\nabla z\|^2_{\underline{L}^2(U_r)} &\leqslant \mu^2\|z\|_{\underline{L}^2(U_r)}\|w-\bar{v}\|_{\underline{L}^2(U_r)} \\ &\quad + \|z\|_{\underline{H}^1(U_r)}\|\nabla\cdot\mathbf{a}\nabla w - \nabla\cdot\bar{\mathbf{a}}\nabla\bar{v}\|_{\underline{H}^{-1}(U_r)} \\ \text{(Young's inequality)} &\leqslant \mu^2\|z\|^2_{\underline{L}^2(U_r)} + \frac{\mu^2}{4}\|w-\bar{v}\|^2_{\underline{L}^2(U_r)} \\ &\quad + \frac{\Lambda^{-1}}{2}\|z\|^2_{\underline{H}^1(U_r)} + \frac{\Lambda}{2}\|\nabla\cdot\mathbf{a}\nabla w - \nabla\cdot\bar{\mathbf{a}}\nabla\bar{v}\|^2_{\underline{H}^{-1}(U_r)} \\ \Longrightarrow \|\nabla z\|_{\underline{L}^2(U_r)} &\leqslant \Lambda\|\nabla\cdot\mathbf{a}\nabla w - \nabla\cdot\bar{\mathbf{a}}\nabla\bar{v}\|_{\underline{H}^{-1}(U_r)} + \sqrt{\Lambda}\mu\|w-\bar{v}\|_{\underline{L}^2(U_r)}. \end{aligned}$$

We use Poincaré's inequality to conclude that

$$\|z\|_{\underline{H}^1(U_r)} \leqslant C(U)\left(\Lambda\|\nabla\cdot\mathbf{a}\nabla w - \nabla\cdot\bar{\mathbf{a}}\nabla\bar{v}\|_{\underline{H}^{-1}(U_r)} + \sqrt{\Lambda}\mu\|w-\bar{v}\|_{\underline{L}^2(U_r)}\right). \tag{2.35}$$

- **Estimate for b.** To estimate b we use the property that it is the optimizer of the problem

$$\inf_{\chi\in b+H_0^1(U_r)} \mu^2\int_{U_r}\chi^2 + \int_{U_r}\nabla\chi\cdot\mathbf{a}\nabla\chi.$$

So we give an upper bound of this functional by the following trial function

$$T_\lambda := \left(\mathbf{1}_{\{\mathbb{R}^d \setminus U_{r,2\ell(\lambda)}\}} \star \zeta_{\ell(\lambda)}\right) \sum_{k=1}^{d} \partial_{x_k}(\bar{v} \star \zeta)\phi_{e_k}^{(\lambda)}, \tag{2.36}$$

where $U_{r,2\ell(\lambda)}$ is defined as

$$U_{r,2\ell(\lambda)} = \{x \in U_r | d(x, \partial U_r) > 2\ell(\lambda)\}.$$

The motivation to test T_λ is the following: If we think the solution of elliptic equation is an average in some sense of the boundary value, then when the coefficient is oscillating, the boundary value is hard to propagate. So one naive candidate is just smoothing the boundary value in a small band of length $2\ell(\lambda)$.

By comparison,

$$\begin{aligned} \mu^2 \int_{U_r} |b|^2 + \int_{U_r} \nabla b \cdot \mathbf{a}\nabla b &\leqslant \mu^2 \int_{U_r} |T_\lambda|^2 + \int_{U_r} \nabla T_\lambda \cdot \mathbf{a}\nabla T_\lambda \\ \Longrightarrow \|\nabla b\|_{\underline{L}^2(U_r)} &\leqslant \mu\sqrt{\Lambda}\|T_\lambda\|_{\underline{L}^2(U_r)} + \Lambda\|\nabla T_\lambda\|_{\underline{L}^2(U_r)}. \end{aligned}$$

Moreover, to estimate the L^2 norm, we use once again Poincaré's inequality

$$\begin{aligned} \|b\|_{\underline{L}^2(U_r)} &= \|b - T_\lambda + T_\lambda\|_{\underline{L}^2(U_r)} \\ &\leqslant \|b - T_\lambda\|_{\underline{L}^2(U_r)} + \|T_\lambda\|_{\underline{L}^2(U_r)} \\ \text{(Poincaré's inequality)} \quad &\leqslant r\|\nabla(b - T_\lambda)\|_{\underline{L}^2(U_r)} + \|T_\lambda\|_{\underline{L}^2(U_r)} \\ &\leqslant r\|\nabla b\|_{\underline{L}^2(U_r)} + r\|\nabla T_\lambda\|_{\underline{L}^2(U_r)} + \|T_\lambda\|_{\underline{L}^2(U_r)}. \end{aligned}$$

We combine the two and get an estimate of b

$$\begin{aligned} \|b\|_{\underline{H}^1(U_r)} &= \frac{1}{|U_r|^{\frac{1}{d}}}\|b\|_{\underline{L}^2(U_r)} + \|\nabla b\|_{\underline{L}^2(U_r)} \\ &\leqslant C(U)\left(\frac{1}{r}\|T_\lambda\|_{\underline{L}^2(U_r)} + \mu\sqrt{\Lambda}\|T_\lambda\|_{\underline{L}^2(U_r)} + (1+\Lambda)\|\nabla T_\lambda\|_{\underline{L}^2(U_r)}\right). \end{aligned}$$

Finally, we put all the estimates above into eq. (2.34)

$$\begin{aligned} \|v - w\|_{\underline{H}^1(U_r)} \leqslant C(U,\Lambda)\Big(&\|\nabla \cdot (\mathbf{a}\nabla w - \bar{\mathbf{a}}\nabla\bar{v})\|_{\underline{H}^{-1}(U_r)} + \mu\|w - \bar{v}\|_{\underline{L}^2(U_r)} \\ &+ \|\nabla T_\lambda\|_{\underline{L}^2(U_r)} + \left(\frac{1}{r} + \mu\right)\|T_\lambda\|_{\underline{L}^2(U_r)}\Big). \end{aligned} \tag{2.37}$$

To complete the proof of Theorem 2.4.1, we have to treat the random norms in eq. (2.37) respectively. It is the main task of the next section. □

2.4.2 Construction of a vector field

In this part, we analyze $\|\nabla \cdot (\mathbf{a}\nabla w - \bar{\mathbf{a}}\nabla\bar{v})\|_{\underline{H}^{-1}(U_r)}$ with the help of the *flux corrector*. Similar formulas appear both in [22, Lemma 3.3] and [25, Chapter 6, Lemma 6.7], here we give the version in our context.

At first, we introduce the *flux corrector* $\mathbf{S}_p$. For every $p \in \mathbb{R}^d$, since $\mathbf{g}_p := \mathbf{a}(p + \nabla\phi_p) - \bar{\mathbf{a}}p$ defines a divergence free field, i.e. $\nabla \cdot \mathbf{g}_p = 0$, it admits a representation as the "curl" of some potential vector by Helmholtz's theorem. That is there exists a $\mathbb{R}^{d\times d}$ skew-symmetric matrix $\mathbf{S}_p$ such that

$$\mathbf{a}(p+\nabla\phi_p)-\bar{\mathbf{a}}p=\nabla\cdot\mathbf{S}_p,$$

where $\nabla\cdot\mathbf{S}_p$ is a $\mathbb{R}^d$ valued vector defined by $(\nabla\cdot\mathbf{S}_p)_i=\sum_{j=1}^d\partial_{x_j}\mathbf{S}_{p,ij}$. In order to "fix the gauge", for each $i,j\in\{1,2,\cdots d\}$, we force

$$\Delta\mathbf{S}_{p,ij}=\partial_{x_j}\mathbf{g}_{p,i}-\partial_{x_i}\mathbf{g}_{p,j},$$

and under this condition, $\mathbf{S}_p$ is unique up to a constant and $\{\nabla\mathbf{S}_p,ij,p\in\mathbb{R}^d,i,j\in\{1,2,\cdots d\}\}$ also forms a family of $\mathbb{Z}^d$-stationary field like $\nabla\phi_p$. This quantity appears in the early work of periodic homogenization [33, Lemma 3.1] and [145, Lemma 4.4]. See [25, Lemma 6.1] for the well-definedness of $\mathbf{S}_p$ in stochatic homogenization setting and this quantity is also adapted in [121, Lemma 1] as *extended corrector*. We also define

$$\mathbf{S}_p^{(\lambda)}=\mathbf{S}_p-\mathbf{S}_p\star\Phi_{\lambda^{-1}},$$

to truncate the constant part, so $\mathbf{S}_p^{(\lambda)}$ is well-defined and $\mathbb{Z}^d$-stationary.

We have the following identity.

Lemma 2.4.1. *For $\lambda>0,\bar{v}\in H_0^1(U_r)\cap H^2(U_r)$ and $w\in H^1(U_r)$ as in Theorem 2.4.1. We construct a vector field $\mathbf{F}$ such that*

$$\nabla\cdot(\mathbf{a}\nabla w-\bar{\mathbf{a}}\nabla\bar{v})=\nabla\cdot\mathbf{F},$$

whose i-th component is given by

$$\begin{aligned}\mathbf{F}_i=&\sum_{j=1}^d(\mathbf{a}_{ij}-\bar{\mathbf{a}}_{ij})\partial_{x_j}(\bar{v}-\bar{v}\star\zeta)+\sum_{j,k=1}^d\left(\mathbf{a}_{ij}\phi_{e_k}^{(\lambda)}-\mathbf{S}_{e_k,ij}^{(\lambda)}\right)\partial_{x_j}\partial_{x_k}(\bar{v}\star\zeta)\\&+\sum_{j,k=1}^d\left(\partial_{x_j}\mathsf{S}_{e_k,ij}\star\Phi_{\lambda^{-1}}-\mathbf{a}_{ij}\partial_{x_j}\phi_{e_k}\star\Phi_{\lambda^{-1}}\right)\partial_{x_k}(\bar{v}\star\zeta).\end{aligned}\tag{2.38}$$

Proof. We develop

$$\begin{aligned}[\mathbf{a}\nabla w-\bar{\mathbf{a}}\nabla\bar{v}]_i&=\left[\mathbf{a}\nabla\left(\bar{v}+\sum_{k=1}^d\partial_{x_k}(\bar{v}\star\zeta)\phi_{e_k}^{(\lambda)}\right)-\bar{\mathbf{a}}\nabla\bar{v}\right]_i\\&=\left[(\mathbf{a}-\bar{\mathbf{a}})\nabla\bar{v}+\mathbf{a}\nabla\sum_{k=1}^d\left(\partial_{x_k}(\bar{v}\star\zeta)\phi_{e_k}^{(\lambda)}\right)\right]_i\\&=\underbrace{[(\mathbf{a}-\bar{\mathbf{a}})\nabla(\bar{v}-\bar{v}\star\zeta)]_i}_{\mathbf{I}}+\underbrace{\left[(\mathbf{a}-\bar{\mathbf{a}})\nabla(\bar{v}\star\zeta)+\mathbf{a}\nabla\sum_{k=1}^d\left(\partial_{x_k}(\bar{v}\star\zeta)\phi_{e_k}^{(\lambda)}\right)\right]_i}_{\mathbf{II}}.\end{aligned}$$

The first term is indeed

$$\mathbf{I}=[(\mathbf{a}-\bar{\mathbf{a}})\nabla(\bar{v}-\bar{v}\star\zeta)]_i=\sum_{j=1}^d(\mathbf{a}_{ij}-\bar{\mathbf{a}}_{ij})\partial_{x_j}(\bar{v}-\bar{v}\star\zeta),$$

as in the right hand side of the identity, so we continue to study the rest of the formula.

$$\begin{aligned}
\mathbf{II} &= \sum_{j=1}^{d}(\mathbf{a}_{ij}-\bar{\mathbf{a}}_{ij})\partial_{x_j}(\bar{v}\star\zeta) + \sum_{j,k=1}^{d}\mathbf{a}_{ij}\partial_{x_j}\left(\partial_{x_k}(\bar{v}\star\zeta)\phi_{e_k}^{(\lambda)}\right) \\
&= \sum_{j,k=1}^{d}(\mathbf{a}_{ij}-\bar{\mathbf{a}}_{ij})\partial_{x_j}(\bar{v}\star\zeta)\delta_{jk} + \sum_{j,k=1}^{d}\mathbf{a}_{ij}\partial_{x_j}\partial_{x_k}(\bar{v}\star\zeta)\phi_{e_k}^{(\lambda)} + \sum_{j,k=1}^{d}\mathbf{a}_{ij}\partial_{x_k}(\bar{v}\star\zeta)\partial_{x_j}\phi_{e_k}^{(\lambda)} \\
&= \underbrace{\sum_{j,k=1}^{d}\left((\mathbf{a}_{ij}-\bar{\mathbf{a}}_{ij})\delta_{jk}+\mathbf{a}_{ij}\partial_{x_j}\phi_{e_k}^{(\lambda)}\right)\partial_{x_k}(\bar{v}\star\zeta)}_{\mathbf{II.1}} + \underbrace{\sum_{j,k=1}^{d}\mathbf{a}_{ij}\partial_{x_j}\partial_{x_k}(\bar{v}\star\zeta)\phi_{e_k}^{(\lambda)}}_{\mathbf{II.2}}.
\end{aligned}$$

$\mathbf{II.2}$ appears in the right hand side of the formula, so it remains $\mathbf{II.1}$ to treat. We use the the definition of $\mathbf{S}_{e_k}^{(\lambda)}$ in $\mathbf{II.1}$

$$\begin{aligned}
\mathbf{II.1} &= \sum_{k=1}^{d}\left[\mathbf{a}(e_k+\nabla\phi_{e_k}^{(\lambda)})-\bar{\mathbf{a}}e_k\right]_i\partial_{x_k}(\bar{v}\star\zeta) \\
&= \sum_{k=1}^{d}\left[\mathbf{a}(e_k+\nabla\phi_{e_k})-\bar{\mathbf{a}}e_k-\mathbf{a}\nabla\phi_{e_k}\star\Phi_{\lambda^{-1}}\right]_i\partial_{x_k}(\bar{v}\star\zeta) \\
&= \sum_{k=1}^{d}\left[\nabla\cdot\mathbf{S}_{e_k}-\mathbf{a}\nabla\phi_{e_k}\star\Phi_{\lambda^{-1}}\right]_i\partial_{x_k}(\bar{v}\star\zeta) \\
&= \underbrace{\sum_{k=1}^{d}\left[\nabla\cdot\mathbf{S}_{e_k}^{(\lambda)}\right]_i\partial_{x_k}(\bar{v}\star\zeta)}_{\mathbf{III}} + \sum_{k=1}^{d}\left[\nabla\cdot\mathbf{S}_{e_k}\star\Phi_{\lambda^{-1}}-\mathbf{a}\nabla\phi_{e_k}\star\Phi_{\lambda^{-1}}\right]_i\partial_{x_k}(\bar{v}\star\zeta).
\end{aligned}$$

All the terms match well except **III**, where we have to look for an equal form after divergence. Thanks to the property of skew-symmetry, we have

$$\begin{aligned}
\nabla\cdot\mathbf{III} &= \nabla\cdot\left(\sum_{k=1}^{d}\left[\nabla\cdot\mathbf{S}_{e_k}^{(\lambda)}\right]_i\partial_{x_k}(\bar{v}\star\zeta)\right) \\
&= \sum_{i,j,k=1}^{d}\partial_{x_i}\left(\partial_{x_j}\mathbf{S}_{e_k,ij}^{(\lambda)}\partial_{x_k}(\bar{v}\star\zeta)\right) \\
\text{(Integration by parts)} &= \sum_{i,j,k=1}^{d}\partial_{x_i}\partial_{x_j}\left(\mathbf{S}_{e_k,ij}^{(\lambda)}\partial_{x_k}(\bar{v}\star\zeta)\right)-\partial_{x_i}\left(\mathbf{S}_{e_k,ij}^{(\lambda)}\partial_{x_j}\partial_{x_k}(\bar{v}\star\zeta)\right) \\
\text{(Skew-symmetry of }\mathbf{S}) &= -\nabla\cdot\left(\sum_{j,k=1}^{d}\mathbf{S}_{e_k,ij}^{(\lambda)}\partial_{x_j}\partial_{x_k}(\bar{v}\star\zeta)\right).
\end{aligned}$$

This finishes the proof. □

2.4.3 Quantitative description of $\phi_{e_k}^{(\lambda)}$ and $\mathbf{S}_{e_k}^{(\lambda)}$

In this subsection, we will give some quantitative descriptions of $\phi_{e_k}^{(\lambda)}$ and $\mathbf{S}_{e_k}^{(\lambda)}$, which serve as the bricks to form $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1$.

Lemma 2.4.2 (Estimate of corrector)**.** *For each $s \in (0,2)$, there exists a positive constant $C(s,\Lambda,d) < \infty$ such that for every $\lambda \in (0,1)$, $i,j,k \in \{1,\cdots d\}, z \in \mathbb{Z}^d$*

$$\begin{aligned}
\|\nabla\phi_{e_k}\star\Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)} \leqslant \mathcal{O}_s(C\lambda^{\frac{d}{2}}), &\quad \|\nabla\mathbf{S}_{e_k,ij}\star\Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)} \leqslant \mathcal{O}_s(C\lambda^{\frac{d}{2}}), \\
\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)} \leqslant \mathcal{O}_s(C\ell(\lambda)), &\quad \|\mathbf{S}_{e_k,ij}^{(\lambda)}\|_{\underline{L}^2(z+\square)} \leqslant \mathcal{O}_s(C\ell(\lambda)).
\end{aligned} \tag{2.39}$$

Proof. We study at first the part $\phi_{e_k}^{(\lambda)}$. [25, Theorem 4.1] gives us three useful estimates

- $d \geqslant 2, r > 1$, for any $x \in \mathbb{R}^d$

$$|\nabla \phi_{e_k} \star \Phi_r(x)| \leqslant \mathcal{O}_s(C(s,d,\Lambda) r^{-\frac{d}{2}}). \tag{2.40}$$

- $d \geqslant 3$,

$$\|\phi_{e_k}\|_{\underline{L}^2(\square)} \leqslant \mathcal{O}_2(C(d,\Lambda)). \tag{2.41}$$

- $d = 2$, for any $2 \leqslant r < R < \infty$, and $x, y \in \mathbb{R}^d$,

$$\|\phi_{e_k} - \phi_{e_k} \star \Phi_r(0)\|_{\underline{L}^2(\square_r)} \leqslant \mathcal{O}_s(C(s,\Lambda) \log^{\frac{1}{2}} r), \tag{2.42a}$$

$$|(\phi_{e_k} \star \Phi_r)(x) - (\phi_{e_k} \star \Phi_R)(y)| \leqslant \mathcal{O}_s\left(C(s,\Lambda) \log^{\frac{1}{2}} \left(2 + \frac{R + |x-y|}{r}\right)\right). \tag{2.42b}$$

Informally speaking, the behavior of corrector when $d \geqslant 3$ is of size constant, but has a logarithm increment when $d = 2$ and this explicates why we have $\ell(\lambda)$ in eq. (2.39).

1. Proof of $\|\nabla \phi_{e_k} \star \Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)} \leqslant \mathcal{O}_s(C\lambda^{\frac{d}{2}})$.

   By choosing $r = \lambda^{-1}$ in eq. (2.40) and using eq. (2.15), we have

$$\begin{aligned} |\nabla \phi_{e_k} \star \Phi_{\lambda^{-1}}(x)| &\leqslant \mathcal{O}_s(C\lambda^{\frac{d}{2}}) \\ \Longrightarrow |\nabla \phi_{e_k} \star \Phi_{\lambda^{-1}}(x)|^2 &\leqslant \mathcal{O}_{s/2}(C^2\lambda^d) \\ \Longrightarrow \int_{z+\square} |\nabla \phi_{e_k} \star \Phi_{\lambda^{-1}}(x)|^2 \, dx &\leqslant \mathcal{O}_{s/2}(C^2\lambda^d) \\ \Longrightarrow \|\nabla \phi_{e_k} \star \Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)} &\leqslant \mathcal{O}_s(C\lambda^{\frac{d}{2}}). \end{aligned}$$

2. Proof that if $d \geqslant 3$, then $\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)} \leqslant \mathcal{O}_s(C)$. We apply eq. (2.41) to get that

$$\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)} \leqslant \|\phi_{e_k}\|_{\underline{L}^2(z+\square)} + \|\phi_{e_k} \star \Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)},$$

   where the first one comes from a modified version of eq. (2.41). In fact, by the $\mathbb{Z}^d$-stationarity of ϕ Definition 2.4.1, we have $\|\phi_{e_k}\|_{\underline{L}^2(z+\square)} \leqslant \mathcal{O}_2(C(d,\Lambda))$, then for every $s \in (0,2)$, we set $C(s,d,\Lambda) = 3^{\frac{1}{s}} C(d,\Lambda)$, then

$$\begin{aligned} \mathbb{E}\left[\exp\left(\left(\frac{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}}{C(s,d,\Lambda)}\right)_+^s\right)\right] &= \mathbb{E}\left[\exp\left(\frac{1}{3}\left(\frac{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}}{C(d,\Lambda)}\right)^s\right)\right] \\ &= \mathbb{E}\left[\left(\exp\left(\left(\frac{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}}{C(d,\Lambda)}\right)^s\right)\right)^{\frac{1}{3}}\right] \\ (\text{Jensen's inequality}) &\leqslant \left(\mathbb{E}\left[\exp\left(\left(\frac{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}}{C(d,\Lambda)}\right)^s\right)\right]\right)^{\frac{1}{3}}. \end{aligned}$$

We decompose the last term into two parts

$$
\begin{aligned}
&\mathbb{E}\left[\exp\left(\left(\frac{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}}{C(s,d,\Lambda)}\right)_+^s\right)\right] \\
\leqslant\ & \left(\mathbb{E}\left[\exp\left(\left(\frac{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}}{C(d,\Lambda)}\right)^s\right)\mathbf{1}_{\left\{\{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}\leqslant C(d,\Lambda)\}\right\}}\right]\right. \\
&\left.+\ \mathbb{E}\left[\exp\left(\left(\frac{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}}{C(d,\Lambda)}\right)^s\right)\mathbf{1}_{\left\{\{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}\geqslant C(d,\Lambda)\}\right\}}\right]\right)^{\frac{1}{3}} \\
\leqslant\ & \left(\underbrace{\mathbb{E}\left[\exp\left(\left(\frac{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}}{C(d,\Lambda)}\right)^s\right)\mathbf{1}_{\left\{\{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}\leqslant C(d,\Lambda)\}\right\}}\right]}_{\leqslant e}\right. \\
&\left.+\ \underbrace{\mathbb{E}\left[\exp\left(\left(\frac{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}}{C(d,\Lambda)}\right)^2\right)\mathbf{1}_{\left\{\{\|\phi_{e_k}\|_{\underline{L}^2(z+\square)}\geqslant C(d,\Lambda)\}\right\}}\right]}_{\leqslant 2}\right)^{\frac{1}{3}} \\
\leqslant\ & (e+2)^{\frac{1}{3}} \leqslant 2.
\end{aligned}
$$

Thus we prove that for any $s \in (0,2), C(s,d,\Lambda) = 3^{\frac{1}{s}}C(d,\Lambda)$

$$
\|\phi_{e_k}\|_{\underline{L}^2(z+\square)} \leqslant \mathcal{O}_2(C(d,\Lambda)) \Longrightarrow \|\phi_{e_k}\|_{\underline{L}^2(z+\square)} \leqslant \mathcal{O}_s(C(s,d,\Lambda)). \tag{2.43}
$$

We focus on the second one that

$$
\begin{aligned}
\|\phi_{e_k} \star \Phi_{\lambda^{-1}}\|^2_{\underline{L}^2(z+\square)} &= \int_{z+\square}\left|\int_{\mathbb{R}^d}\phi_{e_k}(x-y)\Phi_{\lambda^{-1}}(y)\,dy\right|^2 dx \\
&= \int_{z+\square}\left|\int_{\mathbb{R}^d}\phi_{e_k}(x-y)\Phi^{\frac{1}{2}}_{\lambda^{-1}}(y)\Phi^{\frac{1}{2}}_{\lambda^{-1}}(y)\,dy\right|^2 dx \\
\text{(Hölder's inequality)} &\leqslant \int_{z+\square}\left(\int_{\mathbb{R}^d}\phi^2_{e_k}(x-y)\Phi_{\lambda^{-1}}(y)\,dy\right)\underbrace{\left(\int_{\mathbb{R}^d}\Phi_{\lambda^{-1}}(y)\,dy\right)}_{=1} dx \\
&= \int_{z+\square}\int_{\mathbb{R}^d}\phi^2_{e_k}(x-y)\Phi_{\lambda^{-1}}(y)\,dy\,dx \\
(eq.\ (2.15)) &\leqslant \mathcal{O}_s(C).
\end{aligned}
$$

In the last step, we treat $\Phi_{\lambda^{-1}}$ as a weight for different small cubes so we could apply eq. (2.15) and eq. (2.43).

3. Proof that if $d = 2$, then $\|\phi^{(\lambda)}_{e_k}\|_{\underline{L}^2(z+\square)} \leqslant \mathcal{O}_s(C\ell(\lambda))$.

   This part is a little more difficult than the case $d \geqslant 3$ since we have only eq. (2.42a) and eq. (2.42b) instead of eq. (2.41) when $d = 2$. This forces us to do more steps of difference.

We apply eq. (2.42a) eq. (2.42b) to $\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)}$ for any $\lambda \in (0, \frac{1}{2}]$

$$
\begin{aligned}
\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)} &= \|\phi_{e_k} - \phi_{e_k} \star \Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)} \\
&\leqslant \|\phi_{e_k} - \phi_{e_k} \star \Phi_2(z)\|_{\underline{L}^2(z+\square)} + \|\phi_{e_k} \star \Phi_2(z) - \phi_{e_k} \star \Phi_2\|_{\underline{L}^2(z+\square)} \\
&\quad + \|\phi_{e_k} \star \Phi_2 - \phi_{e_k} \star \Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)} \\
&\leqslant 4\underbrace{\|\phi_{e_k} - \phi_{e_k} \star \Phi_2(z)\|_{\underline{L}^2(B_2(z))}}_{\text{Apply } eq.\ (2.42a)} + 4\underbrace{\|\phi_{e_k} \star \Phi_2(z) - \phi_{e_k} \star \Phi_2\|_{\underline{L}^2(B_2(z))}}_{\text{Apply } eq.\ (2.42b)} \\
&\quad + \underbrace{\|\phi_{e_k} \star \Phi_2 - \phi_{e_k} \star \Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)}}_{\text{Apply } eq.\ (2.42b)} \\
&\leqslant \mathcal{O}_s(C) + \mathcal{O}_s(C) + \mathcal{O}_s(C \log^{\frac{1}{2}}(2 + (2\lambda)^{-1})) \\
(eq.\ (2.15)) \quad &\leqslant \mathcal{O}_s(C\ell(\lambda)).
\end{aligned}
$$

Here we use Φ_2 since eq. (2.42a) and eq. (2.42b) require that the scale should be bigger than 2. In last step, we use also the condition $\lambda \leqslant \frac{1}{2}$ to give up the constant term.

Since $\mathbf{S}_{e_k}$ has the same type of estimate eq. (2.40),eq. (2.41),eq. (2.42a),eq. (2.42b) as ϕ_{e_k}, see [25, Proposition 6.2], we apply the same procedure to obtain the other half of the eq. (2.39). □

2.4.4 Detailed H^{-1} and boundary layer estimate

In this subection, we complete the proof of Theorem 2.4.1, which remains to give an explicit random variable in the formula eq. (2.37). This requires to analyze several norms like $\|\nabla \cdot (\mathbf{a}\nabla w - \bar{\mathbf{a}}\nabla \bar{v})\|_{\underline{H}^{-1}(U_r)}$, $\|w - \bar{v}\|_{\underline{L}^2(U_r)}$, $\|\nabla T_\lambda\|_{\underline{L}^2(U_r)}$, $\|T_\lambda\|_{\underline{L}^2(U_r)}$. We will make the use of two technical lemmas in Section 2.3 and Lemma 2.4.2 to estimate them.

Estimate of $\|\nabla \cdot (\mathbf{a}\nabla w - \bar{\mathbf{a}}\nabla \bar{v})\|_{\underline{H}^{-1}(U_r)}$

With the help of Lemma 2.4.1, we have

$$\|\nabla \cdot (\mathbf{a}\nabla w - \bar{\mathbf{a}}\nabla \bar{v})\|_{\underline{H}^{-1}(U_r)} = \|\nabla \cdot \mathbf{F}\|_{\underline{H}^{-1}(U_r)} \leqslant \|\mathbf{F}\|_{\underline{L}^2(U_r)},$$

and we use the identity in eq. (2.38) to obtain

$$
\begin{aligned}
\|\mathbf{F}\|_{\underline{L}^2(U_r)} &\leqslant \underbrace{\sum_{j=1}^{d} \|(\mathbf{a} - \bar{\mathbf{a}})\nabla(\bar{v} - \bar{v} \star \zeta)\|_{\underline{L}^2(U_r)}}_{\mathbf{H.1}} \\
&\quad + \underbrace{\sum_{j,k=1}^{d} \left\| \left(\mathbf{a}\phi_{e_k}^{(\lambda)} - \mathbf{S}_{e_k}^{(\lambda)} \right) \partial_{x_j} \partial_{x_k} (\bar{v} \star \zeta) \right\|_{\underline{L}^2(U_r)}}_{\mathbf{H.2}} \\
&\quad + \underbrace{\sum_{k=1}^{d} \|(\nabla \mathbf{S}_{e_k} \star \Phi_{\lambda^{-1}} - \mathbf{a}\nabla\phi_{e_k} \star \Phi_{\lambda^{-1}})\, \partial_{x_k}(\bar{v} \star \zeta)\|_{\underline{L}^2(U_r)}}_{\mathbf{H.3}}.
\end{aligned}
$$

We treat the three terms respectively. For **H.1**, we recall that $\bar{v} \star \zeta$ means $\mathsf{Ext}(\bar{v}) \star \zeta$ and use the approximation of identity, see for example [25, Lemma 6.8]

$$\mathbf{H.1} \leqslant \frac{d\Lambda}{|U_r|^{\frac{d}{2}}} \|\nabla\, \mathsf{Ext}(\bar{v}) - \nabla\, \mathsf{Ext}(\bar{v}) \star \zeta\|_{L^2(\mathbb{R}^d)} \leqslant \frac{d\Lambda}{|U_r|^{\frac{d}{2}}} \|\nabla^2\, \mathsf{Ext}(\bar{v})\|_{L^2(\mathbb{R}^d)}.$$

We recall the estimate eq. (2.17) that

$$\frac{d\Lambda}{|U_r|^{\frac{d}{2}}}\|\nabla^2 \mathsf{Ext}(\bar{v})\|_{L^2(\mathbb{R}^d)} \leqslant C(U,\Lambda,d)\|\bar{v}\|_{\underline{H}^2(U_r)},$$

so we get $\mathbf{H.1} \leqslant C(U,\Lambda,d)\|\bar{v}\|_{\underline{H}^2(U_r)}$.

For $\mathbf{H.2}$, since $\|\phi^{(\lambda)}_{e_k}\|_{\underline{L}^2(z+\square)}, \|\mathbf{S}^{(\lambda)}_{e_k}\|_{\underline{L}^2(z+\square)}$ are obtained in Lemma 2.4.2, we could use the Lemma 2.3.1 where we treat the cell of the scale $\varepsilon = 1$ and take $g = \partial_{x_j}\partial_{x_k}\mathsf{Ext}(\bar{v})$, and $f = (\mathbf{a}\phi^{(\lambda)}_{e_k} - \mathbf{S}^{(\lambda)}_{e_k})$

$$\begin{aligned}
\mathbf{H.2} &= \sum_{j,k=1}^{d} \left\| \left(\mathbf{a}\phi^{(\lambda)}_{e_k} - \mathbf{S}^{(\lambda)}_{e_k}\right)(\partial_{x_j}\partial_{x_k}\mathsf{Ext}(\bar{v}) \star \zeta)\right\|_{\underline{L}^2(U_r)} \\
&\leqslant \frac{C(\Lambda,d)}{|U_r|^{\frac{d}{2}}}\sum_{k=1}^{d}\max_{z\in\mathbb{Z}^d\cap(U_r+\square)}\left(\|\phi^{(\lambda)}_{e_k}\|_{\underline{L}^2(z+\square)} + \|\mathbf{S}^{(\lambda)}_{e_k}\|_{\underline{L}^2(z+\square)}\right)\|\nabla^2\mathsf{Ext}(\bar{v})\|_{L^2(U_r+3\square)}.
\end{aligned}$$

Once again we apply the Sobolev extension estimate eq. (2.17) that

$$|U_r|^{-\frac{d}{2}}\|\nabla^2\mathsf{Ext}(\bar{v})\|_{L^2(U_r+3\square)} \leqslant |U_r|^{-\frac{d}{2}}\|\nabla^2\mathsf{Ext}(\bar{v})\|_{L^2(\mathbb{R}^d)} \leqslant C(U,d)\|\bar{v}\|_{\underline{H}^2(U_r)}.$$

We extract the term of random variable

$$\mathcal{X}_1 := \sum_{k=1}^{d}\max_{z\in\mathbb{Z}^d\cap(U_r+\square)}\left(\|\phi^{(\lambda)}_{e_k}\|_{\underline{L}^2(z+\square)} + \|\mathbf{S}^{(\lambda)}_{e_k}\|_{\underline{L}^2(z+\square)}\right), \tag{2.44}$$

and obtain that $\mathbf{H.2} \leqslant C(U,\Lambda,d)\mathcal{X}_1\|\bar{v}\|_{\underline{H}^2(U_r)}$. Moreover, Lemma 2.3.2 and Lemma 2.4.2 can be applied here to estimate the size of random variables that

$$\mathcal{X}_1 \leqslant \mathcal{O}_s\left(C(U,s,d)\ell(\lambda)(\log r)^{\frac{1}{s}}\right).$$

The above estimation gives a good recipe for the remaining part. For $\mathbf{H.3}$, we have

$$\begin{aligned}
\mathbf{H.3} &= \sum_{k=1}^{d}\left\|(\nabla\mathbf{S}_{e_k}\star\Phi_{\lambda^{-1}} - \mathbf{a}\nabla\phi_{e_k}\star\Phi_{\lambda^{-1}})(\partial_{x_k}\mathsf{Ext}(\bar{v})\star\zeta)\right\|_{\underline{L}^2(U_r)} \\
&\leqslant C(U,\Lambda,d)\mathcal{X}_2\|\bar{v}\|_{\underline{H}^1(U_r)},
\end{aligned}$$

where we extract that

$$\mathcal{X}_2 := \sum_{k=1}^{d}\max_{z\in\mathbb{Z}^d\cap(U_r+\square)}\left(\|\nabla\phi_{e_k}\star\Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)} + \|\nabla\mathbf{S}_{e_k}\star\Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)}\right), \tag{2.45}$$

and we apply Lemma 2.3.2 and Lemma 2.4.2 to get

$$\mathcal{X}_2 \leqslant \mathcal{O}_s\left(C(U,s,d)\lambda^{\frac{d}{2}}(\log r)^{\frac{1}{s}}\right).$$

Combing $\mathbf{H.1}, \mathbf{H.2}, \mathbf{H.3}$, we get

$$\|\nabla\cdot(\mathbf{a}\nabla w - \bar{\mathbf{a}}\nabla\bar{v})\|_{\underline{H}^{-1}(U_r)} \leqslant C(U,\Lambda,d)\left(\|\bar{v}\|_{\underline{H}^2(U_r)} + \|\bar{v}\|_{\underline{H}^2(U_r)}\mathcal{X}_1, + \|\bar{v}\|_{\underline{H}^1(U_r)}\mathcal{X}_2\right). \tag{2.46}$$

Estimate of $\|w - \bar{v}\|_{\underline{L}^2(U_r)}$

For $\|w - \bar{v}\|_{\underline{L}^2(U_r)}$, we use Lemma 2.3.1 and eq. (2.17) to obtain that

$$\begin{aligned}
\|w - \bar{v}\|_{\underline{L}^2(U_r)} &= \|\sum_{k=1}^{d} \phi_{e_k}^{(\lambda)} \partial_{x_k}(\mathsf{Ext}(\bar{v}) \star \zeta)\|_{\underline{L}^2(U_r)} \\
&\leqslant C(U,\Lambda,d) \sum_{k=1}^{d} \max_{z\in\mathbb{Z}^d\cap(U_r+\square)} \left(\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)}\right) \|\bar{v}\|_{\underline{H}^1(U_r)} \\
&\leqslant C(U,\Lambda,d)\|\bar{v}\|_{\underline{H}^1(U_r)}\mathcal{X}_1. \\
\Longrightarrow \|w - \bar{v}\|_{\underline{L}^2(U_r)} &\leqslant C(U,\Lambda,d)\|\bar{v}\|_{\underline{H}^1(U_r)}\mathcal{X}_1. \qquad (2.47)
\end{aligned}$$

Estimate of $\|\nabla T_\lambda\|_{\underline{L}^2(U_r)}, \|T_\lambda\|_{\underline{L}^2(U_r)}$

Finally, we come to the estimate of $\|\nabla T_\lambda\|_{\underline{L}^2(U_r)}, \|T_\lambda\|_{\underline{L}^2(U_r)}$. We study $\|\nabla T_\lambda\|_{\underline{L}^2(U_r)}$ at first.

$$\begin{aligned}
\|\nabla T_\lambda\|_{\underline{L}^2(U_r)} &= \underbrace{\left\|\left(\mathbf{1}_{\{\mathbb{R}^d\setminus U_{r,2l(\lambda)}\}} \star \frac{1}{\ell^{\frac{d}{2}+1}(\lambda)}(\nabla\zeta)\left(\frac{\cdot}{\ell(\lambda)}\right)\right)\sum_{k=1}^{d}\partial_{x_k}(\bar{v}\star\zeta)\phi_{e_k}^{(\lambda)}\right\|_{\underline{L}^2(U_r)}}_{\mathbf{T.1}} \\
&+ \underbrace{\left\|\left(\mathbf{1}_{\{\mathbb{R}^d\setminus U_{r,2\ell(\lambda)}\}} \star \zeta_{\ell(\lambda)}\right)\sum_{k=1}^{d}\partial_{x_k}(\nabla\bar{v}\star\zeta)\phi_{e_k}^{(\lambda)}\right\|_{\underline{L}^2(U_r)}}_{\mathbf{T.2}} \\
&+ \underbrace{\left\|\left(\mathbf{1}_{\{\mathbb{R}^d\setminus U_{r,2\ell(\lambda)}\}} \star \zeta_{\ell(\lambda)}\right)\sum_{k=1}^{d}\partial_{x_k}(\bar{v}\star\zeta)\nabla\phi_{e_k}^{(\lambda)}\right\|_{\underline{L}^2(U_r)}}_{\mathbf{T.3}}.
\end{aligned}$$

For the term **T.1**, we use eq. (2.17) and eq. (2.18) that

$$\begin{aligned}
\mathbf{T.1} &\leqslant C\frac{1}{\ell(\lambda)}\left\|\sum_{k=1}^{d}\mathbf{1}_{\{U_r\setminus U_{r,2\ell(\lambda)}\}}\partial_{x_k}(\mathsf{Ext}(\bar{v})\star\zeta)\phi_{e_k}^{(\lambda)}\right\|_{\underline{L}^2(U_r)} \\
&\leqslant \frac{C(U,d)}{\ell(\lambda)}\sum_{k=1}^{d}\max_{z\in\mathbb{Z}^d\cap(U_r\setminus U_{r,2\ell(\lambda)}+\square)}\left(\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)}\right)\|\mathbf{1}_{\{U_r\setminus U_{r,2\ell(\lambda)}\}}\bar{v}\|_{\underline{H}^1(U_r)}.
\end{aligned}$$

Here we should pay attention to one small improvement: The domain of integration is in fact restricted in $U_r\setminus U_{r,2\ell(\lambda)}$, so we would like to give it a bound in terms of $\underline{H}^2(U_r)$ rather than $\underline{H}^1(U_r)$. We borrow a trace estimate in [22, Proposition A.1] that for $f\in H^1(U_r)$

$$\|f\mathbf{1}_{\{U_r\setminus U_{r,2\ell(\lambda)}\}}\|_{\underline{L}^2(U_r)} \leqslant C(U,d)\ell(\lambda)^{\frac{1}{2}}\|f\|^{\frac{1}{2}}_{\underline{H}^1(U_r)}\|f\|^{\frac{1}{2}}_{\underline{L}^2(U_r)}, \qquad (2.48)$$

using eq. (2.48) then we obtain an estimate

$$\mathbf{T.1} \leqslant \frac{C(U,d)}{\ell^{\frac{1}{2}}(\lambda)}\sum_{k=1}^{d}\max_{z\in\mathbb{Z}^d\cap(U_r\setminus U_{r,2\ell(\lambda)}+\square)}\left(\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)}\right)\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^2(U_r)}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^1(U_r)},$$

so we define the random variable

$$\mathcal{Y}_1 := \sum_{k=1}^{d}\max_{z\in\mathbb{Z}^d\cap(U_r\setminus U_{r,2\ell(\lambda)}+\square)}\left(\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)}\right), \qquad (2.49)$$

and we have the estimate by Lemma 2.3.2 and Lemma 2.4.2

$$\mathcal{Y}_1 \leqslant \mathcal{O}_s\left(C(U,s,d)\ell(\lambda)(\log r)^{\frac{1}{s}}\right).$$

We skip the details since they are analogue to the previous part. **T.3** follows from the same type of estimate as **T.1** and **T.2** is routine where we suffices to apply Lemma 2.3.1 and eq. (2.48). We find that

$$\begin{aligned}
\mathbf{T.2} &\leqslant C(U,d)\|\bar{v}\|_{\underline{H}^2(U_r)}\mathcal{Y}_1,\\
\mathbf{T.3} &\leqslant C(U,d)\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^2(U_r)}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^1(U_r)}\ell^{\frac{1}{2}}(\lambda)\mathcal{X}_2.
\end{aligned}$$

The three estimates of **T.1**, **T.2**, **T.3** implies that

$$\begin{aligned}
\|\nabla T_\lambda\|_{\underline{L}^2(U_r)} \leqslant{}& C(U,d)\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^2(U_r)}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^1(U_r)}\frac{1}{\ell^{\frac{1}{2}}(\lambda)}\mathcal{Y}_1\\
&+C(U,d)\left(\|\bar{v}\|_{\underline{H}^2(U_r)}\mathcal{Y}_1+\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^2(U_r)}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^1(U_r)}\ell^{\frac{1}{2}}(\lambda)\mathcal{X}_2\right).
\end{aligned} \tag{2.50}$$

Finally, we find that $\|T_\lambda\|_{\underline{L}^2(U_r)}$ has been contained in the estimate **T.1** that

$$\|T_\lambda\|_{\underline{L}^2(U_r)} \leqslant C(U,d)\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^2(U_r)}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^1(U_r)}\ell^{\frac{1}{2}}(\lambda)\mathcal{Y}_1. \tag{2.51}$$

eq. (2.37), eq. (2.46), eq. (2.47), eq. (2.50), eq. (2.51) conclude the proof of Theorem 2.4.1. We have

$$\begin{aligned}
\|v-w\|_{\underline{H}^1(U_r)} &\leqslant C(U,\Lambda)\Big(\|\mathbf{a}\nabla w-\bar{\mathbf{a}}\nabla\bar{v}\|_{\underline{H}^{-1}(U_r)}+\mu\|w-\bar{v}\|_{\underline{L}^2(U_r)}\\
&\quad+\|\nabla T_\lambda\|_{\underline{L}^2(U_r)}+(\frac{1}{r}+\mu)\|T_\lambda\|_{\underline{L}^2(U_r)}\Big)\\
&\leqslant C(U,\Lambda,d)\Big(\|\bar{v}\|_{\underline{H}^2(U_r)}+\|\bar{v}\|_{\underline{H}^2(U_r)}\mathcal{X}_1+\|\bar{v}\|_{\underline{H}^1(U_r)}\mathcal{X}_2+\mu\|\bar{v}\|_{\underline{H}^1(U_r)}\mathcal{X}_1\\
&\quad+\ell(\lambda)^{\frac{1}{2}}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^2(U_r)}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^1(U_r)}\left(\left(\mu+\frac{1}{r}+\frac{1}{\ell(\lambda)}\right)\mathcal{Y}_1+\mathcal{X}_2\right)+\|\bar{v}\|_{\underline{H}^2(U_r)}\mathcal{Y}_1\Big)\\
&= C(U,\Lambda,d)\Big[\|\bar{v}\|_{\underline{H}^2(U_r)}+\left(\|\bar{v}\|_{\underline{H}^2(U_r)}+\mu\|\bar{v}\|_{\underline{H}^1(U_r)}\right)\mathcal{X}_1\\
&\quad+\left(\ell(\lambda)^{\frac{1}{2}}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^2(U_r)}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^1(U_r)}+\|\bar{v}\|_{\underline{H}^1(U_r)}\right)\mathcal{X}_2\\
&\quad+\left(\ell(\lambda)^{\frac{1}{2}}\left(\mu+\frac{1}{r}+\frac{1}{\ell(\lambda)}\right)\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^2(U_r)}\|\bar{v}\|^{\frac{1}{2}}_{\underline{H}^1(U_r)}+\|\bar{v}\|_{\underline{H}^2(U_r)}\right)\mathcal{Y}_1\Big].
\end{aligned}$$

We add one table of $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1$ to check the its typical size.

2.5 Iteration estimate

In this part, we use Theorem 2.4.1 to analyze the algorithm, and we give at first an H^1, H^2 *a priori* estimate.

R.V	Expression	$\mathcal{O}_s$ size
$\mathcal{X}_1$	$\sum_{k=1}^d \max_{z\in\mathbb{Z}^d\cap(U_r+\square)} \left(\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)} + \|\mathbf{S}_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)}\right)$	$\mathcal{O}_s\left(C\ell(\lambda)(\log r)^{\frac{1}{s}}\right)$
$\mathcal{X}_2$	$\sum_{k=1}^d \max_{z\in\mathbb{Z}^d\cap(U_r+\square)} \left(\|\nabla\phi_{e_k} \star \Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)} + \|\nabla\mathbf{S}_{e_k} \star \Phi_{\lambda^{-1}}\|_{\underline{L}^2(z+\square)}\right)$	$\mathcal{O}_s\left(C\lambda^{\frac{d}{2}}(\log r)^{\frac{1}{s}}\right)$
$\mathcal{Y}_1$	$\sum_{k=1}^d \max_{z\in\mathbb{Z}^d\cap(U_r\setminus U_{r,2\ell(\lambda)}+\square)} \left(\|\phi_{e_k}^{(\lambda)}\|_{\underline{L}^2(z+\square)}\right)$	$\mathcal{O}_s\left(C\ell(\lambda)(\log r)^{\frac{1}{s}}\right)$

Figure 2.3: A table of random variables $\mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1$.

2.5.1 Proof of an H^1, H^2 estimate

Lemma 2.5.1. *In eq. (2.3), we have a control*

$$\|\bar{u}\|_{\underline{H}^1(U_r)} + \lambda^{-1}\|\bar{u}\|_{\underline{H}^2(U_r)} \leqslant C(U,\Lambda,d)\|v-u\|_{\underline{H}^1(U_r)}.$$

Proof. We test the first equation $(\lambda^2 - \nabla\cdot\mathbf{a}\nabla)u_0 = -\nabla\cdot\mathbf{a}\nabla(u-v)$ in eq. (2.3) by u_0 and use the ellipticity condition to obtain

$$\begin{aligned}
\lambda^2\|u_0\|^2_{L^2(U_r)} + \Lambda^{-1}\|\nabla u_0\|^2_{L^2(U_r)} &\leqslant \lambda^2\|u_0\|^2_{L^2(U_r)} + \int_{U_r} \nabla u_0\cdot\mathbf{a}\nabla u_0 \\
&= \int_{U_r} \nabla u_0\cdot\mathbf{a}\nabla(u-v) \\
&\leqslant \Lambda\|\nabla(v-u)\|_{L^2(U_r)}\|\nabla u_0\|_{L^2(U_r)} \\
\Longrightarrow \|\nabla u_0\|_{\underline{L}^2(U_r)} &\leqslant \Lambda^2\|\nabla(v-u)\|_{\underline{L}^2(U_r)}.
\end{aligned}$$

We put back this term in the inequality, we also obtain that

$$\lambda\|u_0\|_{\underline{L}^2(U_r)} \leqslant \Lambda^{\frac{3}{2}}\|\nabla(v-u)\|_{\underline{L}^2(U_r)}. \tag{2.52}$$

Using this estimate, we obtain that of $\nabla\bar{u}$ by testing $-\nabla\cdot\bar{\mathbf{a}}\nabla\bar{u} = -\nabla\cdot\mathbf{a}\nabla(u-v-u_0)$ with $\bar{u}$

$$\begin{aligned}
\int_{U_r} \nabla\bar{u}\cdot\bar{\mathbf{a}}\nabla\bar{u} &= \int_{U_r} \nabla\bar{u}\cdot\mathbf{a}\nabla(u-v-u_0) \\
\Longrightarrow \|\nabla\bar{u}\|_{\underline{L}^2(U_r)} &\leqslant \Lambda^2\|\nabla(u-v-u_0)\|_{\underline{L}^2(U_r)} \\
&\leqslant \Lambda^2\|\nabla(u-v)\|_{\underline{L}^2(U_r)} + \Lambda^2\|\nabla u_0\|_{\underline{L}^2(U_r)} \\
&\leqslant C(U,\Lambda,d)\|\nabla(u-v)\|_{\underline{L}^2(U_r)}.
\end{aligned}$$

Finally, we calculate the H^2 norm of $\bar{u}$. Because it is the solution of $-\nabla\cdot\bar{\mathbf{a}}\nabla\bar{u} = \lambda^2 u_0$, we apply the classical H^2 estimate of elliptic equation (see [105, Theorem 6.3.2.4])

$$\begin{aligned}
\Lambda^{-1}\|\bar{u}\|_{\underline{H}^2(U_r)} &\leqslant \lambda^2\|u_0\|_{L^2(U_r)} \\
(\text{ Using } \textit{eq. } (2.52)) &\leqslant \lambda\Lambda^{\frac{3}{2}}\|\nabla(v-u)\|_{L^2(U_r)} \\
\Longrightarrow \|\bar{u}\|_{\underline{H}^2(U_r)} &\leqslant \lambda\Lambda^{\frac{5}{2}}\|\nabla(v-u)\|_{\underline{L}^2(U_r)}.
\end{aligned}$$

□

2.5.2 Proof of the main theorem

With all these tools in hand, we can now prove Theorem 2.1.1. We denote by $\mathbb{R}(\lambda, \mu, r, \mathbf{a}, d, U, \bar{v})$ the right hand side of eq. (2.32), that is

$$\|v - w\|_{\underline{H}^1(U_r)} \leqslant \mathbb{R}(\lambda, \mu, r, \mathbf{a}, d, U, \bar{v}).$$

Proof. We take the first and second equations in the eq. (2.3) and use the equation eq. (2.2)

$$\begin{aligned} -\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} &= \lambda^2 u_0 \\ &= f + \nabla \cdot \mathbf{a} \nabla (v + u_0) \\ &= -\nabla \cdot \mathbf{a} \nabla (u - v - u_0). \end{aligned}$$

This is in the frame of Theorem 2.4.1 thanks to the classical H^2 theory that $\bar{u} \in H^2(U_r)$. We apply Theorem 2.4.1 with abuse of notation of the two scale expansion

$$w := \bar{u} + \sum_{k=1}^{d} \partial_{x_k} (\mathsf{Ext}(\bar{u}) \star \zeta) \phi_{e_k}^{(\lambda)},$$

with $\mathsf{Ext}(\bar{u})$ satisfying eq. (2.17). Then we obtain that

$$\|w - (u - v - u_0)\|_{\underline{H}^1(U_r)} \leqslant \mathbb{R}(\lambda, 0, r, \mathbf{a}, d, U, \bar{u}). \tag{2.53}$$

The third equation of eq. (2.3) $(\lambda^2 - \nabla \cdot \mathbf{a} \nabla)\widetilde{u} = (\lambda^2 - \nabla \cdot \bar{\mathbf{a}} \nabla)\bar{u}$ is also of the form of the Theorem 2.4.1, so we obtain

$$\|\widetilde{u} - w\|_{\underline{H}^1(U_r)} \leqslant \mathbb{R}(\lambda, \lambda, r, \mathbf{a}, d, U, \bar{u}). \tag{2.54}$$

We combine this two estimates and use the triangle inequality to obtain

$$\|(v + u_0 + \widetilde{u}) - u\|_{\underline{H}^1(U_r)} \leqslant \mathbb{R}(\lambda, 0, r, \mathbf{a}, d, U, \bar{u}) + \mathbb{R}(\lambda, \lambda, r, \mathbf{a}, d, U, \bar{u}). \tag{2.55}$$

It remains to see how to adapt $\mathbb{R}(\lambda, \mu, r, \mathbf{a}, d, U, \bar{u})$ in a proper way in the context of eq. (2.2). We plug in the formula in Lemma 2.5.1 to separate all the norms of H^1 and H^2 and use $0 < \mu < \lambda$.

$$\begin{aligned} \mathbb{R}(\lambda, \mu, r, \mathbf{a}, d, U, \bar{u}) \leqslant C(U, \Lambda, d) \Big[\lambda + \lambda \mathcal{X}_1 + \left(1 + \ell(\lambda)^{\frac{1}{2}} \lambda^{\frac{1}{2}}\right) \mathcal{X}_2 \\ + \left(\ell(\lambda)^{\frac{1}{2}} \lambda^{\frac{1}{2}} + 1\right) \left(\lambda + \frac{1}{r} + \frac{1}{\ell(\lambda)}\right) \mathcal{Y}_1 \Big] \|v - u\|_{\underline{H}^1(U_r)}. \end{aligned}$$

By checking fig. 2.3 and notice that the largest term is $\ell(\lambda)^{-\frac{1}{2}} \lambda^{\frac{1}{2}} \mathcal{Y}_1$, so we obtain that the factor is of type $\mathcal{O}_s\left(C(U, \Lambda, s, d)(\log r)^{\frac{1}{s}} \ell(\lambda)^{\frac{1}{2}} \lambda^{\frac{1}{2}}\right)$ as desired. □

Chapter 3

AHKM iterative algorithm on percolation clusters

We present an efficient algorithm to solve elliptic Dirichlet problems defined on the cluster of $\mathbb{Z}^d$ supercritical Bernoulli percolation, as a generalization of the iterative method proposed by S. Armstrong, A. Hannukainen, T. Kuusi and J.-C. Mourrat. We also explore the two-scale expansion on the infinite cluster of percolation, and use it to give a rigorous analysis of the algorithm.

This chapter corresponds to the article [134].

3.1 Introduction

3.1.1 Motivation and main result

The main goal of this paper is to study a fast algorithm for computing the solution of Dirichlet problems with random coefficients on Bernoulli percolation clusters. For dimension $d \geqslant 2$, let $(\mathbb{Z}^d, E_d)$ be the Eculidean lattice, where E_d denotes the set of (unoriented) nearest-neighbor *bonds* (or *edges*), that is, two-element sets $\{x, y\}$ with $x, y \in \mathbb{Z}^d$ satisfying $|x - y| = 1$. We also write $x \sim y$ whenever $\{x, y\} \in E_d$. Then we give ourselves a constant $\Lambda > 1$ and a *random conductance* $\mathbf{a} : E_d \to \{0\} \cup [\Lambda^{-1}, 1]$ such that the random variables $\{\mathbf{a}(e)\}_{e \in E_d}$ are independent and identically distributed. The Bernoulli percolation in this work is defined by the random conductance $\{\mathbf{a}(e)\}_{e \in E_d}$: for every bond $e \in E_d$, we say that e is an *open bond* if $\mathbf{a}(e) > 0$, and that e is a *closed bond* otherwise. The connected components on $(\mathbb{Z}^d, E_d)$ generated by the open bonds are called *clusters*, and we are interested in the supercritical percolation case, that is, we assume that $\mathfrak{p} := \mathbb{P}[\mathbf{a}(e) > 0]$ is strictly larger than the critical percolation parameter, which we denote by $\mathfrak{p}_c(d)$. As a consequence, there exists a unique infinite percolation cluster $\mathscr{C}_\infty$ [149].

The configuration of clusters is random in a finite cube $\square_m := \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d \cap \mathbb{Z}^d$. However, under the supercritical percolation setting and when the cube $\square_m$ is large, typically we will see a giant cluster $\mathscr{C}_*(\square_m)$, which takes most of the volume in $\square_m$, and the other clusters are very small. (See Figure 1.1 for an illustration.) We call this situation "$\square_m$ is a good cube" and informally one can think of $\mathscr{C}_*(\square_m)$ as the largest cluster of $\mathscr{C}_\infty \cap \square_m$. The rigorous definitions of "$\square_m$ is a good cube" and of the maximal cluster $\mathscr{C}_*(\square_m)$ will be given in Definitions 3.2.2 and 3.2.5 below, and they are typical since there exists a positive constant $C(d, \mathfrak{p})$ such that

$$\mathbb{P}[\square_m \text{ is a good cube}] \geqslant 1 - C(d, \mathfrak{p}) \exp(-C(d, \mathfrak{p})^{-1} 3^m).$$

Our goal is to find an algorithm for solving Dirichlet problems on $\mathscr{C}_*(\square_m)$. That is, given two functions $f, g : \square_m \to \mathbb{R}$, we aim to define and study an efficient method for calculating the solution u of

$$\begin{cases} -\nabla \cdot \mathbf{a} \nabla u = f & \text{in } \mathscr{C}_*(\square_m), \\ u = g & \text{on } \mathscr{C}_*(\square_m) \cap \partial \square_m, \end{cases} \tag{3.1}$$

where the divergence-form operator is defined as

$$-\nabla \cdot \mathbf{a} \nabla u(x) := \sum_{y \sim x} \mathbf{a}(x, y) \left(u(x) - u(y)\right). \tag{3.2}$$

Equation (3.1) is very natural to describe many models in applied mathematics and other disciplines. For example, one can think of the electric potential in a porous medium: a domain is made of two types of composites, represented respectively by the open bonds and the closed bonds on the lattice graph $(\mathbb{Z}^d, E_d)$, and only the open bonds are available for the current to flow, while the closed bonds are insulating. See [220] for a comprehensive introduction and [95, 163, 175] for some examples of its applications in nanomaterials.

The complex geometry of the percolation cluster causes significant perturbations to the electric potential, and this makes efficient numerical calculations challenging. Naive finite-difference schemes will become very costly as the size of the domain is increased, and the perforated geometry and low regularity of solutions does not allow for simple coarsening mechanisms. As is well-known, using the *effective conductance* $\bar{\mathbf{a}}$, which is a constant matrix (in fact a scalar by the symmetries in our assumptions) whose definition will be recalled in eqs. (3.134) and (3.135), one can replace the heterogeneous operator $-\nabla\cdot\mathbf{a}\nabla$ by the constant-coefficient operator $-\bar{\mathbf{a}}\Delta$ defined by

$$-\bar{\mathbf{a}}\Delta u(x) := \bar{\mathbf{a}} \sum_{y\sim x} (u(x) - u(y)), \tag{3.3}$$

and thus obtain an approximation $\bar{u}$ as the solution of a homogenized equation. This is a nice idea, but the gap between $\bar{u}$ and u always exists: on small scales, the homogenized solution $\bar{u}$ will typically be very smooth, while u has oscillations. Indeed, the homogenized solution $\bar{u}$ can only approximate u in L^2, but not in H^1. Moreover, the L^2 norm of $(u - \bar{u})$ depends on the size of $\square_m$ and only goes to zero in the limit $m \to \infty$. In other words, $\bar{u}$ converges to u in L^2 only in the limit of "infinite separation of scales".

The goal of the present work is to go beyond these limitations: we will devise an algorithm that produces a sequence of approximations which rapidly converges to u in H^1, in a regime of large but finite separation of scales. The main idea is to look for a way to use the homogenized operator as a *coarse operator* in a scheme analogous to a multigrid method. In fact, the algorithm here is at first proposed in [22] by Armstrong, Hannukainen, Kuusi and Mourrat for the same equation under uniform ellipticity condition on $\mathbb{R}^d$, where the authors believe that the their method can be extended to a more degenerate case like percolation model. This generalization is more challenging, since we have to figure out not only the *coarse operator* but also the *projection operator*, which comes from the perturbation of the geometry. We use a new idea of *mask operator* to resolve it, see Section 3.1.3 for more detailed discussions. Thus the present work also confirms the robustness of their algorithm by stating clearly how to adapt it on percolation clusters and giving rigorous analysis for the rate of convergence.

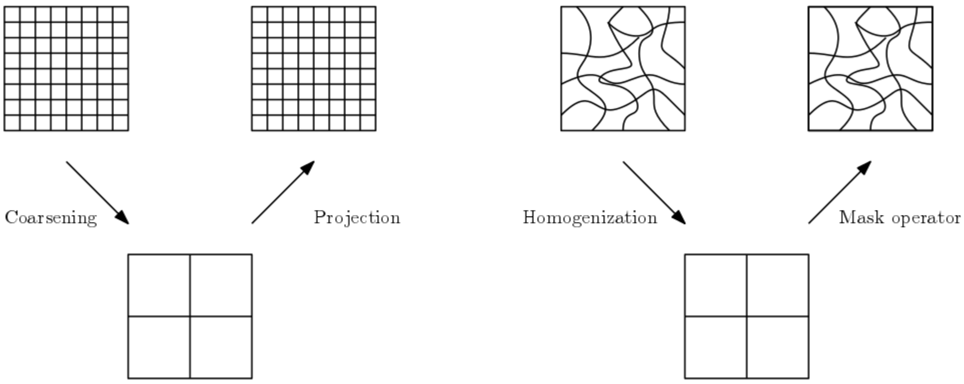

Figure 3.1: The classical multigrid algorithm contains two main steps: coarsening and projection. The algorithm in [22] gives idea how to use homogenization as the coarse operator for random conductance, and this work proposes to use the mask operator as a counterpart of projection in degenerated random conductance case.

Let us introduce some more notations and state the main theorem. For any $V \subseteq \mathbb{Z}^d$, the *interior* of V is defined as $\mathrm{int}(V) := \{x \in V : y \sim x \Rightarrow y \in V\}$, and the *boundary* is defined as $\partial(V) := V \backslash \mathrm{int}(V)$. The function space $C_0(V)$ is the set of functions with zero boundary condition. The L^2 integration of the gradient of v on the percolation cluster is defined as

$$\left\| \nabla v \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{L^2(V)} := \left(\frac{1}{2} \sum_{x,y\in V, x\sim y} (v(y) - v(x))^2 \mathbf{1}_{\{\mathbf{a}(x,y)\neq 0\}} \right)^{\frac{1}{2}}.$$

For any $V \subseteq \mathbb{Z}^d$ we define its associated σ-algebra $\mathcal{F}(V) := \sigma\left(\{\mathbf{a}(e)\}_{e \cap V \neq \emptyset}\right)$ and $\mathcal{F}$ shorthand for $\mathcal{F}_{\mathbb{Z}^d}$. We denote the probability space by $(\{\mathbf{a}(e)\}_{e \in E_d}, \mathcal{F}, \mathbb{P})$. For a random variable X, we use two positive parameters s, θ, and the notation $\mathcal{O}$ to measure its size by

$$X \leqslant \mathcal{O}_s(\theta) \Longleftrightarrow \mathbb{E}\left[\exp((\theta^{-1}X)_+^s)\right] \leqslant 2,$$

where $(\theta^{-1}X)_+ := \max\{\theta^{-1}X, 0\}$. Roughly speaking, the statement $X \leqslant \mathcal{O}_s(\theta)$ tells us that X has a tail lighter than $\exp(-(\theta^{-1}x)^s)$. We also define, for each $\lambda > 0$, the mappings $\lambda_{\mathscr{C},m} : \mathbb{Z}^d \to \mathbb{R}$, and $\ell : \mathbb{R}^+ \to \mathbb{R}^+$ by

$$\lambda_{\mathscr{C},m}(x) := \begin{cases} \lambda & \text{if } x \in \mathscr{C}_*(\square_m), \\ 0 & \text{otherwise.} \end{cases} \qquad \ell(\lambda) := \begin{cases} \log^{\frac{1}{2}}(1 + \lambda^{-1}) & \text{if } d = 2, \\ 1 & \text{if } d > 2. \end{cases} \tag{3.4}$$

Theorem 3.1.1 (Main theorem)**.** *There exist two finite positive constants $s(d, \mathfrak{p}, \Lambda)$, $C(d, \mathfrak{p}, \Lambda, s)$, and for every integer $m > 1$ and $\lambda \in \left(\frac{1}{3^m}, \frac{1}{2}\right)$, an $\mathcal{F}$-measurable random variable $\mathcal{Z}$ satisfying*

$$\mathcal{Z} \leqslant \mathcal{O}_s\left(C\ell(\lambda)^{\frac{1}{2}}\lambda^{\frac{1}{2}}m^{\frac{1}{s}+d}\right),$$

such that the following holds. Let $f, g : \square_m \to \mathbb{R}, u_0 \in g + C_0(\square_m)$ and $u \in g + C_0(\square_m)$ be the solution of eq. (3.1). On the event that $\square_m$ is a good cube, for $u_1, \bar{u}, u_2 \in C_0(\square_m)$ solving (with null Dirichlet boundary condition)

$$\begin{cases} (\lambda^2 - \nabla \cdot \mathbf{a}\nabla)u_1 &= f + \nabla \cdot \mathbf{a}\nabla u_0 & \text{in } \mathscr{C}_*(\square_m)\backslash\partial\square_m, \\ -\nabla \cdot \bar{\mathbf{a}}\nabla\bar{u} &= \lambda^2_{\mathscr{C},m}u_1 & \text{in } \operatorname{int}(\square_m), \\ (\lambda^2 - \nabla \cdot \mathbf{a}\nabla)u_2 &= (\lambda^2 - \nabla \cdot \bar{\mathbf{a}}\nabla)\bar{u} & \text{in } \mathscr{C}_*(\square_m)\backslash\partial\square_m, \end{cases} \tag{3.5}$$

and for $\hat{u} := u_0 + u_1 + u_2$, we have the contraction estimate

$$\left\|\nabla(\hat{u} - u)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} \leqslant \mathcal{Z}\left\|\nabla(u_0 - u)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))}. \tag{3.6}$$

We explain a little more why this theorem ensures the good performance of the algorithm. We are mainly interested in two aspects: the convergence rate of the algorithm and its numerical complexity. To better illustrate the typical size, we denote by $r = 3^m$ the diameter of the domain.

- *Convergence rate of the algorithm.* We start by an arbitrary guess $u_0 \in g + C_0(\square_m)$ as an approximation of u, and repeat the eq. (3.5) several rounds. At the end of every round, we use the $\hat{u}$ just obtained in place of u_0 in the new round of iteration. The contraction rate of iteration has a bound $\mathcal{Z}$, which is a random factor only depending on the conductance $\mathbf{a}$, the choice of our regularization λ, and the size of the cube $\square_m$, but independent of the data f, g, u_0. We can choose λ such that $\frac{1}{r} \ll \lambda \ll (\log r)^{-\left(\frac{1}{s}+d\right)}$, then $\mathcal{Z} \leqslant \mathcal{O}_s\left(C\ell(\lambda)^{\frac{1}{2}}\lambda^{\frac{1}{2}}(\log r)^{\frac{1}{s}+d}\right)$ tells us that $\mathcal{Z}$ has large probability to be smaller than 1 and implies a geometric rate of convergence.

- *Complexity analysis.* The numerical costs come from three parts: the iteration eq. (3.5) itself, the cost to calculate the homogenized coefficient $\bar{\mathbf{a}}$ and to determine the maximal cluster $\mathscr{C}_*(\square_m)$. Sometimes we also omit that the last two are already known, as they cost less numerical resources compared to the first one.

- *The cost for the iteration.* We notice that if we solve a Dirichlet problem for $-\nabla\cdot\mathbf{a}\nabla$ naively, for a precision ε it requires total $O(r\log(\varepsilon^{-1}))$ iterations of *conjugate gradient method* (CGD), while for the problem with regularization like the first and third equation in eq. (3.5), it can be reduced to $O(\lambda^{-1}\log(\varepsilon^{-1}))$ rounds of CGM. The second equation in eq. (3.5) can be solved by a standard multigrid algorithm with $O(\log(\varepsilon^{-1}))$ iterations of CGM (see [65, Chapter 4]). Therefore, applying eq. (3.5) with more detailed choice of resolution in every iteration, it allows us to solve the problem for precision ε with $O((\log r)^{\left(\frac{1}{s}+d\right)}(\log(\varepsilon^{-1}))^2)$ rounds of CGM; see Section 2.1.2 for details.
- *The homogenized coefficient* $\bar{\mathbf{a}}$. There exist many excellent methods to calculate $\bar{\mathbf{a}}$ quickly, which can be naturally generalized to the percolation setting; see for example [119, 102, 184, 107, 138]. The result from [184, Proposition 1.1] tells us the best precision in a domain of size r is $\varepsilon = r^{-\frac{d}{2}}$ with $O(r^d\log r)$ operations, which corresponds to about $O(\log r)$ rounds of CGM.
- *The maximal cluster* $\mathscr{C}_*(\square_m)$. This is a supplementary step compared to the problem on $\mathbb{R}^d$, and one can use the “UnionFind” algorithm [79, Chapter 21] which requires at most $O(r^d\log r)$ operations, which corresponds to about $O(\log r)$ rounds of CGM.

In conclusion, from the discussion above we know the limit for the precision is about $\varepsilon = r^{-n}$ for $n \leqslant \frac{d}{2}$, thus our algorithm does reduce the numerical complexity.

The rest of this paper focuses more on the theoretical proof of Theorem 3.1.1 and we add two remarks to conclude the introduction part. Firstly, eq. (3.1) can be defined in a more general domain $\mathbb{Z}^d \cap U_r$ where U is a convex domain with $C^{1,1}$ boundary, $r > 0$ is a length scale which we think of as being large, and $U_r := \{rx | x \in U\}$. In this case $\mathscr{C}_*(U_r)$ can be informally thought as the largest cluster in U_r. Our iterative algorithm eq. (3.5) and its analysis can be adapted to this more general setting by following very similar arguments.

Secondly, in eq. (3.1) one can simply write $-\nabla\cdot\bar{\mathbf{a}}\nabla$ as $-\bar{\mathbf{a}}\Delta$ defined in eq. (3.3) as $\bar{\mathbf{a}}$ is in fact a scalar coefficient. However, for some other models like inhomogeneous percolation (see [131, Chapter 11.9] and [132]), where $\bar{\mathbf{a}}$ can be an effective matrix rather than a scalar. One example is in $\mathbb{Z}^2$, we choose two different parameters p_1, p_2 and $p_1 + p_2 > 1$, then let $\mathbb{P}[\mathbf{a}(e) > 0] = p_1$ for the horizonal bonds and $\mathbb{P}[\mathbf{a}(e) > 0] = p_2$ for the vertical bonds. We believe that our algorithm also works in these models by re-establishing all the quantitative homogenization theory from [19, 83] and repeating all the analysis in this paper. Thus, to state the algorithm more generally, we choose to use the notation of $\bar{\mathbf{a}}$ as a matrix in Theorem 3.1.1 and in the rest of the paragraph, especially Section 3.C.

3.1.2 Previous work

The homogenization theory was first developed for elliptic or parabolic equations with periodic coefficients, and then generalised to the case of random stationary coefficients. There exist many classical references such as [47, 161, 219, 145, 6]. Quantitative results in stochastic homogenization took a long time to emerge. The first partial results result were obtained by Yurinskii [225]. Recently, thanks to the work of Gloria, Neukamm and Otto [123, 124, 121, 125, 122], and Armstrong, Kuusi, Mourrat and Smart [30, 23, 31, 24], we understand better the typical size of the fundamental quantities in the stochastic homogenization of uniformly elliptic equations, which provides us with the possibility to analyze the performance of numerical algorithms in this context.

The homogenization of environments that do not satisfy a uniform ellipticity condition also drew attention. In [226], Zhikov and Piatnitski establish many results qualitatively and explain how to formulate the effective equation on various types of degenerate stationary environments. In [164], Lamacz, Neukamm and Otto obtain a bound of correctors on a simplified percolation model by imposing all the bonds in the first coordinate direction to be open. In [45], the Liouville regularity problem in a general context of random graphs is studied by Benjamini, Duminil-Copin, Kozma, and Yadin using the entropy method, and its complete description on infinite cluster of Bernoulli percolation is given by Armstrong and Dario in [19]. Dario also gives the moment estimate of the correctors of the same model in [83].

Homogenization has a natural probabilistic interpertation in terms of *random walks in random environment*, as a generalised central limit theorem. One fundamental work in this context is the paper [153] by Kipnis and Varadhan, where the case of general reversible Markov chains is studied. The case of random walks on the supercritical percolation cluster attracted particular interest, and the *quenched* central limit theorem was obtained at first by Sidoravicius and Sznitman in [211] for dimension $d \geqslant 4$, then generalized by Berger, Biskup, Mathieu and Piatnitski in [49, 180] for any dimension $d \geqslant 2$. We also refer to [53, 157, 162] for overviews of this line of research.

Finally, concerning the construction of efficient numerical methods, our algorithm is inspired by the one introduced in [22] by Armstrong, Hannukainen, Kuusi and Mourrat, which is designed to treat the same question in a uniform ellipticity context, and also [133] where a uniform estimate is obtained. Besides the fact that the problem we consider here is not uniformly elliptic, we stress that a fundamental issue we need to address relates to the fact that the *geometry* of the domain itself must be modified as we move from fine to coarse scales. Indeed, the fine scales must be resolved on the original, highly perforated domain, while the coarse scales are resolved in a homogeneous medium in which the wholes have been "filled up". As far as I know, this is the first work proposing a practical and rigorous method for the numerical approximation of elliptic problems posed in rapidly oscillating perforated domains. For the homogenized coefficient $\bar{\mathbf{a}}$, we have many excellent works like [119, 102, 184, 107, 138], which can be adapted naturally on the percolation setting. Alternative numerical methods for computing the solution of elliptic problems in non-perforated domains have been studied extensively; we refer in particular to [42, 36, 101, 128, 197, 178, 159, 196], as well as to [129, 154, 103, 104] where the concept of homogenization is used explicitly.

3.1.3 Ideas of the proof and main contributions

In this part, we introduce some key concepts underlying the analysis of the algorithm and the proof of Theorem 3.1.1. We also present our main contributions, including the mask operator and some other results like estimates on the flux and a quantitative version of the two-scale expansion on the cluster of percolation, which are of independent interest. Some notations are explained quickly in the statement and their rigorous definitions will be given in Section 3.2 or in the later part when they are used.

Main strategy

The main strategy of the algorithm is very similar to an algorithm proposed in the previous work [22, 133] where we study the classical Dirichlet problem in $\mathbb{R}^d$ setting with symmetric $\mathbb{R}^{d\times d}$-valued coefficient matrix $\mathbf{a}$, which is random, stationary, of finite range correlation and satisfies the uniform ellipticity condition. We recall the idea in the previous work with a little abuse of notation that $\square_m$ stands $\left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d$ in this paragraph: to solve a divergence-form equation $-\nabla \cdot \mathbf{a}\nabla u = f$ in $\square_m$ with boundary condition g, we propose to compute $(u_1, \bar{u}, u_2)$ with null Dirichlet boundary condition solving

$$\begin{cases} (\lambda^2 - \nabla \cdot \mathbf{a}\nabla)u_1 & = f + \nabla \cdot \mathbf{a}\nabla u_0 & \text{in } \square_m, \\ -\nabla \cdot \bar{\mathbf{a}}\nabla \bar{u} & = \lambda^2 u_1 & \text{in } \square_m, \\ (\lambda^2 - \nabla \cdot \mathbf{a}\nabla)u_2 & = (\lambda^2 - \nabla \cdot \bar{\mathbf{a}}\nabla)\bar{u} & \text{in } \square_m. \end{cases} \tag{3.7}$$

In [22, 133] we proved that $\hat{u} := u_0 + u_1 + u_2$ satisfies

$$\|\hat{u} - u\|_{H^1(\square_m)} \leqslant \mathcal{Z} \, \|u_0 - u\|_{H^1(\square_m)},$$

with a random factor $\mathcal{Z}$ of size $\mathcal{Z} \leqslant \mathcal{O}_s\left(C(\Lambda, s, d)\ell(\lambda)^{\frac{1}{2}}\lambda^{\frac{1}{2}}m^{\frac{1}{s}}\right)$ for any $s \in (0,2)$ and independent of u, u_0, f, g.

The main ingredient in the proof is the two-scale expansion theorem: for $v, \bar{v}$ with the same boundary condition and satisfying

$$(\mu^2 - \nabla \cdot \mathbf{a}\nabla)v = (\mu^2 - \nabla \cdot \bar{\mathbf{a}}\nabla)\bar{v} \qquad \text{in } \square_m, \tag{3.8}$$

one can use $\bar{v} + \sum_{k=1}^d (\partial_{x_k}\bar{v})\phi_{e_k}$ to approximate v in H^1. Here $\{e_k\}_{1 \leqslant k \leqslant d}$ stands for the canonical basis in $\mathbb{R}^d$, and ϕ_{e_k} is *the first order corrector* associated with the direction e_k. In our algorithm eq. (3.7), combining the first equation, the second equation of eq. (3.7) and $-\nabla \cdot \mathbf{a}\nabla u = f$, we can obtain that

$$-\nabla \cdot \bar{\mathbf{a}}\nabla \bar{u} = -\nabla \cdot \mathbf{a}\nabla(u - u_0 - u_1) \qquad \text{in } \square_m,$$

which is an equation of type eq. (3.8) with $\mu = 0$. Moreover, the third equation in eq. (3.7) also follows the form of eq. (3.8), this time with $\mu = \lambda$. Thus, we have

$$(u - u_0 - u_1) \simeq w := \bar{u} + \sum_{k=1}^d (\partial_{x_k}\bar{u})\phi_{e_k} \simeq u_2,$$

up to a small error, so we can estimate $|\hat{u} - u|$ by studying

$$|\hat{u} - u| = |u - (u_0 + u_1 + u_2)| \leqslant |(u - u_0 - u_1) - w| + |w - u_2|.$$

In [22] the error in the two-scale expansion theorem is made quantitative, and in [133] we refine this bound so that the contraction bound is uniform over the relevant data (most importantly: the bound is uniform over u_0, which guarantees that the algorithm can indeed be iterated).

Mask operator trick

In order to figure out the generalization of the algorithm on clusters, we recall the interpretation from multigrid method for eq. (3.7): the first equation in eq. (3.7) is the scheme in fine grid, which runs several rounds thanks to the regularization. The second step of eq. (3.7) is a coarse grid, where we use the homogenized matrix as a coarse operator. The third equation in eq. (3.7) is a post-treatment to project the error in the coarse grid back to the fine grid. However, in the $\mathbb{R}^d$ setting, the projection step is natural, but we should treat it more carefully in percolation setting, since the fine grid is defined on $\mathscr{C}_*(\square_m)$, which is random and depends on the realization of $\mathbf{a}$, while the coarse grid is defined on $\square_m$; see Figure 3.1 for an illustration.

To resolve the problem of projection, we use an idea called mask operator, which is defined as

$$\lambda_{\mathscr{C},m}(x) := \begin{cases} \lambda & \text{if } x \in \mathscr{C}_*(\square_m), \\ 0 & \text{otherwise.} \end{cases} \qquad \mathbf{a}_{\mathscr{C},m}(x,y) := \begin{cases} \mathbf{a}(x,y) & \text{if } x,y \in \mathscr{C}_*(\square_m), \\ 0 & \text{otherwise.} \end{cases}$$

We remark that some simliar idea also appears in the early work [226], where they call it *singular random measure* in the degenerate ergodic enviornment. In our algorithm, the mask operator is already implicitly used in the third equation of eq. (3.5), but the following nice observation Proposition 3.1.1 shows all its power: it allows us to treat the problem on percolation as if it is on $\mathbb{Z}^d$.

Proposition 3.1.1 (Arbitary extension)**.** *After an arbitrary extension of the function u_0, u_1, u_2 defined in eq. (3.5) on $\operatorname{int}(\square_m)\backslash\mathscr{C}_*(\square_m)$, the functions $u_1, \bar{u}, u_2$ also satisfy*

$$\begin{cases} (\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m}\nabla)u_1 & = f_{\mathscr{C},m} + \nabla \cdot \mathbf{a}_{\mathscr{C},m}\nabla u_0 & \text{in } \operatorname{int}(\square_m), \\ -\nabla \cdot \bar{\mathbf{a}}\nabla\bar{u} & = \lambda^2_{\mathscr{C},m}u_1 & \text{in } \operatorname{int}(\square_m), \\ (\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m}\nabla)u_2 & = (\lambda^2_{\mathscr{C},m} - \nabla \cdot \bar{\mathbf{a}}\nabla)\bar{u} & \text{in } \operatorname{int}(\square_m). \end{cases} \tag{3.9}$$

Proof. In the first equation of eq. (3.9) the left hand side can be rewritten as

$$(\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m}\nabla)u_1(x) = \lambda^2_{\mathscr{C},m}(x)u_1(x) + \sum_{y\sim x}(\mathbf{a}_{\mathscr{C},m}(x,y))(u_1(x) - u_1(y)),$$

while the right hand side equals

$$f_{\mathscr{C},m}(x) + \nabla \cdot \mathbf{a}_{\mathscr{C},m}\nabla u_0(x) = f_{\mathscr{C},m}(x) + \sum_{y\sim x}(\mathbf{a}_{\mathscr{C},m}(x,y))(u_0(y) - u_0(x)).$$

If $x \in \mathscr{C}_*(\square_m)\backslash\partial\square_m$ the left hand side and the right hand side both equal to the first equation in eq. (3.5), so the equation is established. If $x \in \operatorname{int}(\square_m)\backslash\mathscr{C}_*(\square_m)$, no matter what values u_1, u_0 takes on the extension, the factors and function $f_{\mathscr{C},m}(x) = \lambda_{\mathscr{C},m}(x) = \mathbf{a}_{\mathscr{C},m}(x,y) = 0$ make both left hand side and right hand side 0.

In the second equation, on the right hand side $\lambda^2_{\mathscr{C},m}u_1$ coincides with that in eq. (3.5) so the equation is also established.

The third equation is valid, if $x \in \mathscr{C}_*(\square_m)\backslash\partial\square_m$ for the similar reason as described in the first equation. If $x \in \operatorname{int}(\square_m)\backslash\mathscr{C}_*(\square_m)$, the left hand side equals 0 since all the factors and conductance are 0. The right hand side is also 0 thanks to a simple manipulation using the second equation of eq. (3.9)

$$(\lambda^2_{\mathscr{C},m} - \nabla \cdot \bar{\mathbf{a}}\nabla)\bar{u}(x) = \lambda^2_{\mathscr{C},m}(x)(\bar{u}(x) + u_1(x)) = 0,$$

and this finishes the proof. □

The same idea also works for u defined in eq. (3.1), which can also be defined as the solution

$$\begin{cases} -\nabla \cdot \mathbf{a}_{\mathscr{C},m}\nabla u = f & \text{in } \operatorname{int}(\square_m), \\ u = g & \text{on } \mathscr{C}_*(\square_m) \cap \partial\square_m, \end{cases} \tag{3.10}$$

with an arbitrary extension outside $\mathscr{C}_*(\square_m)$.

Two-scale expansion on $\mathscr{C}_\infty$

Once we obtain the description of the algorithm eq. (3.9), we can repeat the argument for eq. (3.7) to explore the two-scale convergence theorem, which should define its left hand side on the cluster $\mathscr{C}_*(\square_m)$ and its right hand side on the homogenized geometry $\square_m$,

$$(\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m}\nabla)v = (\lambda^2_{\mathscr{C},m} - \nabla \cdot \bar{\mathbf{a}}\nabla)\bar{v} \qquad \text{in } \mathrm{int}(\square_m), \tag{3.11}$$

and we hope to use a modified two-scale expansion

$$w := \bar{v} + \sum_{k=1}^{d} (\Upsilon \mathcal{D}_{e_k}\bar{v})\phi^{(\lambda)}_{e_k}, \tag{3.12}$$

to approximate v. Here $\mathcal{D}_{e_k}\bar{v}(x) := \bar{v}(x+e_k) - \bar{v}(x)$ and Υ is a cut-off function supported in $\square_m$, constant 1 in the interior and decreases to 0 linearly near the boundary defined as

$$\Upsilon := \mathbf{1}_{\{\square_m\}} \wedge \left(\frac{\mathrm{dist}(\cdot, \partial\square_m) - \ell(\lambda)}{\ell(\lambda)}\right)_+, \tag{3.13}$$

so the function Υ can help reduce the boundary layer effect of the two-scale expansion. The modified corrector $\{\phi^{(\lambda)}_{e_k}\}_{1\leqslant k\leqslant d}$ is defined as

$$\phi^{(\lambda)}_{e_k} := \phi_{e_k} - [\phi_{e_k}]^{\eta}_{\mathcal{P}} \star \Phi_{\lambda^{-1}}, \tag{3.14}$$

where $\Phi_{\lambda^{-1}}$ is a heat kernel of scale λ^{-1}, i.e. $\Phi_{\lambda^{-1}}(x) := \frac{1}{(4\pi\lambda^{-2})^{d/2}}\exp\left(-\frac{x^2}{4\lambda^{-2}}\right)$ and $[\phi_{e_k}]^{\eta}_{\mathcal{P}}$ is a coarsened version of ϕ_{e_k}, whose proper definition will be given in Definition 3.2.2. Although the corrector is only well-defined up to a constant, notice that eq. (3.14) is well-defined. Notice also that by (3.11), the function $\bar{v}$ is discrete-harmonic outside of $\mathscr{C}_*(\square_m)$.

In Section 3.4, we will prove the following quantitative two-scale expansion theorem as a main tool to prove the contraction estimate (Theorem 3.1.1). We remark here that we also add a technical condition $\square_m \in \mathcal{P}_*$, which is defined in Definition 3.2.6 and it comes from the partition of the cluster $\mathscr{C}_\infty$. It is stronger than "$\square_m$ is good", and means that "the cluster $\mathscr{C}_*(\square_m)$ is indeed a subset of $\mathscr{C}_\infty$".

Theorem 3.1.2 (Two-scale expansion on percolation)**.** *There exist two positive constants $s := s(d, \mathfrak{p}, \Lambda)$, $C := C(d, \mathfrak{p}, \Lambda, s)$, and for every integer $m > 1$ such that $\square_m \in \mathcal{P}_*$ and every $\lambda \in \left(\frac{1}{3^m}, \frac{1}{2}\right)$, there exists a random variable $\widetilde{\mathcal{Z}}$ controlled by*

$$\widetilde{\mathcal{Z}} \leqslant \mathcal{O}_s\left(C(d, \mathfrak{p}, \Lambda, s)\ell(\lambda)m^{\frac{1}{s}+d}\right),$$

such that the following is valid: for any $\mu \in [0, \lambda]$ and any $v, \bar{v} \in C_0(\square_m)$ satisfying

$$(\mu^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m}\nabla)v = (\mu^2_{\mathscr{C},m} - \nabla \cdot \bar{\mathbf{a}}\nabla)\bar{v} \qquad \textit{in } \mathrm{int}(\square_m), \tag{3.15}$$

defining a two-scale expansion $w := \bar{v} + \sum_{k=1}^{d}(\Upsilon\mathcal{D}_{e_k}\bar{v})\phi^{(\lambda)}_{e_k}$, we have

$$\begin{aligned} \left\|\nabla(w - v)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} \leqslant \widetilde{\mathcal{Z}}\Big(\left(3^{-\frac{m}{2}}\ell^{-\frac{1}{2}}(\lambda) + \mu\right)\|\nabla\bar{v}\|_{L^2(\square_m)} + \|\Delta\bar{v}\|_{L^2(\mathrm{int}(\square_m))} \\ + \|\nabla\bar{v}\|^{\frac{1}{2}}_{L^2(\square_m)}\|\Delta\bar{v}\|^{\frac{1}{2}}_{L^2(\mathrm{int}(\square_m))}\Big). \end{aligned} \tag{3.16}$$

Remark. We add some more explanations for the technical condition $\square_m \in \mathcal{P}_*$. In fact, when one implements the algorithm, the condition "$\square_m$ is a good cube" is very natural, but this condition is $\mathcal{F}_{\square_{m+1}}$-measurable thus only uses local information, and it does not necessarily imply that the maximal cluster $\mathscr{C}_*(\square_m)$ is part of the cluster $\mathscr{C}_\infty$. On the other hand, the corrector theory is established for the infinite cluster $\mathscr{C}_\infty$, so is two-scale expansion. We have to pay attention to this minor difference, although it is well-known that "the maximal cluster in a good cube is part of infinite cluster with high probability". One can remove the condition $\square_m \in \mathcal{P}_*$ in Theorem 3.1.2 with more technical analysis, but we choose to fill this gap in the part analysis of algorithm in Section 3.5 with a very simple inequality eq. (3.119).

Centered flux $\mathbf{g}_p$

Another topic studied in detail in this paper is an object called *centered flux* defined for each $p \in \mathbb{R}^d$ by

$$\mathbf{g}_p := \mathbf{a}_{\mathscr{C}}(\mathcal{D}\phi_p + p) - \bar{\mathbf{a}}p, \tag{3.17}$$

where $\mathbf{a}_{\mathscr{C}}(x,y) := \mathbf{a}(x,y)\mathbf{1}_{\{x,y \in \mathscr{C}_\infty\}}$. Together with the corrector ϕ_p, they are two quantities required for the proof of the two-scale convergence Theorem 3.1.2. Its physical interpretation is clear: we define $l_p(x) := p \cdot x$ and recall that the harmonic function can be seen as an electric potential. Then $(l_p + \phi_p)$ is the electric potential defined on $\mathscr{C}_\infty$ with conductance $\mathbf{a}$ associated to the direction p, while l_p is the one for the homogenized conductance $\bar{\mathbf{a}}$. We know that ϕ_p as the difference between the electric potentials is small compared to l_p, and heuristically, it should also be the case for the electric current. By Ohm's law, the two electric currents are defined by $\mathbf{a}_{\mathscr{C}}\nabla(\phi_p + l_p)$ and $\bar{\mathbf{a}}\nabla l_p$, so we expect indeed that $\mathbf{g}_p$ will be small. This is however only true in a weak sense, or equivalently, after a spatial convolution. In fact, we expect that $\mathbf{g}_p$ satisfies estimates that are very similar to those satisfied by $\nabla\phi_p$, and we will indeed prove an analogue of the result of [83, Proposition 3.1]. Here we use the notation $[\cdot]$ to represent the constant extension on every cube of the form $z + (-\frac{1}{2}, \frac{1}{2})^d$ for some $z \in \mathbb{Z}^d$.

Proposition 3.1.2 (Spatial average)**.** *There exist two positive constants $s(d,\mathfrak{p},\Lambda)$, $C(d,\mathfrak{p},\Lambda,s)$ such that for every $R \geqslant 1$ and every kernel $K_R : \mathbb{R}^d \to \mathbb{R}^+$ integrable and satisfying*

$$\exists C_{K,R} < \infty, \quad \forall x \in \mathbb{R}^d, \quad K_R(x) \leqslant \frac{C_{K,R}}{R^d\left(|\frac{x}{R}| \vee 1\right)^{\frac{d+1}{2}}}, \tag{3.18}$$

the quantity $(K_R \star [\mathbf{g}_p])(x)$ is well defined for every $x \in \mathbb{R}^d, p \in \mathbb{R}^d$ and it satisfies

$$|K_R \star [\mathbf{g}_p]|(x) \leqslant \mathcal{O}_s(C_{K,R} C |p| R^{-\frac{d}{2}}). \tag{3.19}$$

3.1.4 Organization of the paper

In Section 3.2, we define all the notations precisely and restate some important theorems in previous work. Section 3.3 is devoted to the study of the centered flux $\mathbf{g}_p$ and to the proof of Proposition 3.1.2. Section 3.4 gives the proof of the two-scale expansion on the cluster of percolation (Theorem 3.1.2). In Section 3.5, we use the two-scale expansion to analyze our algorithm. Finally, in Section 3.6, we present numerical experiments confirming the usefulness of the algorithm.

3.2 Preliminaires

This part defines rigorously all the notations used throughout this article. We also record some important results developed in previous work.

3.2.1 Notations $\mathcal{O}_s(1)$ and its operations

We recall the definition of $\mathcal{O}_s$

$$X \leqslant \mathcal{O}_s(\theta) \iff \mathbb{E}\left[\exp((\theta^{-1}X)_+^s)\right] \leqslant 2, \tag{3.20}$$

where $(\theta^{-1}X)_+$ means $\max\{\theta^{-1}X, 0\}$. One can use the Markov inequality to obtain that

$$X \leqslant \mathcal{O}_s(\theta) \Longrightarrow \forall x > 0, \mathbb{P}[X \geqslant \theta x] \leqslant 2\exp(-x^s).$$

For $\lambda \in \mathbb{R}^+, X \leqslant \mathcal{O}_s(\theta) \Longrightarrow \lambda X \leqslant \mathcal{O}_s(\lambda\theta)$. We list some results on the estimates of the random variables with respect of $\mathcal{O}_s$ in [25, Appendix A]. For the product of random variables, we have

$$|X| \leqslant \mathcal{O}_{s_1}(\theta_1), |Y| \leqslant \mathcal{O}_{s_2}(\theta_2) \Longrightarrow |XY| \leqslant \mathcal{O}_{\frac{s_1 s_2}{s_1+s_2}}(\theta_1\theta_2). \tag{3.21}$$

By choosing $Y = 1$, one can always use the estimate above to get an estimate for smaller exponent, i.e. for $0 < s' < s$, there exists a constant $C_{s'} < \infty$ such that

$$X \leqslant \mathcal{O}_s(\theta) \Longrightarrow X \leqslant \mathcal{O}_{s'}(C_{s'}\theta). \tag{3.22}$$

We have an estimate on the sum of a series of random variables: for a measure space $(E, \mathcal{S}, m)$ and $\{X(z)\}_{z\in E}$ a family of random variables, we have

$$\forall z \in E, X(z) \leqslant \mathcal{O}_s(\theta(z)) \Longrightarrow \int_E X(z)m(dz) \leqslant \mathcal{O}_s\left(C_s \int_E \theta(z)m(dz)\right), \tag{3.23}$$

where $0 < C_s < \infty$ is a constant defined by

$$C_s = \begin{cases} \left(\frac{1}{s\log 2}\right)^{\frac{1}{s}} & s < 1, \\ 1 & s \geqslant 1. \end{cases} \tag{3.24}$$

Finally, we can also obtain the estimate of the maximum of a finite number of random variables, which is proved in [133, Lemma 3.2] (see also Lemma 2.3.2): for all $N \geqslant 1$ and family of random variables $\{X_i\}_{1\leqslant i\leqslant N}$ satisfying that $X_i \leqslant \mathcal{O}_s(1)$, we have

$$\left(\max_{1\leqslant i\leqslant N} X_i\right) \leqslant \mathcal{O}_s\left(\left(\frac{\log(2N)}{\log(3/2)}\right)^{\frac{1}{s}}\right). \tag{3.25}$$

3.2.2 Discrete analysis

This part is devoted to introducing notations and some functional inequalities on graphs or on lattices. We take two systems of derivative in our setting: ∇ on graph and the finite difference $\mathcal{D}$ on $\mathbb{Z}^d$. The notation ∇ is more general, but it loses the sense of derivative with respect to a given direction, which is very natural in the system of $\mathcal{D}$.

Spaces and functions

For every $V \subseteq \mathbb{Z}^d$, we can construct two types of geometry $(V, E_d(V))$ and $(V, E_d^{\mathbf{a}}(V))$. The set of edges $E_d(V)$ inherited from $(\mathbb{Z}^d, E_d)$ and $E_d^{\mathbf{a}}(V)$ inherited from the open bonds of the percolation are defined as

$$E_d(V) := \{\{x,y\} | x, y \in V, x \sim y\}, \qquad E_d^{\mathbf{a}}(V) := \{\{x,y\} | x, y \in V, \mathbf{a}(x,y) \neq 0\}.$$

The *interior* of V with respect to $(V, E_d(V))$ and $(V, E_d^{\mathbf{a}}(V))$ are defined

$$\mathrm{int}(V) := \{x \in V | y \sim x \Longrightarrow y \in V\}, \qquad \mathrm{int}_{\mathbf{a}}(V) := \{x \in V | y \sim x, \mathbf{a}(x,y) \neq 0 \Longrightarrow y \in V\},$$

and the *boundaries* are defined as $\partial(V) := V \backslash \mathrm{int}(V)$ and $\partial_{\mathbf{a}}(V) := V \backslash \mathrm{int}_{\mathbf{a}}(V)$. For any $x, y \in \mathbb{Z}^d$, we say $x \overset{\mathbf{a}}{\leftrightarrow} y$ if there exists an open path connecting x and y.

We denote by $\overrightarrow{E_d}$ *the oriented bonds* of $(\mathbb{Z}^d, E_d)$, i.e. $\overrightarrow{E_d} := \{(x,y) \in \mathbb{Z}^d \times \mathbb{Z}^d : |x-y| = 1\}$, and for any $E \subseteq E_d$, we can associate it to a natural oriented bonds set $\overrightarrow{E}$. An *(anti-symmetric) vector field* $\overrightarrow{F}$ on $\overrightarrow{E_d}$ is a function $\overrightarrow{F} : \overrightarrow{E_d} \to \mathbb{R}$ such that $\overrightarrow{F}(x,y) = -\overrightarrow{F}(y,x)$. Sometimes we also write $\overrightarrow{F}(e)$ for $e = \{x,y\} \in E_d$ to give its value with an arbitrary orientation for e, in the case it is well defined (for example $|\overrightarrow{F}|(e)$). The *discrete divergence* of $\overrightarrow{F}$ is defined as $\nabla \cdot \overrightarrow{F} : \mathbb{Z}^d \to \mathbb{R}$

$$\forall x \in \mathbb{Z}^d, \quad \nabla \cdot \overrightarrow{F}(x) := \sum_{y \sim x} \overrightarrow{F}(x,y).$$

For any $u : \mathbb{Z}^d \to \mathbb{R}$, we define the discrete derivative $\nabla u : \overrightarrow{E_d} \to \mathbb{R}$ as a vector field

$$\forall (x,y) \in \overrightarrow{E_d}, \quad \nabla u(x,y) := u(y) - u(x),$$

and $\mathbf{a}\nabla u, \nabla u \mathbf{1}_{\{\mathbf{a} \neq 0\}}$ are vector fields $\overrightarrow{E_d} \to \mathbb{R}$ defined by

$$\mathbf{a}\nabla u(x,y) := \mathbf{a}(x,y)\nabla u(x,y), \qquad \nabla u \mathbf{1}_{\{\mathbf{a} \neq 0\}}(x,y) := \nabla u(x,y) \mathbf{1}_{\{\mathbf{a}(x,y) \neq 0\}}.$$

Then, the $\mathbf{a}$-Laplacian operator $-\nabla \cdot \mathbf{a} \nabla$ is well defined and we have

$$-\nabla \cdot \mathbf{a}\nabla u(x) := \sum_{y \sim x} \mathbf{a}(x,y)(u(x) - u(y)).$$

Finite difference derivative

We start by introducing the notation of translation: let B be a Banach space, then for any $h \in \mathbb{Z}^d$ and $u : \mathbb{Z}^d \to B$ a B-valued function, we define T_h as an operator

$$\forall x \in \mathbb{Z}^d, \quad (T_h u)(x) = u(x+h).$$

We also define the operator $\mathcal{D}_h$ and its conjugate operator $\mathcal{D}_h^*$ for any $u : \mathbb{Z}^d \to \mathbb{R}$,

$$\mathcal{D}_h u := T_h u - u, \qquad \mathcal{D}_h^* u := T_{-h} u - u.$$

It is easy to check $\mathcal{D}_h^* = -T_{-h}(\mathcal{D}_h u)$ and for two functions $f, g : \mathbb{Z}^d \to \mathbb{R}$, we have

$$\mathcal{D}_h(fg) = (\mathcal{D}_h f)g + (T_h f)(\mathcal{D}_h g). \tag{3.26}$$

In this system, we also define *vector field* $\widetilde{F} : \mathbb{Z}^d \to \mathbb{R}^d, \widetilde{F}(x) = (\widetilde{F}_1(x), \widetilde{F}_2(x) \cdots \widetilde{F}_d(x))$ and this can be distinguished with the one defined on E_d by the context. We use $(e_1, e_2 \cdots e_d)$ to represent the d canonical directions in $\mathbb{Z}^d$, and a discrete gradient $\mathcal{D}u : \mathbb{Z}^d \to \mathbb{R}^d$ is a vector field

$$\mathcal{D}u(x) := (\mathcal{D}_{e_1}u(x), \mathcal{D}_{e_2}u(x) \cdots \mathcal{D}_{e_d}u(x)).$$

Then the finite difference divergence operator is defined as the conjugate operator of $\mathcal{D}$

$$\mathcal{D}^* \cdot \widetilde{F} := \sum_{j=1}^{d} \mathcal{D}^*_{e_j} \widetilde{F}_j.$$

As convention, we use the notation $\mathbf{a}\mathcal{D}_{e_j}u$ and $\mathbf{a}\mathcal{D}u$ to represent

$$\mathbf{a}\mathcal{D}_{e_j}u(x) := \mathbf{a}(x, x+e_j)\mathcal{D}_{e_j}u(x), \qquad \mathbf{a}\mathcal{D}u := (\mathbf{a}\mathcal{D}_{e_1}u, \cdots \mathbf{a}\mathcal{D}_{e_d}u),$$

and $\mathbf{1}_{\{\mathbf{a}\neq 0\}}\mathcal{D}_{e_j}u, \mathbf{1}_{\{\mathbf{a}\neq 0\}}\mathcal{D}u$, to represent

$$\mathbf{1}_{\{\mathbf{a}\neq 0\}}\mathcal{D}_{e_j}u(x) := \mathbf{1}_{\{\mathbf{a}(x,x+e_j)\neq 0\}}\mathcal{D}_{e_j}u(x), \qquad \mathbf{1}_{\{\mathbf{a}\neq 0\}}\mathcal{D}u := \left(\mathbf{1}_{\{\mathbf{a}\neq 0\}}\mathcal{D}_{e_1}u, \cdots \mathbf{1}_{\{\mathbf{a}\neq 0\}}\mathcal{D}_{e_d}u\right).$$

Thus the $\mathbf{a}$-Laplacian operator $-\nabla \cdot \mathbf{a}\nabla$ can also be expressed by the finite difference $\mathcal{D}^* \cdot \mathbf{a}\mathcal{D}$. We can prove it by a simple calculation that

$$-\nabla \cdot \mathbf{a}\nabla u = \mathcal{D}^* \cdot \mathbf{a}\mathcal{D}u. \tag{3.27}$$

Inner product and norm

For $V \subseteq \mathbb{Z}^d$ and $E \subseteq E_d$, we define inner product $\langle \cdot, \cdot \rangle_V$ for any function $u, v : V \to \mathbb{R}$ and $\langle \cdot, \cdot \rangle_E$ for any vector field $\vec{F}, \vec{G} : \vec{E} \to \mathbb{R}$

$$\langle u, v \rangle_V := \sum_{x \in V} u(x)v(x), \qquad \left\langle \vec{F}, \vec{G} \right\rangle_E := \sum_{\{x,y\} \in E} \vec{F}(x,y)\vec{G}(x,y),$$

and this defines a norm $\|u\|_{L^2(V)} := \sqrt{\langle u, u \rangle_V}$ and $\left\|\vec{F}\right\|_{L^2(E)} := \sqrt{\left\langle \vec{F}, \vec{F} \right\rangle_E}$. We also abuse a little the notation to define $\langle \cdot, \cdot \rangle_V$ for vector field

$$\left\langle \vec{F}, \vec{G} \right\rangle_V := \left\langle \vec{F}, \vec{G} \right\rangle_{E_d(V)} = \sum_{\{x,y\} \in E_d(V)} \vec{F}(x,y)\vec{G}(x,y) = \frac{1}{2} \sum_{x,y \in V, y \sim x} \vec{F}(x,y)\vec{G}(x,y).$$

We use the notation $\langle \cdot, \cdot \rangle_{E^{\mathbf{a}}_d(V)}$ to represent the inner product of the vector field on $(V, E^{\mathbf{a}}_d(V))$. For two vector fields $\widetilde{F}, \widetilde{G} : V \to \mathbb{R}^d$, the inner product is defined as

$$\left\langle \widetilde{F}, \widetilde{G} \right\rangle_V := \sum_{x \in V} \sum_{j=1}^{d} \widetilde{F}_j(x)\widetilde{G}_j(x),$$

and similarly $\left\|\widetilde{F}\right\|_{L^2(V)} = \sqrt{\left\langle \widetilde{F}, \widetilde{F} \right\rangle_E}$ also defines a norm.

To define a general $L^p(V)(p \geqslant 1)$ norm for vector fields, we have to introduce its modules. For any $\vec{F} : \vec{E_d} \to \mathbb{R}$ or $\widetilde{F} : \mathbb{Z}^d \to \mathbb{R}^d$, we write

$$|\vec{F}|(x) := \left(\frac{1}{2} \sum_{y \sim x} \vec{F}^2(x,y) \right)^{\frac{1}{2}}, \qquad |\widetilde{F}|(x) := \left(\sum_{j=1}^{d} \widetilde{F}^2_j(x) \right)^{\frac{1}{2}}.$$

Then for f (a function, an $\mathbb{R}^d$-valued vector field or a vector field on $\overrightarrow{E_d}$)

$$\|f\|_{L^p(V)}^p := \left(\sum_{x\in V} |f|^p(x)\right)^{\frac{1}{p}}, \qquad \|f\|_{\underline{L}^p(V)}^p := \left(\frac{1}{|V|}\sum_{x\in V} |f|^p(x)\right)^{\frac{1}{p}}.$$

We recall $C_0(V)$ the space of functions supported on V with null boundary condition. Then one can deduce integration by part formula: for any function $v \in C_0(V)$, $\overrightarrow{F} : \overrightarrow{E_d}(V) \to \mathbb{R}$ and $\widetilde{F} : \mathbb{Z}^d \to \mathbb{R}^d$, one can check

$$\left\langle v, -\nabla \cdot \overrightarrow{F}\right\rangle_{\mathrm{int}(V)} = \left\langle \nabla v, \overrightarrow{F}\right\rangle_V, \qquad \left\langle v, \mathcal{D}^* \cdot \widetilde{F}\right\rangle_{\mathrm{int}(V)} = \left\langle \mathcal{D}v, \widetilde{F}\right\rangle_V. \tag{3.28}$$

Some functional inequalities

Here are some discrete functional inequalities used throughout the article.

Lemma 3.2.1 (Discrete functional inequality)**.** *1. (A naive estimate) Given a $V \subseteq \mathbb{Z}^d$ and for a function $v : V \to \mathbb{R}$, we have*

$$\langle \nabla v, \nabla v\rangle_V \leqslant 2d \langle v, v\rangle_V. \tag{3.29}$$

2. (Poincaré's inequality) For every $v \in C_0(\square_m)$, we have

$$\|v\|_{L^2(\square_m)} \leqslant C(d)3^m \|\nabla v\|_{L^2(\square_m)}. \tag{3.30}$$

3. (H^2 interior regularity for discrete harmonic function) Given two functions $v, f \in C_0(\square_m)$ satisfying the discrete elliptic equation ($\Delta v = \nabla \cdot \nabla v$)

$$-\Delta v = f, \qquad \textit{in } \mathrm{int}(\square_m), \tag{3.31}$$

then we have an interior estimate

$$\|\mathcal{D}^*\mathcal{D}v\|_{L^2(\mathrm{int}(\square_m))}^2 := \sum_{i,j=1}^d \left\|\mathcal{D}_{e_i}^*\mathcal{D}_{e_j} v\right\|_{L^2(\mathrm{int}(\square_m))}^2 \leqslant d\,\|f\|_{L^2(\mathrm{int}(\square_m))}^2. \tag{3.32}$$

4. (Trace inequality) For every $u : \square_m \to \mathbb{R}$ and $0 \leqslant K \leqslant \frac{3^m}{4}$, we have the following inequality

$$\left\|u\mathbf{1}_{\{\mathrm{dist}(\cdot,\partial\square_m)\leqslant K\}}\right\|_{L^2(\square_m)}^2 \leqslant C(d)(K+1)\left(3^{-m}\|u\|_{L^2(\square_m)}^2 + \|u\|_{L^2(\square_m)}\|\nabla u\|_{L^2(\square_m)}\right). \tag{3.33}$$

The inequality (3.29) is very elementary, and the proof of eq. (3.30) is similar to the standard case, so we skip their proofs. The inequality (3.32) is also relatively standard, but involves a careful calculation. The argument for eq. (3.33) is more combinatorial and non-trivial. We provide their proofs in Section 3.A.

3.2.3 Partition of good cubes

One difficulty to treat the function defined on the percolation clusters comes from its random geometry. To overcome this problem, [19] introduces a Calderón-Zygmund type partition of good cubes, and we recall it here.

We denote by $\mathcal{T}$ the *triadic cube* and $\square_m(z)$ is defined by

$$\square_m(z) := \mathbb{Z}^d \bigcap \left(z + \left(-\frac{1}{2}3^m, \frac{1}{2}3^m\right)\right), z \in 3^m\mathbb{Z}^d, m \in \mathbb{N},$$

where *center* and *size* of the cube above is respectively z and 3^m, and we use the notation $\operatorname{size}(\cdot)$ to refer to the size, i.e. $\operatorname{size}(\square_m(z)) = 3^m$. In this paper, without further mention, we use the word "cube" for short of "triadic cube" and $\square_m$ for short of $\square_m(0)$. The collection of all the cubes of size 3^n is defined by $\mathcal{T}_n$, i.e. $\mathcal{T}_n := \{z + \square_n : z \in 3^n\mathbb{Z}^d\}$. Then we have naturally $\mathcal{T} = \bigcup_{n\in\mathbb{N}} \mathcal{T}_n$. Every cube of size 3^m can be divided into a partition of $3^{(m-n)}$ cubes in $\mathcal{T}_n$, and two cubes in $\mathcal{T}$ can be either disjoint or included one by the other. For each $\square \in \mathcal{T}$, the *predecessor* of $\square$ is the unique triadic cube $\widetilde{\square} \in \mathcal{T}$ satisfying $\square \subseteq \widetilde{\square}$, and $\frac{\operatorname{size}(\widetilde{\square})}{\operatorname{size}(\square)} = 3$, and reciprocally, we say $\square$ is a *successor* of $\widetilde{\square}$.

The distance between two points $x, y \in \mathbb{R}^d$ is defined to be $\operatorname{dist}(x, y) = \max_{i\in\{1,2\cdots d\}} |x_i - y_i|$ and the distance for $U, V \subseteq \mathbb{Z}^d$ is $\operatorname{dist}(U, V) = \inf_{x\in U, y\in V} \operatorname{dist}(x, y)$. In particular, two $\square, \square'$ are neighbors if and only if $\operatorname{dist}(\square, \square') = 1$ and one is included in the other if and only if $\operatorname{dist}(\square, \square') = 0$.

General setting

We state at first the general setting of partition of good cubes.

Proposition 3.2.1 (Proposition 2.1 of [19]). *Let $\mathcal{G} \subseteq \mathcal{T}$ a sub-collection of triadic cubes satisfying the following: for every $\square = z + \square_n \in \mathcal{T}$, $\{\square \notin \mathcal{G}\} \in \mathcal{F}(z + \square_{n+1})$, and there exist two finite positive constants K, s*

$$\sup_{z\in 3^n\mathbb{Z}^d} \mathbb{P}[z + \square_n \notin \mathcal{G}] \leqslant K \exp(-K^{-1}3^{ns}).$$

Then, $\mathbb{P}$-almost surely there exists $\mathcal{S} \subseteq \mathcal{T}$ a partition of $\mathbb{Z}^d$ with the following properties:

1. *Cubes containing elements of $\mathcal{S}$ are good: for every $\square, \square' \in \mathcal{T}$, $\square \subseteq \square', \square \in \mathcal{S} \Longrightarrow \square' \in \mathcal{G}$.*

2. *Neighbors of elements of $\mathcal{S}$ are comparable: for every $\square, \square' \in \mathcal{S}$ such that $\operatorname{dist}(\square, \square') \leqslant 1$, we have $\frac{1}{3} \leqslant \frac{\operatorname{size}(\square)}{\operatorname{size}(\square')} \leqslant 3$.*

3. *Estimate for the coarseness: we use $\square_{\mathcal{S}}(x)$ to represent the unique element in $\mathcal{S}$ containing a point $x \in \mathbb{Z}^d$, then there exists a finite positive constant $C := C(s, K, d)$ such that, for every $x \in \mathbb{Z}^d$, $\operatorname{size}(\square_{\mathcal{S}}(x)) \leqslant \mathcal{O}_s(C)$.*

Case of well-connected cubes

The construction in Proposition 3.2.1 works for all collection of good cubes $\mathcal{G}$, here we give the concrete definition of good cubes we use in our context of percolation, as appearing in the work [200], [201] and [17] of Antal, Pisztora and Penrose. We remark that in Definition 3.2.1 and Definition 3.2.2 we use "cube" exceptionally for a general lattice cube, and we will highlight explicitly "triadic cube" when using it. The notation $\frac{3}{4}\overline{\square}$ indicates that we take the convex hull of the lattice cube, and then change its size by multiplying by $\frac{3}{4}$ while keeping the center fixed.

Definition 3.2.1 (Crossability and crossing cluster). We say that a cube $\square$ is *crossable* with respect to the open edges defined by $\mathbf{a}$ if each of the d pairs of opposite $(d-1)$-dimensional faces of $\square$ can be joined by an open path in $\square$. We say that a cluster $\mathscr{C} \subseteq \square$ is a *crossing cluster* for $\square$ if $\mathscr{C}$ intersects each of the $(d-1)$-dimensional faces of $\square$.

Definition 3.2.2 (Well-connected cube and good cube, Theorem 3.2 of [201]). We say that $\square \in \mathcal{T}$ is *well-connected* if there exists a crossing cluster $\mathscr{C}$ for $\square$ such that :

1. each cube $\square'$ with $\frac{1}{10}\operatorname{size}(\square) \leqslant \operatorname{size}(\square') \leqslant \frac{1}{2}\operatorname{size}(\square)$ and $\square' \cap \frac{3}{4}\overline{\square} \neq \varnothing$ is crossable.
2. every path $\gamma \subseteq \square'$ defined above with $\operatorname{diam}(\gamma) \geqslant \frac{1}{10}\operatorname{size}(\square)$ is connected to $\mathscr{C}$ within $\square'$.

We say that $\square \in \mathcal{T}$ is a *good cube* if $\operatorname{size}(\square) \geqslant 3$, $\square$ is connected and all his 3^d successors are well-connected. Otherwise, we say that $\square \in \mathcal{T}$ is a *bad cube*.

The following estimates makes the construction defined in Proposition 3.2.1 work.

Lemma 3.2.2 ((2.24) of [17]). *For each $\mathfrak{p} \in (\mathfrak{p}_c, 1]$, there exists a positive constant $C(d,\mathfrak{p})$ such that for every $n \in \mathbb{N}$,*

$$\mathbb{P}[\square_n \in \mathcal{G}] \geqslant 1 - C\exp(-C^{-1}3^n).$$

Definition 3.2.3 (Partition of good cubes in percolation context). We let $\mathcal{P} \subseteq \mathcal{T}$ be the partition $\mathcal{S}$ of $\mathbb{Z}^d$ obtained by applying Proposition 3.2.1 to the collection of good cubes defined in Definition 3.2.2

$$\mathcal{G} := \{\square \in \mathcal{T} : \square \text{ is good cube }\}.$$

A direct application of Lemma 3.2.2 and Proposition 3.2.1 gives us:

Corollary 3.2.1. *There exists a positive constant $C(d,\mathfrak{p})$, such that for every $z \in \mathbb{Z}^d$, we have the two estimates*

$$\operatorname{size}(\square_{\mathcal{P}}(z)) \leqslant \mathcal{O}_1(C), \qquad \mathbf{1}_{\{\operatorname{size}(\square_{\mathcal{P}}(z)) \geqslant n\}} \leqslant \mathcal{O}_1(C3^{-n}). \tag{3.34}$$

The maximal cluster is well defined on every good cube by Definition 3.2.2.

Definition 3.2.4 (Maximal cluster in good cubes). For every good cube $\square$, there exists a unique maximal crossing cluster in it, and we denote this cluster by $\mathscr{C}_*(\square)$.

Although $\mathscr{C}_*(\square)$ only uses local information, the next lemma shows that, for a $\square \in \mathcal{P}$ (stronger than $\square$ is good), its maximal cluster $\mathscr{C}_*(\square)$ must belong to the infinite cluster $\mathscr{C}_\infty$.

Lemma 3.2.3 (Lemma 2.8 of [19]). *Let $n, n' \in \mathbb{N}$ with $|n - n'| \leqslant 1$ and $z, z' \in 3^n\mathbb{Z}^d$ such that*

$$\operatorname{dist}(\square_n(z), \square_{n'}(z')) \leqslant 1.$$

Suppose also that $\square_n(z)$ and $\square_{n'}(z')$ are all good cubes, then there exists a cluster $\mathscr{C}$ such that

$$\mathscr{C}_*(\square_n(z)) \cup \mathscr{C}_*(\square_{n'}(z')) \subseteq \mathscr{C} \subseteq \square_n(z) \cup \square_{n'}(z').$$

This lemma helps us generalize the definition of maximal cluster in a general set $U \subseteq \mathbb{Z}^d$, the idea is to define the union of the partition cubes that cover U, and then find the maximal cluster in it.

Definition 3.2.5 (Maximal cluster in general set)**.** For a general set $U \subseteq \mathbb{Z}^d$, we define its *closure with respect to* $\mathcal{P}$ by

$$\mathrm{cl}_{\mathcal{P}}(U) := \bigcup_{z \in U} \square_{\mathcal{P}}(z), \tag{3.35}$$

and $\mathscr{C}_*(U)$ to be the cluster contained in $\mathrm{cl}_{\mathcal{P}}(U)$ which contains all the clusters of $\mathscr{C}_*(\square_{\mathcal{P}}(z))$, $z \in U$.

One can check easily that Lemma 3.2.3 makes the definition $\mathscr{C}_*(U)$ well-defined. However, we do not have necessarily $\mathscr{C}_*(U) = \bigcup_{z \in U} \mathscr{C}_*(\square_{\mathcal{P}}(z))$. We provide with a detailed discussion of this point in Section 3.B.

Since the cubes in $\mathcal{T}$ can be either included in one another or disjoint, if one cube $\square \in \mathcal{T}$ contains an element in $\mathcal{P}$, then it can be decomposed as the disjoint union of elements in $\mathcal{P}$ without enlarging the domain. Thus, we define:

Definition 3.2.6 (Minimal scale for partition)**.**

$$\mathcal{P}_* = \left\{ \square \in \mathcal{T} : \exists \square' \subseteq \square \text{ and } \square' \in \mathcal{P} \right\}. \tag{3.36}$$

The following observations are very useful and can be checked easily: for every $\square \in \mathcal{T}$, we have

$$\square \in \mathcal{P}_* \Longrightarrow \mathrm{cl}_{\mathcal{P}}(\square) = \square, \qquad \mathbf{1}_{\{\square \notin \mathcal{P}_*\}} \leqslant \mathbf{1}_{\{\mathrm{size}(\square_{\mathcal{P}}(z)) > \mathrm{size}(\square)\}} \leqslant \mathcal{O}_1(C(\mathrm{size}(\square))^{-1}). \tag{3.37}$$

Mask operator and coarsened function

To overcome the problem of the passage between the two geometries $(\mathbb{Z}^d, E_d)$ and $(\mathscr{C}_\infty, E_d^{\mathbf{a}})$, one useful technique is the mask operator.

Definition 3.2.7 (Mask operator and local mask operator)**.** For $f : \mathbb{Z}^d \to \mathbb{R}$ and $\mathbf{a} : E_d \to \mathbb{R}$, we define a *mask operator* $(\cdot)_{\mathscr{C}}$ to restrict their support on $\mathscr{C}_\infty$ and $E_d^{\mathbf{a}}(\mathscr{C}_\infty)$ respectively

$$f_{\mathscr{C}}(x) := \begin{cases} f(x) & x \in \mathscr{C}_\infty, \\ 0 & \text{otherwise.} \end{cases} \qquad \mathbf{a}_{\mathscr{C}}(x,y) := \begin{cases} \mathbf{a}(x,y) & x, y \in \mathscr{C}_\infty, \\ 0 & \text{otherwise.} \end{cases} \tag{3.38}$$

Moreover, we also define a *local mask operator* for $\square_m \in \mathcal{G}$ as

$$f_{\mathscr{C},m}(x) := \begin{cases} f(x) & x \in \mathscr{C}_*(\square_m), \\ 0 & \text{otherwise.} \end{cases} \qquad \mathbf{a}_{\mathscr{C},m}(x,y) := \begin{cases} \mathbf{a}(x,y) & x, y \in \mathscr{C}_*(\square_m), \\ 0 & \text{otherwise.} \end{cases} \tag{3.39}$$

Then we call $f_{\mathscr{C}} (f_{\mathscr{C},m}), \mathbf{a}_{\mathscr{C}} (\mathbf{a}_{\mathscr{C},m})$ *(local) masked function and (local) masked conductance.*

Reciprocally, for a function only defined on the clusters, sometimes we have to extend them to the whole space. We can apply the technique of coarsening the function defined on the percolation cluster.

Definition 3.2.8 (Coarsened function)**.** Given $\square \in \mathcal{P}$, we let $\bar{z}(\square)$ represent the vertex in $\mathscr{C}_*(\square)$ which is closest to its center. For a function $u : \mathscr{C}_\infty \to \mathbb{R}$, we define the coarsened function with respect to $\mathcal{P}$ to be $[u]_{\mathcal{P}} : \mathbb{Z}^d \to \mathbb{R}$ that

$$[u]_{\mathcal{P}}(x) := u(\bar{z}(\square_{\mathcal{P}}(x))).$$

We also use the notation $[\cdot]$ to mean doing constant extension on every cube, i.e. given $v : \mathbb{Z}^d \to \mathbb{R}$, we define $[v] : \mathbb{R}^d \to \mathbb{R}$ such that for every $z \in \mathbb{Z}^d$ and every $x \in z + \left[-\frac{1}{2}, \frac{1}{2}\right)^d$, $[v](x) := v(z)$.

The advantage of the coarsened function is that it allows to extend the support of function from $\mathscr{C}_\infty$ to the whole space, and constant in every cube by paying a small cost of errors.

Proposition 3.2.2 (Lemmas 3.2 and 3.3 of [19]). *For every $1 \leqslant s < \infty$, there exists a finite positive constant $C(s, d, \mathfrak{p})$, such that for every $\square \in \mathcal{P}_*$, $u : \mathscr{C}_\infty \to \mathbb{R}$, we have*

$$\sum_{x \in \mathscr{C}_*(\square)} |u(x) - [u]_{\mathcal{P}}(x)|^s \leqslant C^s \sum_{\{y,z\} \in E_d^{\mathbf{a}}(\mathscr{C}_*(\square))} \operatorname{size}(\square_{\mathcal{P}}(y))^{sd} |\nabla u|^s(y,z), \tag{3.40}$$

$$\sum_{\{x,y\} \in E_d(\square)} |\nabla [u]_{\mathcal{P}}(x,y)|^s \leqslant C^s \sum_{\{x,y\} \in E_d^{\mathbf{a}}(\mathscr{C}_*(\square))} \operatorname{size}(\square_{\mathcal{P}}(x))^{sd-1} |\nabla u|^s(x,y). \tag{3.41}$$

Remark. The main idea of coarsened function is to give function a constant value in every cube, but the value does not have to be of the one closest to the center. Following the same idea of proof of [19, Lemmas 3.2 and 3.3], one can prove that for $\square \in \mathcal{P}_*, u \in C_0(\square)$

$$[u]_{\mathcal{P},\square}(x) = \begin{cases} [u]_{\mathcal{P}}(x) & \operatorname{dist}(\square_{\mathcal{P}}(x), \partial \operatorname{cl}_{\mathcal{P}}(\square)) \geqslant 1, \\ 0 & \operatorname{dist}(\square_{\mathcal{P}}(x), \partial \operatorname{cl}_{\mathcal{P}}(\square)) = 0, \end{cases} \tag{3.42}$$

we have the same inequality as eq. (3.40) and eq. (3.41) by putting $[u]_{\mathcal{P},\square}$ in the place of $[u]_{\mathcal{P}}$.

3.2.4 Harmonic functions on the infinite cluster

We define $\mathcal{A}(U)$, the set of $\mathbf{a}$-harmonic functions on $U \subseteq \mathbb{Z}^d$, by

$$\mathcal{A}(U) := \{v : \mathscr{C}_\infty \to \mathbb{R} | -\nabla \cdot \mathbf{a}_{\mathscr{C}} \nabla v = 0, \forall x \in \operatorname{int}_{\mathbf{a}}(U)\},$$

and $\mathcal{A}(\mathscr{C}_\infty)$ the set $\mathbf{a}$-harmonic functions on $\mathscr{C}_\infty$. The $\mathbf{a}$-harmonic function $\mathcal{A}_k(\mathscr{C}_\infty)$ is the subspace of $\mathbf{a}$-harmonic functions which grows more slowly than a polynomial of degree $k+1$:

$$\mathcal{A}_k(\mathscr{C}_\infty) := \left\{ u \in \mathcal{A}(\mathscr{C}_\infty) \middle| \limsup_{R \to \infty} R^{-(k+1)} \|u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_R)} = 0 \right\}.$$

Similarly, we can define the spaces $\bar{\mathcal{A}}, \bar{\mathcal{A}}_k$ for harmonic functions on $\mathbb{R}^d$. It is well-known that the space $\bar{\mathcal{A}}_k$ is a finite-dimensional vector space of polynomials. A recent remarkable result about $\mathbf{a}$-harmonic functions on the infinite cluster of percolation conjectured in [45] and proved in [19] is that the space $\mathcal{A}_k(\mathscr{C}_\infty)$ also has this property, and in fact has the same dimension as $\bar{\mathcal{A}}_k$. Here we only recall the structure of $\mathcal{A}_1(\mathscr{C}_\infty)$: for every $\mathbf{a}$-harmonic functions $u \in \mathcal{A}_1(\mathscr{C}_\infty)$, there exists $c \in \mathbb{R}, p \in \mathbb{R}$ such that

$$\forall x \in \mathscr{C}_\infty, \quad u(x) = c + p \cdot x + \phi_p(x),$$

where the functions $\{\phi_p\}_{p \in \mathbb{R}^d}$ are called *the first order correctors*. The first order correctors have sublinear growth: there exists a positive exponent $\delta(d, \mathfrak{p}, \Lambda) < 1$ and a minimal scale $\mathcal{M} \leqslant \mathcal{O}_s(C(d, \mathfrak{p}, \Lambda))$ such that, for every $r \geqslant \mathcal{M}$ and $p \in \mathbb{R}^d$,

$$\|\phi_p\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r)} \leqslant C|p| r^{1-\delta}. \tag{3.43}$$

Combining eq. (3.43) and Cacciopolli's inequality [19, Lemma 3.5], for every $r \geqslant 2\mathcal{M}$, we have

$$\|\nabla(\phi_p + l_p)\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_{r/2})} \leqslant \frac{1}{r} \|\phi_p + l_p\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r)} \leqslant \frac{C}{r}(r^{1-\delta} + r)|p| \leqslant C|p|,$$

so it also implies that the estimate for the gradient of corrector that

$$\|\nabla\phi_p\|_{\underline{L}^2(\mathscr{C}_\infty\cap B_{r/2})} \leqslant C|p|. \tag{3.44}$$

The corrector plays an important role in the homogenization theory, and [83] gives a more precise description of these correctors. We recall that $\Phi_R(x) := \frac{1}{(4\pi R^2)^{d/2}}\exp\left(-\frac{x^2}{4R^2}\right)$, and $[\phi_p]^\eta_{\mathcal{P}} := [\phi_p]_{\mathcal{P}} \star \eta$ where $\eta \in C_0^\infty(B_1)$ is positive, and $\eta \equiv 1$ in $B_{\frac{1}{2}}$.

Proposition 3.2.3 (Local estimate and spatial average estimate, Proposition 3.1 of [83])**.** *There exist two finite positive constants $s := s(d,\mathfrak{p},\Lambda)$, $C := C(d,\mathfrak{p},\Lambda)$ such that for each $R \geqslant 1$ and each $p \in \mathbb{R}^d$,*

$$\forall x \in \mathbb{Z}^d, \quad |\nabla\phi_p \mathbf{1}_{\{\mathbf{a}\neq 0\}}|(x) \leqslant \mathcal{O}_s(C|p|), \tag{3.45}$$

$$\forall x \in \mathbb{R}^d, \quad \left|\nabla\left(\Phi_R \star [\phi_p]^\eta_{\mathcal{P}}\right)(x)\right| \leqslant \mathcal{O}_s(C|p|R^{-\frac{d}{2}}). \tag{3.46}$$

Proposition 3.2.4 (Theorem 1 and 2 of [83], L^q estimates on $\mathscr{C}_\infty$)**.** *There exist three finite positive constants $s := s(d,\mathfrak{p},\Lambda), k := k(d,\mathfrak{p},\Lambda)$ and $C := C(d,\mathfrak{p},\Lambda)$ such that for each $q \in [1,\infty)$, $R \geqslant 1$ and $p \in \mathbb{R}^d$,*

$$\left(R^{-d}\int_{\mathscr{C}_\infty\cap B_R} |\phi_p - (\phi_p)_{\mathscr{C}_\infty\cap B_R}|^q\right)^{\frac{1}{q}} \leqslant \begin{cases} \mathcal{O}_s(C|p|q^k\log^{\frac{1}{2}}(R)) & d=2, \\ \mathcal{O}_s(C|p|q^k) & d=3, \end{cases} \tag{3.47}$$

and for every $x,y \in \mathbb{Z}^d$ and $p \in \mathbb{R}^d$,

$$|\phi_p(x) - \phi_p(y)|\mathbf{1}_{\{x,y\in\mathscr{C}_\infty\}} \leqslant \begin{cases} \mathcal{O}_s(C|p|\log^{\frac{1}{2}}|x-y|) & d=2, \\ \mathcal{O}_s(C|p|) & d=3. \end{cases} \tag{3.48}$$

3.3 Centered flux on the cluster

In this part, we will study an object $\mathbf{g}_p : \mathbb{Z}^d \to \mathbb{R}^d$ called *centered flux* defined by

$$\mathbf{g}_p := \mathbf{a}_{\mathscr{C}}(\mathcal{D}\phi_p + p) - \bar{\mathbf{a}}p,$$

where $\mathbf{a}_{\mathscr{C}}$ is the masked conductance defined in eq. (3.38) and it is $\mathbf{a}$ restricted on the infinite cluster $\mathscr{C}_\infty$. Because $\mathbf{g}_p$ satisfies $\mathcal{D}^*\cdot\mathbf{g}_p = 0$ on $\mathbb{Z}^d$, following the spirit of Helmholtz-Hodge decomposition, in the later part of this section we will also study another object $\mathbf{S}_p : \mathbb{Z}^d \to \mathbb{R}^{d\times d}$ called *flux corrector* such that $\mathbf{g}_p = \mathcal{D}^*\cdot\mathbf{S}_p$ on $\mathbb{Z}^d$, in the sense $\mathbf{g}_{p,i} = \sum_{j=1}^d \mathcal{D}^*_{e_j}\mathbf{S}_{p,ij}$.

The quantities $\mathbf{g}_p$ and $\mathbf{S}_p$ are fundamental to the quantitative analysis of the two-scale expansion, see for instance [123] and [25, Chapter 6]. Roughly speaking, $\mathcal{D}\phi_p, \mathbf{g}_p$ and $\mathcal{D}\mathbf{S}_p$ should satisfy similar estimates. The goal of this section is to study various quantities like spatial averages and L^p and L^∞ estimates on $\mathbf{S}_p$, as a counterpart of the work [83] concerning ϕ_p.

We can prove at first a very simple result.

Proposition 3.3.1 (Local average)**.** *There exit two finite positive constants $s := s(d,\mathfrak{p},\Lambda)$ and $C := C(d,\mathfrak{p},\Lambda)$ such that*

$$\forall x \in \mathbb{Z}^d, \quad |\mathbf{g}_p|(x) \leqslant \mathcal{O}_s(C|p|). \tag{3.49}$$

Proof. We have, by eq. (3.45)

$$|\mathbf{g}_p| = |\mathbf{a}_{\mathscr{C}}(\mathcal{D}\phi_p + p) - \bar{\mathbf{a}}p| \leqslant |\mathbf{a}_{\mathscr{C}}\mathcal{D}\phi_p| + |\mathbf{a}_{\mathscr{C}}p| + |\bar{\mathbf{a}}p| \leqslant |\nabla\phi_p\mathbf{1}_{\{\mathbf{a}\neq 0\}}| + 2|p| \leqslant \mathcal{O}_s(C|p|).$$

□

3.3.1 Spatial average of centered flux

In this part, we focus on the spatial average quantity $K_R \star [\mathbf{g}_p]$ and prove Proposition 3.1.2. The spirit of the proof can go back to the spectral gap method (or Efron-Stein type inequality) in the work of Naddaf and Spencer [188], which is also employed in the work of Gloria and Otto [123, 124, 120]. Proposition 3.1.2 is more technical in two aspects:

- In the percolation context, the perturbation of the geometry of clusters has to be taken into consideration when applying the spectral gap method.
- The result stated with $\mathcal{O}_s$ notation requires a stronger concentration analysis.

Our proof follows generally the main idea of eq. (3.46) appearing in [83, Proposition 3.1], and the main tool used in this proof is a variant of the Efron-Stein type inequality, combined with the Green's function and Meyers' inequality on $\mathscr{C}_\infty$.

Proof of Proposition 3.1.2. Without loss of generality we suppose that $|p| = 1$, and the proof is decomposed into 4 steps.

Step 1: Spectral gap inequality and double environment. We introduce the Efron-Stein type inequality used for the proof, which is proved first in [27, Proposition 2.2] and also used in [83, Proposition 2.17]. (We remark kindly that there is a typo in the exponent in [27, Proposition 2.2], which should be $\frac{2-\beta}{2}$; see also [27, Appendix A] where the exponent is correct.)

Proposition 3.3.2 (Exponential Efron-Stein inequality, Proposition 2.2 of [27])**.** *Fix $\beta \in (0,2)$ and let X be a random variable defined in the random space $(\Omega, \mathcal{F}, \mathbb{P})$ generated by $\{\mathbf{a}(e)\}_{e\in E_d}$, and we define*

$$\mathcal{F}(E_d\backslash\{e\}) := \sigma\left(\{\mathbf{a}(e')\}_{e'\in E_d\backslash e}\right), \tag{3.50}$$

$$X_e := \mathbb{E}\left[X|\mathcal{F}(E_d\backslash\{e\})\right], \qquad \mathbb{V}[X] := \sum_{e\in E_d}(X - X_e)^2. \tag{3.51}$$

Then, there exists a positive constant $C := C(d,\beta)$ such that

$$\mathbb{E}\left[\exp\left(|X - \mathbb{E}[X]|^\beta\right)\right] \leqslant C\mathbb{E}\left[\exp\left((C\mathbb{V}[X])^{\frac{\beta}{2-\beta}}\right)\right]^{\frac{2-\beta}{2}}. \tag{3.52}$$

In the proof of Proposition 3.3.1, we apply this inequality by posing $X := (K_R \star [\mathbf{g}_p])(x)$ and we claim that it suffices to verify two conditions

$$\mathbb{E}[X] \leqslant C_1 R^{-\frac{d}{2}}, \tag{3.53}$$

$$\mathbb{V}[X] \leqslant \mathcal{O}_{s'}\left(C_2 R^{-d}\right). \tag{3.54}$$

It is also very natural, because the two conditions say that the average and fluctuation of X are of the order of $R^{-\frac{d}{2}}$. We choose a s such that $\frac{s}{2-s} = s'$ where s' is the exponent in eq. (3.54) and $C_3 := (C_1 \vee C_2)C(d,\beta)$ where $C(d,\beta)$ is the constant in eq. (3.52) and C_1, C_2

the one in eq. (3.53), eq. (3.54), then

$$\begin{aligned}\mathbb{E}\left[\exp\left(\left(\frac{X}{C_3R^{-\frac{d}{2}}}\right)^s\right)\right] &\leqslant \mathbb{E}\left[\exp\left(\left(\frac{X-\mathbb{E}[X]}{C_3R^{-\frac{d}{2}}}+\frac{\mathbb{E}[X]}{C_3R^{-\frac{d}{2}}}\right)^s\right)\right] \\ &\leqslant C\,\underbrace{\mathbb{E}\left[\exp\left(\left(\frac{|X-\mathbb{E}[X]|}{C_3R^{-\frac{d}{2}}}\right)^s\right)\right]}_{\text{Using } eq.\ (3.52)} \\ &\leqslant C\mathbb{E}\left[\exp\left(\left(\frac{\mathbb{V}[X]}{C_2R^{-d}}\right)^{\frac{s}{2-s}}\right)\right]^{\frac{2-s}{2}} \\ &\leqslant 2C.\end{aligned}$$

Finally, we increase C_3 with respect to s so that we get $X \leqslant \mathcal{O}_s(CR^{-\frac{d}{2}})$.

We focus on the two conditions eq. (3.53), eq. (3.54). In fact, we can check the condition eq. (3.53) by proving $\mathbb{E}[\mathbf{a}_{\mathscr{C}}(\mathcal{D}\phi_p + p)] = \bar{\mathbf{a}}p$, which is a well-known result in classic homogenization. In percolation context, it is also true by a careful check of the several equivalent definitions of $\bar{\mathbf{a}}$. We put its proof in Theorem 3.C.1.

To prove the condition eq. (3.54), we use a useful technique in Efron-Stein type inequality of "doubling" the probability space: we sample a copy of random conductance $\{\widetilde{\mathbf{a}}(e')\}_{e'\in E_d}$ with the same law but independent to $\{\mathbf{a}(e')\}_{e'\in E_d}$, and the two probability spaces generated by the two copies are denoted respectively by $(\Omega_{\mathbf{a}}, \mathcal{F}_{\mathbf{a}}, \mathbb{P}_{\mathbf{a}}), (\Omega_{\widetilde{\mathbf{a}}}, \mathcal{F}_{\widetilde{\mathbf{a}}}, \mathbb{P}_{\widetilde{\mathbf{a}}})$. Then we put the two copies of random conductance together and make a larger probability space $(\Omega', \mathcal{F}', \mathbb{P}') = (\Omega_{\mathbf{a}} \times \Omega_{\widetilde{\mathbf{a}}}, \mathcal{F}_{\mathbf{a}} \otimes \mathcal{F}_{\widetilde{\mathbf{a}}}, \mathbb{P}_{\mathbf{a}} \otimes \mathbb{P}_{\widetilde{\mathbf{a}}})$, and we also use the notation $\mathcal{O}'_s$ to represent the same definition eq. (3.20) in the larger probability space $(\Omega', \mathcal{F}', \mathbb{P}')$. We also introduce the another random environment $\{\mathbf{a}^e(e')\}_{e'\in E_d}$, obtained by replacing one conductance $\mathbf{a}(e)$ by $\widetilde{\mathbf{a}}(e)$, i.e.

$$\mathbf{a}^e(e') = \begin{cases} \mathbf{a}(e') & e' \neq e, \\ \widetilde{\mathbf{a}}(e') & e' = e. \end{cases} \tag{3.55}$$

We use $X^e, \mathscr{C}^e_\infty, \phi^e_p$ to represent respectively the random variable, the infinite cluster and the corrector in the environment $\{\mathbf{a}^e(e')\}_{e'\in Ed}$. The definition of $\mathbb{V}[X]$ says that the variance comes from the fluctuation caused by the perturbation of every conductance, which suggests the following lemma:

Lemma 3.3.1. *We have the following estimate*

$$\sum_{e\in E_d}(X - X^e)^2 \leqslant \mathcal{O}'_s(CR^{-d}) \Longrightarrow \mathbb{V}[X] \leqslant \mathcal{O}_s(C_sR^{-d}). \tag{3.56}$$

Proof. We use the double environment trick to see that

$$X_e = \mathbb{E}\left[X|\mathcal{F}(E_d\backslash\{e\})\right] = \mathbb{E}_{\widetilde{\mathbf{a}}}[X^e],$$

and Jensen's inequality to reformulate at first the inequality

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(\left(\frac{\mathbb{V}[X]}{CR^{-d}}\right)^s\right)\right] &= \int_\Omega \exp\left(\left(\frac{\sum_{e\in E_d}(X-X_e)^2}{CR^{-d}}\right)^s\right) d\mathbb{P}_{\mathbf{a}}(\omega) \\
&= \int_\Omega \exp\left(\left(\frac{\sum_{e\in E_d}\left(\int_{\Omega_{\widetilde{\mathbf{a}}}}(X-X^e)\, d\mathbb{P}_{\widetilde{\mathbf{a}}}(\omega)\right)^2}{CR^{-d}}\right)^s\right) d\mathbb{P}_{\mathbf{a}}(\omega) \\
&\leqslant \int_\Omega \exp\left(\left(\int_{\Omega_{\widetilde{\mathbf{a}}}} \frac{\sum_{e\in E_d}(X-X^e)^2}{CR^{-d}}\, d\mathbb{P}_{\widetilde{\mathbf{a}}}(\omega)\right)^s\right) d\mathbb{P}_{\mathbf{a}}(\omega)
\end{aligned}
$$

In the next step, we want to add a constant t_s to make $\exp((\cdot+t_s)^s)$ convex, and then exchange the expectation and $\exp((\cdot+t_s)^s)$ by Jensen's inequality. We can choose $t_s=0$ for $s\geqslant 1$, and $t_s=\left(\frac{1-s}{s}\right)^{\frac{1}{s}}$ for $0<s<1$. (The spirit is the same as eq. (3.23) and see [25, Lemma A.4] for details of this proof.)

$$
\begin{aligned}
\mathbb{E}\left[\exp\left(\left(\frac{\mathbb{V}[X]}{CR^{-d}}\right)^s\right)\right] &\leqslant \int_\Omega\int_{\Omega_{\widetilde{\mathbf{a}}}} \exp\left(\left(\frac{\sum_{e\in E_d}(X-X^e)^2}{CR^{-d}}+t_s\right)^s\right) d\mathbb{P}_{\widetilde{\mathbf{a}}}(\omega)\, d\mathbb{P}_{\mathbf{a}}(\omega) \\
&\leqslant \widetilde{C}\int_\Omega\int_{\Omega_{\widetilde{\mathbf{a}}}} \exp\left(\left(\frac{\sum_{e\in E_d}(X-X^e)^2}{CR^{-d}}\right)^s\right) d\mathbb{P}_{\widetilde{\mathbf{a}}}(\omega)\, d\mathbb{P}_{\mathbf{a}}(\omega) \\
&\leqslant 2\widetilde{C}.
\end{aligned}
$$

In the last step we use the condition $\sum_{e\in E_d}(X-X^e)^2 \leqslant \mathcal{O}_s'(CR^{-d})$ and we reduce the constant C to get the desired result. □

By Lemma 3.3.1, to prove eq. (3.54) it suffices to focus on the quantity $\sum_{e\in E_d}(X-X^e)^2$, and in our context it is

$$
\sum_{e\in E_d}\left|K_R\star\left([\mathbf{g}_p]-[\mathbf{g}_p^e]\right)\right|^2(x)\leqslant \mathcal{O}_s'(CR^{-d}). \tag{3.57}
$$

Since we have

$$
\begin{aligned}
&\left|K_R\star\left([\mathbf{g}_p]-[\mathbf{g}_p^e]\right)\right|(x) \\
&\qquad\leqslant \left|K_R\star\left([\mathbf{g}_p]-[\mathbf{g}_p^e]\right)\right|(x)\mathbf{1}_{\{\mathbf{a}^e(e)\leqslant\mathbf{a}(e)\}}+\left|K_R\star\left([\mathbf{g}_p]-[\mathbf{g}_p^e]\right)\right|(x)\mathbf{1}_{\{\mathbf{a}(e)\leqslant\mathbf{a}^e(e)\}},
\end{aligned}
$$

and the two terms have the same law, without loss of generality, we suppose

$$
\mathbf{a}^e(e)\leqslant\mathbf{a}(e), \tag{3.58}
$$

is always valid in the following paragraphs in order to avoid the indicator function everywhere. We will then distinguish several cases and attack them one by one.

Step 2: Case $\mathscr{C}_\infty\neq\mathscr{C}_\infty^e$, proof of $\mathbf{g}_p=\mathbf{g}_p^e$. We have to consider the perturbation of the geometry between $\mathscr{C}_\infty$ and $\mathscr{C}_\infty^e$. We prove the following lemma, which has a typical realization in Figure 3.2.

Lemma 3.3.2 (Pivot edge)**.** *Under the condition eq. (3.58) and in the case $\mathscr{C}_\infty\neq\mathscr{C}_\infty^e$, we have:*

1. *The part $\mathscr{C}_\infty\backslash\mathscr{C}_\infty^e$ is connected to $\mathscr{C}_\infty^e$ by e (called the pivot edge), and $|\mathscr{C}_\infty\backslash\mathscr{C}_\infty^e|<\infty$.*

2. *We denote by* $e := \{e_*, e^*\}, e_* \in \mathscr{C}_\infty^e \cap \mathscr{C}_\infty$ *and* $e^* \in \mathscr{C}_\infty \backslash \mathscr{C}_\infty^e$*, then the function* $(\phi_p + l_p)$ *is constant on* $\mathscr{C}_\infty \backslash \mathscr{C}_\infty^e$ *and equals to* $(\phi_p + l_p)(e_*)$*.*

3. *The function* ϕ_p^e *has a representation that* $\phi_p^e = \phi_p \mathbf{1}_{\{\mathscr{C}_\infty^e\}}$ *up to a constant and satisfies* $\mathbf{a}_{\mathscr{C}} \nabla(\phi_p + l_p) = \mathbf{a}_{\mathscr{C}}^e \nabla(\phi_p^e + l_p)$ *on* E_d*.*

Proof. 1. It comes from the fact that $\mathbf{a}$ and $\mathbf{a}^e$ are different only by one edge, thus $\mathscr{C}_\infty^e \subsetneq \mathscr{C}_\infty$ means that $\mathbf{a}(e) > 0$ in $\mathscr{C}_\infty$ but $\mathbf{a}^e(e) = 0$ in $\mathscr{C}_\infty^e$ and makes one part disconnected from $\mathscr{C}_\infty$. It is well-known that in the supercritical percolation, almost surely there exists one unique infinite cluster, thus we have $|\mathscr{C}_\infty \backslash \mathscr{C}_\infty^e| < \infty$.

2. We study the harmonic function $-\nabla \cdot \mathbf{a}_{\mathscr{C}} \nabla(\phi_p + l_p) = 0$ on the part $\mathscr{C}_\infty \backslash \mathscr{C}_\infty^e$. This is a non-degenerate linear system with $|\mathscr{C}_\infty \backslash \mathscr{C}_\infty^e|$ equations and $|\mathscr{C}_\infty \backslash \mathscr{C}_\infty^e| + 1$ variables, thus the solution is of 1 dimension and we know this constant is $(\phi_p + l_p)(e_*)$.

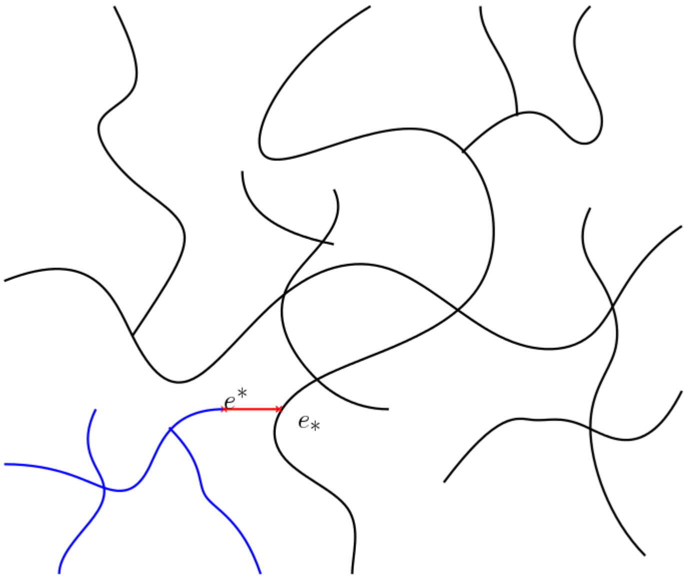

Figure 3.2: In the image the segment in red is the edge $e = \{e_*, e^*\}$ and the part in blue is the cluster $\mathscr{C}_\infty \backslash \mathscr{C}_\infty^e$, where $\mathbf{a}$-harmonic function $(\phi_p + l_p)$ is constant of value $(\phi_p + l_p)(e_*)$.

3. We prove that at first that $\mathbf{a}_{\mathscr{C}} \nabla(\phi_p + l_p) = \mathbf{a}_{\mathscr{C}}^e \nabla(\phi_p \mathbf{1}_{\{\mathscr{C}_\infty^e\}} + l_p)$ for every $e \in E_d$.

   - For the edge e' such that $\mathbf{a}_{\mathscr{C}}(e') = 0$, as $\mathbf{a}_{\mathscr{C}}^e(e') \leqslant \mathbf{a}_{\mathscr{C}}(e')$, the two functions $\mathbf{a}_{\mathscr{C}} \nabla(\phi_p + l_p)(e')$ and $\mathbf{a}_{\mathscr{C}}^e \nabla(\phi_p \mathbf{1}_{\{\mathscr{C}_\infty^e\}} + l_p)(e')$ are null.
   - For the only pivot edge e that $\mathbf{a}_{\mathscr{C}}(e) > 0, \mathbf{a}_{\mathscr{C}}^e(e) = 0$, thanks to the second term of Lemma 3.3.2, we have $\nabla(\phi_p + l_p)(e) = 0$. Thus, the equation also establishes.
   - For the edge that $\mathbf{a}_{\mathscr{C}}(e') > 0, \mathbf{a}_{\mathscr{C}}^e(e') > 0$, we know that this implies that the two endpoints are on $\mathscr{C}_\infty^e$ so that we have
   $$\nabla(\phi_p + l_p)(e') = \nabla(\phi_p \mathbf{1}_{\{\mathscr{C}_\infty^e\}} + l_p)(e'),$$
   and $\mathbf{a}_{\mathscr{C}}(e') = \mathbf{a}_{\mathscr{C}}^e(e')$, so the equation is also established.

   $\mathbf{a}_{\mathscr{C}} \nabla(\phi_p + l_p) = \mathbf{a}_{\mathscr{C}}^e \nabla(\phi_p \mathbf{1}_{\{\mathscr{C}_\infty^e\}} + l_p)$ implies directly that $-\nabla \cdot \mathbf{a}_{\mathscr{C}}^e \nabla(\phi_p \mathbf{1}_{\{\mathscr{C}_\infty^e\}} + l_p) = 0$ on $\mathbb{Z}^d$, therefore, by the Liouville regularity, we obtain that $\phi_p^e = \phi_p \mathbf{1}_{\{\mathscr{C}_\infty^e\}}$ on $\mathscr{C}_\infty^e$ up to a constant.

□

A direct corollary of the third part of Lemma 3.3.2 is that $\mathbf{g}_p = \mathbf{g}_p^e$ on E_d when $\mathscr{C}_\infty \neq \mathscr{C}_\infty^e$, thus $K_R \star ([\mathbf{g}_p] - [\mathbf{g}_p^e]) = 0$. So, it suffices to consider $\sum_{e\in E_d} |K_R \star ([\mathbf{g}_p] - [\mathbf{g}_e^e])|^2 (x)$ under the condition $\mathscr{C}_\infty = \mathscr{C}_\infty^e$. Then, we can reformulate the quantity in eq. (3.57) as following:

$$\begin{aligned}
&K_R \star ([\mathbf{g}_p] - [\mathbf{g}_p^e])(x)\\
&= K_R \star \left([\mathbf{a}_{\mathscr{C}}\mathcal{D}(\phi_p + l_p)] - [\mathbf{a}_{\mathscr{C}}^e\mathcal{D}(\phi_p^e + l_p)]\right)(x)\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}}\\
&= \underbrace{K_R \star \left[(\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C}}^e)\mathcal{D}(\phi_p^e + l_p)\right](x)\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}}}_{:=A_e(x)} + \underbrace{K_R \star \left[\mathbf{a}_{\mathscr{C}}\mathcal{D}(\phi_p - \phi_p^e)\right](x)\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}}}_{:=B_e(x)}.
\end{aligned}$$

In order to prove eq. (3.57), we study $A_e(x)$ and $B_e(x)$ separately.

Step 3: Case $\mathscr{C}_\infty = \mathscr{C}_\infty^e$, proof of $\sum_{e\in E_d} |A_e(x)|^2 \leqslant \mathcal{O}_s'(CR^{-d})$. Lemma 3.3.2 helps us simplify the discussion on the case $\mathscr{C}_\infty = \mathscr{C}_\infty^e$ and following lemma carries the convolution to the cluster $\mathscr{C}_\infty$.

Lemma 3.3.3. *For a kernel K_R as in Proposition 3.1.2 and every $x \in \mathbb{R}^d$, there exists a function $\Gamma_{K,R}^x : \mathbb{Z}^d \to \mathbb{R}^+$ such that for every function ξ supported on $\mathscr{C}_\infty$, we have*

$$(K_R \star [\xi])(x) = \langle \Gamma_{K,R}^x, \xi \rangle_{\mathscr{C}_\infty}, \tag{3.59}$$

and we have the estimate

$$\Gamma_{K,R}^x(z) \leqslant \frac{2^d C_{K,R}}{R^d\left(|\frac{x-z}{R}| \vee 1\right)^{\frac{d+1}{2}}}. \tag{3.60}$$

Proof. We can do the calculation directly

$$\begin{aligned}
(K_R \star [\xi])(x) &= \int_{\mathbb{R}^d} [\xi](y) K_R(x-y)\, dy\\
&= \int_{\mathbb{R}^d} \left(\sum_{z\in\mathscr{C}_\infty} \mathbf{1}_{\{y\in z+\square\}} [\xi](z)\right) K_R(x-y)\, dy\\
&= \sum_{z\in\mathscr{C}_\infty} \left(\int_{\mathbb{R}^d} \mathbf{1}_{\{y\in z+\square\}} K_R(x-y)\, dy\right) \xi(z).
\end{aligned}$$

Thus we can define

$$\Gamma_{K,R}^x(z) := \int_{y\in z+\square} K_R(x-y)\, dy. \tag{3.61}$$

The estimate eq. (3.60) comes directly from this expression and $K_R \leqslant \frac{C_{K,R}}{R^d(|\frac{x}{R}|\vee 1)^{\frac{d+1}{2}}}$. □

We want to apply directly Lemma 3.3.3 to every random environment $\mathbf{a}^e$ to $A_e(x)$. We see that it suffices to study the case $e \in E_d(\mathscr{C}_\infty)$, otherwise the condition $\mathscr{C}_\infty = \mathscr{C}_\infty^e$ will not be satisfied or $(\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C}}^e)\mathcal{D}(\phi_p^e + l_p)$ will be 0. Thus, $\operatorname{supp}\left((\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C}}^e)\mathcal{D}_{e_i}(\phi_p^e + l_p)\right) \subseteq \mathscr{C}_\infty$ and Lemma 3.3.3 works.

$$\begin{aligned}
\sum_{e\in E_d} |A_e|^2(x) &= \sum_{i=1}^d \sum_{e\in E_d(\mathscr{C}_\infty)} \left|K_R \star \left[(\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C}}^e)\mathcal{D}_{e_i}(\phi_p^e + l_p)\right]\right|^2 (x)\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}}\\
&= \sum_{i=1}^d \sum_{e\in E_d(\mathscr{C}_\infty)} \left|\left\langle \Gamma_{K,R}^x, \left[(\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C}}^e)\mathcal{D}(\phi_p^e + l_p)\right]_i \right\rangle_{\mathscr{C}_\infty}\right|^2 \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \qquad (3.62)\\
&\leqslant \sum_{ß=1}^d \sum_{z\in\mathscr{C}_\infty} |\Gamma_{K,R}^x(\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C}}^{\{z,z+e_i\}})\mathcal{D}_{e_i}(\phi_e^{\{z,z+e_i\}} + l_p)|(z)^2 \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}}.
\end{aligned}$$

The third line comes from the fact that under the condition $\mathscr{C}_\infty = \mathscr{C}_\infty^e$, only one conductance $\mathbf{a}(e)$ and $\mathbf{a}^e(e)$ is different.

Using eq. (3.60), we know the part $|\Gamma_{K,R}^x|^2$ is integrable. It remains to estimate the vector field Θ

$$\forall e = \{e_*, e^*\} \in E_d, \qquad \Theta(e_*, e^*) := (\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C}}^e)\nabla(\phi_p^e + l_p)(e_*, e^*), \tag{3.63}$$

and prove a stochastic integrability for eq. (3.62). Since the quantity $\Theta(e)$ plays an important role in our analysis and we will use it several times, we prove the following lemma:

Lemma 3.3.4. *Under the condition eq. (3.58), there exist two finite positive constants $s(d, \mathfrak{p}, \Lambda)$ and $C(d, \mathfrak{p}, \Lambda)$ such that*

$$\forall e \in E_d, \qquad |\Theta(e)|\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \leqslant \mathcal{O}_s'(C). \tag{3.64}$$

Proof. This estimate is very easy when $\mathbf{a}_{\mathscr{C}}^e(e) > 0$, since it implies $\mathbf{a}_{\mathscr{C}}^e(e) > \Lambda^{-1}$ and we obtain $\mathbf{a}_{\mathscr{C}}(e) \leqslant \Lambda \mathbf{a}_{\mathscr{C}}^e(e)$ together with eq. (3.58). We then use Proposition 3.3.1 directly that

$$|(\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C}}^e)\mathcal{D}(\phi_p^e + l_p)|\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \leqslant (1+\Lambda)|\mathbf{a}_{\mathscr{C}}^e \mathcal{D}(\phi_p^e + l_p)|\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \leqslant \mathcal{O}_s'(C).$$

The less immediate part comes from the case $\mathbf{a}_{\mathscr{C}}^e(e) = 0$ while $\mathbf{a}_{\mathscr{C}}(e) > 0$, where $\mathbf{a}_{\mathscr{C}}^e \mathcal{D}(\phi_p^e + l_p) = 0$ and cannot be used to dominate $|\Theta|(e)$. We treat this case differently: we denote by $e = \{e_*, e^*\}$, $\mathscr{C}_\infty = \mathscr{C}_\infty^e$ implies the existence of another open path γ in $\mathscr{C}_\infty^e$ connecting e_* and e^* (see Figure 3.3). This path can be chosen in $\mathscr{C}_*(\square_{\mathcal{P}^e}(e_*)) \cup \mathscr{C}_*(\square_{\mathcal{P}^e}(e^*))$ applying Lemma 3.2.3 to the partition cube $\mathcal{P}^e$.

$$\begin{aligned} |\Theta(e_*, e^*)|\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} &= |(\mathbf{a}_{\mathscr{C}}^e - \mathbf{a}_{\mathscr{C}})\nabla(\phi_p^e + l_p)|(e_*, e^*)\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \\ &\leqslant 2 \sum_{e' \in \gamma \subseteq \mathscr{C}_*(\square_{\mathcal{P}^e}(e_*)) \cup \mathscr{C}_*(\square_{\mathcal{P}^e}(e^*))} |\nabla(\phi_p^e + l_p)|(e')\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \end{aligned} \tag{3.65}$$

In this sum, we have to notice that not only the corrector ϕ_p^e is random, but also the path γ and its length. This forbids us to use directly eq. (3.23) or eq. (3.25), so we apply the minimal scale argument and eq. (3.44) in the environment $\{\mathbf{a}^e(\widetilde{e})\}_{\widetilde{e} \in E_d}$: there exists $s := s(d, \mathfrak{p}, \Lambda) > 0$ and $C := C(d, \mathfrak{p}, \Lambda) < \infty$ such that for any $x \in \mathscr{C}_\infty^e$, we have a random variable $\mathcal{M}^e(x) \leqslant \mathcal{O}_s'(C)$ and for every $r \geqslant \mathcal{M}^e(x)$,

$$\left\| \nabla \phi_p^e \mathbf{1}_{\{\mathbf{a}^e \neq 0\}} \right\|_{\underline{L}^2(\mathscr{C}_\infty^e \cap B_r(x))} \leqslant C. \tag{3.66}$$

Then we take

$$\widetilde{\mathcal{M}}(e) = \max \left\{ \mathcal{M}^e(e^*), \operatorname{size}(\square_{\mathcal{P}^e}(e^*)), \operatorname{size}(\square_{\mathcal{P}^e}(e_*)) \right\},$$

and it is clear that $\widetilde{\mathcal{M}}(e) \leqslant \mathcal{O}_s'(C)$ and the ball $B_{\widetilde{\mathcal{M}}(e)}(e^*)$ contains $\square_{\mathcal{P}^e}(e^*)$ and $\square_{\mathcal{P}^e}(e_*)$. Thus we can use eq. (3.66) and Cauchy-Schwarz inequality to control the sum over the path γ

$$\begin{aligned} |\Theta(e_*, e^*)|\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} &\leqslant 2 \sum_{e' \in \mathscr{C}_\infty^e \cap B_{\widetilde{\mathcal{M}}(e)}(e^*)} |\nabla(\phi_p^e + l_p)|(e')\mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \\ &\leqslant (\widetilde{\mathcal{M}}(e))^d \left\| \nabla(\phi_p^e + l_p)\mathbf{1}_{\{\mathbf{a}^e \neq 0, \mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \right\|_{\underline{L}^2(\mathscr{C}_\infty^e \cap B_{\widetilde{\mathcal{M}}(e)}(e^*))} \\ &\leqslant C(\widetilde{\mathcal{M}}(e))^d. \end{aligned}$$

Finally we use $(\widetilde{\mathcal{M}}(e))^d \leqslant \mathcal{O}_{s/d}'(C)$ to conclude the proof of Lemma 3.3.4. □

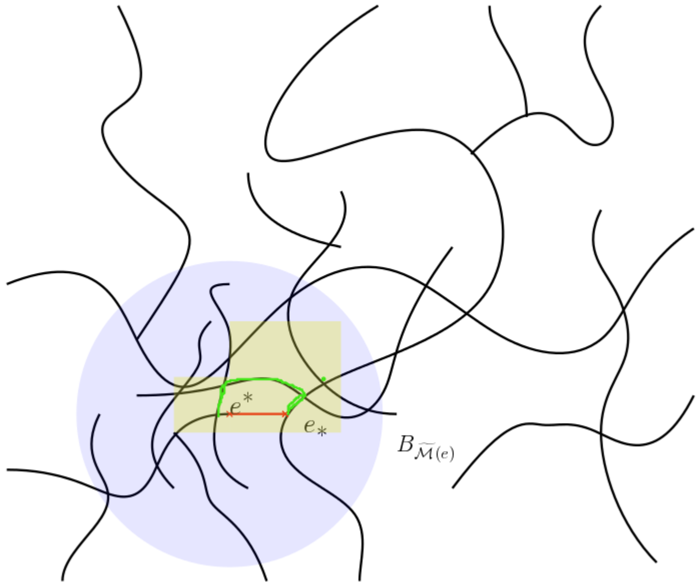

Figure 3.3: This image shows the case $\mathscr{C}_\infty = \mathscr{C}_\infty^e$ and $\mathbf{a}^e_{\mathscr{C}}(e) = 0$ while $\mathbf{a}_{\mathscr{C}}(e) > 0$, so $e = \{e_*, e^*\}$ (the segment in red) is an open bond in $\mathscr{C}_\infty$ but not in $\mathscr{C}_\infty^e$. The condition $\mathscr{C}_\infty = \mathscr{C}_\infty^e$ ensures another open path γ (the segment in green) in $\mathscr{C}_\infty^e$ connecting e_* and e^*. The path γ is contained in the union of the partition cube $\square_{\mathcal{P}^e}(e_*)$ and $\square_{\mathcal{P}^e}(e^*)$ (the cubes in yellow). To estimate the sum of $\nabla \phi_p^e$ over this path γ, we choose a minimal scale $\widetilde{\mathcal{M}}(e)$ and study the average in this scale (the ball in blue).

We conclude from eq. (3.62), eq. (3.60) and Lemma 3.3.4 that

$$\begin{aligned}
\sum_{e \in E_d} |A_e|^2(x) &\leqslant \sum_{i=1}^d \sum_{z \in \mathscr{C}_\infty} |\Gamma^x_{K,R}|^2(z) |\Theta|^2(z, z+e_i) \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \\
&\leqslant \sum_{i=1}^d \sum_{z \in \mathbb{Z}^d} \frac{4^d C^2_{K,R}}{R^{2d}(\left|\frac{x-z}{R}\right| \vee 1)^{d+1}} |\Theta|^2(z, z+e_i) \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \\
&\leqslant \mathcal{O}'_s\left(\frac{C^2_{K,R} C(d, \mathfrak{p}, \Lambda)}{R^d}\right).
\end{aligned}$$

Step 4: Case $\mathscr{C}_\infty = \mathscr{C}_\infty^e$, proof of $\sum_{e \in E_d} |B_e|^2(x) \leqslant \mathcal{O}'_s(CR^{-d})$. This step is similar to that for A_e but more technical. We define a space

$$\dot{H}^1(\mathscr{C}_\infty) := \{v : \mathscr{C}_\infty \to \mathbb{R}, \langle \nabla v, \nabla v \rangle_{E_d^{\mathbf{a}}(\mathscr{C}_\infty)} < \infty\},$$

and use the Green's function on $(\mathscr{C}_\infty, E_d^{\mathbf{a}}(\mathscr{C}_\infty))$ [83, Proposition 2.11]:

Proposition 3.3.3 (Green's function on $\mathscr{C}_\infty$)**.** *Let $\mathbf{a} \in \Omega$ be an environment with an infinite cluster $\mathscr{C}_\infty$ and $x, y \in \mathscr{C}_\infty$, then there exist a constant $C := C(d, \Lambda) < \infty$ and a Green's function $G^{x,y} \in \dot{H}^1(\mathscr{C}_\infty)$ such that*

$$-\nabla \cdot \mathbf{a}_{\mathscr{C}} \nabla G^{x,y} = \delta_y - \delta_x \ \text{ on } \mathscr{C}_\infty,$$

in the sense for any $v \in \dot{H}^1(\mathscr{C}_\infty)$, we have

$$\langle \nabla G^{x,y}, \mathbf{a}_{\mathscr{C}} \nabla v \rangle_{E_d^{\mathbf{a}}(\mathscr{C}_\infty)} = v(y) - v(x).$$

In the case that $e = (x,y) \in \overrightarrow{E_d^{\mathbf{a}}}(\mathscr{C}_\infty)$, *we note* $G^{x,y} := G^e$. *The Green's function* $G^{x,y}$ *has the following properties*

- *Symmetry: For every* $x, y, x', y' \in \mathscr{C}_\infty$, *we have* $G^{x,y}(y') - G^{x,y}(x') = G^{x',y'}(y) - G^{x',y'}(x)$.
- *Representation: For every* $v \in \dot{H}^1(\mathscr{C}_\infty)$, *every vector field* $\xi : \overrightarrow{E_d^{\mathbf{a}}}(\mathscr{C}_\infty) \to \mathbb{R}$, *and* $u_\xi \in \dot{H}^1(\mathscr{C}_\infty)$ *such that*

$$\left\langle \nabla u_\xi, \mathbf{a}_{\mathscr{C}} \nabla v \right\rangle_{E_d^{\mathbf{a}}(\mathscr{C}_\infty)} = \langle \xi, \nabla v \rangle_{E_d^{\mathbf{a}}(\mathscr{C}_\infty)},$$

we have the representation

$$\nabla u_\xi = \sum_{e \in E_d^{\mathbf{a}}(\mathscr{C}_\infty)} \xi(e) \nabla G^e. \tag{3.67}$$

In this formula, we give an arbitrary orientation for $e \in E_d^{\mathbf{a}}(\mathscr{C}_\infty)$, *and the equation is well-defined.*

Proof. The proof of the existence and uniqueness up to a constant for the function $G^{x,y}$ comes from the Lax-Milgram theorem on the space $\dot{H}^1(\mathscr{C}_\infty)$ where the conductance satisfies the quenched uniform ellipticity condition. The symmetry comes from testing the equation $-\nabla \cdot \mathbf{a}_{\mathscr{C}} \nabla G^{x,y} = \delta_y - \delta_x$ by $G^{x',y'}$ and testing the equation $-\nabla \cdot \mathbf{a}_{\mathscr{C}} \nabla G^{x',y'} = \delta_{y'} - \delta_{x'}$ by $G^{x,y}$ that

$$G^{x,y}(y') - G^{x,y}(x') = \left\langle \nabla G^{x,y}, \mathbf{a}_{\mathscr{C}} \nabla G^{x',y'} \right\rangle_{E_d^{\mathbf{a}}(\mathscr{C}_\infty)} = G^{x',y'}(y) - G^{x',y'}(x).$$

The final representation formula can be checked easily by the linear combination of the Green's function. □

The proof of $\sum_{e \in E_d} |B_e|^2(x) \leqslant \mathcal{O}'_s(CR^{-d})$ can be divided in 4 steps.
Step 4.1: Identification of $\mathcal{D}(\phi_p - \phi_p^e)$ *using Green's function.* We identify at first $\mathcal{D}(\phi_p - \phi_p^e)$ by using the Green's function G^e introduced in Proposition 3.3.3 and then estimate its size by this representation. We want to carry all the analysis on the geometry $(\mathscr{C}_\infty, E_d^{\mathbf{a}}(\mathscr{C}_\infty))$ and to claim the following lemma:

Lemma 3.3.5. *We denote* $e := \{e_*, e^*\} \in E_d^{\mathbf{a}}(\mathscr{C}_\infty)$, *under the condition* $\mathscr{C}_\infty = \mathscr{C}_\infty^e$, *then we have the following representation for* $(\phi_p^e - \phi_p)$, *using Proposition 3.3.3 and the definition* Θ *in eq.* (3.63),

$$\nabla(\phi_p^e - \phi_p)(\cdot) = \Theta(e_*, e^*) \nabla G^{e_*, e^*}(\cdot) \text{ on } \overrightarrow{E_d^{\mathbf{a}}}(\mathscr{C}_\infty). \tag{3.68}$$

Proof. Using the $\mathbf{a}$-harmonic equation and $\mathbf{a}^e$-harmonic equation for their correctors, we have at first

$$-\nabla \cdot \mathbf{a}_{\mathscr{C}} \nabla(\phi_p + l_p) = -\nabla \cdot \mathbf{a}_{\mathscr{C}}^e \nabla(\phi_p^e + l_p) \text{ on } \mathbb{Z}^d,$$

then we obtain that

$$-\nabla \cdot \mathbf{a}_{\mathscr{C}} \nabla(\phi_p - \phi_p^e) = -\nabla \cdot (\mathbf{a}_{\mathscr{C}}^e - \mathbf{a}_{\mathscr{C}}) \nabla(\phi_p^e + l_p) \text{ on } \mathbb{Z}^d. \tag{3.69}$$

Using the definition $\Theta(e_*, e^*) = (\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C}}^e) \nabla(\phi_p^e + l_p)(e_*, e^*)$ and under the condition $\mathscr{C}_\infty = \mathscr{C}_\infty^e$, the right hand side of eq. (3.69) equals to $\Theta(e)(\delta_{e^*} - \delta_{e_*})$. Moreover, since $e_*, e^* \in \mathscr{C}_\infty$, eq. (3.69) can be seen restricted on the cluster $(\mathscr{C}_\infty, E_d^{\mathbf{a}}(\mathscr{C}_\infty))$. Thus we solve the in $\dot{H}^1(\mathscr{C}_\infty)$ the equation

$$-\nabla \cdot \mathbf{a}_{\mathscr{C}} \nabla \widetilde{w}^e = -\nabla \cdot (\mathbf{a}_{\mathscr{C}}^e - \mathbf{a}_{\mathscr{C}}) \nabla(\phi_p^e + l_p) \text{ on } \mathscr{C}_\infty,$$

and by Proposition 3.3.3, it has a unique solution up to a constant that $\widetilde{w}^e = \Theta(e_*, e^*)G^{e_*,e^*}$.

Now we have $(\phi_p - \phi_p^e)$ and $\widetilde{w}^e$ solving the same equation, but we do not yet know if $(\phi_p - \phi_p^e)$ belongs to $\dot{H}^1(\mathscr{C}_\infty)$. We hope to identify that $\phi_p - \phi_p^e = \widetilde{w}^e$ and the argument is to use the Liouville regularity theorem: notice that $(\phi_p - \phi_p^e - \widetilde{w}^e)$ is an $\mathbf{a}$-harmonic function on $\mathscr{C}_\infty$ and $\langle \nabla \widetilde{w}^e, \nabla \widetilde{w}^e \rangle_{E_d^{\mathbf{a}}} < \infty$ implies that $(\phi_p - \phi_p^e - \widetilde{w}^e) \in \mathcal{A}_1$. We claim that it is in fact in $\mathcal{A}_0$ and prove by contradiction: suppose that $(\phi_p - \phi_p^e - \widetilde{w}^e) \in \mathcal{A}_1 \backslash \mathcal{A}_0$, then by the Liouville regularity there exists $h \neq 0$ such that

$$\phi_p - \phi_p^e - \widetilde{w}^e = l_h + \phi_h.$$

However, this implies that $\widetilde{w}^e = \phi_p - \phi_p^e - \phi_h - l_h$, so $\widetilde{w}^e$ has an asymptotic linear increment at infinity, which contradicts the fact that $\widetilde{w}^e \in \dot{H}^1(\mathscr{C}_\infty)$. In conclusion, we have $\phi_p - \phi_p^e - \Theta(e_*, e^*)G^{e_*,e^*} = c$ and we get eq. (3.68). □

Step 4.2: Carry the analysis on $(\mathscr{C}_\infty, E_d^{\mathbf{a}}(\mathscr{C}_\infty))$. Observing that we do the sum of $\mathbf{a}_{\mathscr{C}} \mathcal{D}(\phi_p - \phi_p^e)$, it suffices to do the sum over $E_d^{\mathbf{a}}(\mathscr{C}_\infty)$ and with the help of Lemma 3.3.5 , we have the formula

$$\begin{aligned}
\sum_{e \in E_d} |B_e|^2(x) &= \sum_{i=1}^{d} \sum_{e \in E_d^{\mathbf{a}}(\mathscr{C}_\infty)} \left|K_R \star \left[\mathbf{a}_{\mathscr{C}} \mathcal{D}_{e_i}(\phi_p - \phi_p^e)\right]\right|^2(x) \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \\
&= \sum_{i=1}^{d} \sum_{e \in E_d^{\mathbf{a}}(\mathscr{C}_\infty)} \left|K_R \star \left[\mathbf{a}_{\mathscr{C}} \mathcal{D}_{e_i} G^e\right]\right|^2 (x) \Theta^2(e) \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} \qquad (3.70) \\
&= \sum_{i=1}^{d} \sum_{e \in E_d^{\mathbf{a}}(\mathscr{C}_\infty)} \left|\left\langle \Gamma_{K,R}^x, \mathbf{a}_{\mathscr{C}} \mathcal{D}_{e_i} G^e \right\rangle_{\mathscr{C}_\infty}\right|^2 (x) \Theta^2(e) \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}}.
\end{aligned}$$

We analyze $\left\langle \Gamma_{K,R}^x, \mathbf{a}_{\mathscr{C}} \mathcal{D}_{e_i} G^e \right\rangle_{\mathscr{C}_\infty} \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}}$ by defining the notation that

$$\mathbf{1}_{\{E_d^i\}}(e') = \begin{cases} 1 & \text{if } \exists z \in \mathbb{Z}^d \text{ such that } e' = \{z, z + e_i\} \\ 0 & \text{Otherwise} \end{cases},$$

and a vector field $\widetilde{\Gamma}_{K,R,i}^x : E_d^{\mathbf{a}}(\mathscr{C}_\infty) \to \mathbb{R}$ that

$$\widetilde{\Gamma}_{K,R,i}^x \coloneqq \Gamma_{K,R}^x(e'_{*,i}) \mathbf{a}_{\mathscr{C}}(e') \mathbf{1}_{\{E_d^i\}}(e'), \qquad e'_{*,i} \in \mathscr{C}_\infty \text{ such that } e' = \left\{e'_*, e'_* + e_i\right\}. \qquad (3.71)$$

Then, we can send $\left\langle \Gamma_{K,R}^x, \mathbf{a}_{\mathscr{C}} \mathcal{D}_{e_i} G^e \right\rangle_{\mathscr{C}_\infty} \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}}$ to the inner product of vector field on $E_d^{\mathbf{a}}(\mathscr{C}_\infty)$

$$\left\langle \Gamma_{K,R}^x, \mathbf{a}_{\mathscr{C}} \mathcal{D}_{e_i} G^e \right\rangle_{\mathscr{C}_\infty} \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}} = \left\langle \widetilde{\Gamma}_{K,R,i}^x, \nabla G^e \right\rangle_{E_d^{\mathbf{a}}(\mathscr{C}_\infty)} \mathbf{1}_{\{\mathscr{C}_\infty = \mathscr{C}_\infty^e\}}. \qquad (3.72)$$

Step 4.3 : Apply once again the representation with the Green's function. Since $\widetilde{\Gamma}_{K,R,i}^x$ is defined on $E_d^{\mathbf{a}}(\mathscr{C}_\infty)$, we can apply Proposition 3.3.3 to define $w_{\widetilde{\Gamma}_{K,R,i}^x} \in \dot{H}_1(\mathscr{C}_\infty)$ the solution of the equation

$$-\nabla \cdot \mathbf{a}_{\mathscr{C}} \nabla w_{\widetilde{\Gamma}_{K,R,i}^x} = -\nabla \cdot \widetilde{\Gamma}_{K,R,i}^x, \text{ on } \mathscr{C}_\infty, \qquad (3.73)$$

and it has a representation $\nabla w_{\widetilde{\Gamma}_{K,R,i}^x}(e) = \sum_{e' \in E_d^{\mathbf{a}}(\mathscr{C}_\infty)} \widetilde{\Gamma}_{K,R,i}^x(e') \nabla G^{e'}(e)$. We use the symmetry $\nabla G^{e'}(e) = \nabla G^e(e')$

$$\nabla w_{\widetilde{\Gamma}_{K,R,i}^x}(e) = \sum_{e' \in E_d^{\mathbf{a}}(\mathscr{C}_\infty)} \widetilde{\Gamma}_{K,R,i}^x(e') \nabla G^e(e') = \left\langle \widetilde{\Gamma}_{K,R,i}^x, \nabla G^e \right\rangle_{E_d^{\mathbf{a}}(\mathscr{C}_\infty)}. \qquad (3.74)$$

We combine eq. (3.70) eq. (3.72) and eq. (3.74) together and obtain that

$$\sum_{e\in E_d}|B_e|^2(x) \leqslant \sum_{i=1}^{d}\sum_{e\in E_d^{\mathbf{a}}(\mathscr{C}_\infty)} |\nabla w_{\widetilde{\Gamma}^x_{K,R,i}}|^2(e)\Theta^2(e)\mathbf{1}_{\{\mathscr{C}_\infty=\mathscr{C}_\infty^\varepsilon\}}. \tag{3.75}$$

Step 4.4: Meyers' inequality and minimal scale. From the eq. (3.73), we obtain a $\dot{H}^1(\mathscr{C}_\infty)$ estimate using eq. (3.60)

$$\begin{aligned}&\left\langle \nabla w_{\widetilde{\Gamma}^x_{K,R,i}}, \mathbf{a}_{\mathscr{C}}\nabla w_{\widetilde{\Gamma}^x_{K,R,i}}\right\rangle^{\frac{1}{2}}_{E_d^{\mathbf{a}}(\mathscr{C}_\infty)} = \left\langle \nabla w_{\widetilde{\Gamma}^x_{K,R,i}}, \widetilde{\Gamma}^x_{K,R,i}\right\rangle^{\frac{1}{2}}_{E_d^{\mathbf{a}}(\mathscr{C}_\infty)} \\ \Longrightarrow &\left\|\nabla w_{\widetilde{\Gamma}^x_{K,R,i}}\right\|_{L^2(E_d^{\mathbf{a}}(\mathscr{C}_\infty))} \leqslant \Lambda \left\|\widetilde{\Gamma}^x_{K,R,i}\right\|_{L^2(E_d^{\mathbf{a}}(\mathscr{C}_\infty))} \leqslant \Lambda \left\|\Gamma^x_{K,R}\right\|_{L^2(E_d^{\mathbf{a}}(\mathscr{C}_\infty))} \leqslant C^2_{K,R}R^{-\frac{d}{2}}.\end{aligned}$$

Combining eq. (3.75) and the estimate on $\Theta(e)$, one may want to argue that

$$\sum_{e\in E_d}|B_e|^2(x) \leqslant \sum_{i=1}^{d}\sum_{e\in E_d^{\mathbf{a}}(\mathscr{C}_\infty)} \mathcal{O}'_s\left(|\nabla w_{\widetilde{\Gamma}^x_{K,R,i}}|^2(e)\right) \leqslant \mathcal{O}'_s(C^2_{K,R}R^{-d}).$$

However, this argument is not correct since eq. (3.23) does not work for our case where $w_{\widetilde{\Gamma}^x_{K,R,i}}$ is stochastic. A rigorous proof needs an argument as in [83, Lemma 3.7] using the minimal scale: We construct a collection of good cubes $\mathcal{G}'$ such that not only Meyers' inequality [83, Proposition 3.6] is established, but also there exists $\varepsilon(d,\mathfrak{p},\Lambda) > 0$ and $C(d,\mathfrak{p},\Lambda) < \infty$ for all $\square \in \mathcal{G}'$

$$\frac{1}{|\square|}\left(\int_{E_d^{\mathbf{a}}(\mathscr{C}_\infty\cap\square)} \Theta^{\frac{2(2+\varepsilon)}{\varepsilon}}(e)\right)^{\frac{\varepsilon}{2+\varepsilon}} < C(d,\mathfrak{p},\Lambda). \tag{3.76}$$

Then we do the Calderón-Zygmund decomposition Proposition 3.2.1 for $\mathcal{G}'$ to obtain a partition of cubes $\mathcal{U}$, and apply Meyers' inequality for eq. (3.75)

$$\begin{aligned}\sum_{e\in E_d}|B_e|^2(x) &\leqslant \sum_{i=1}^{d}\sum_{\square\in\mathcal{U}}\sum_{e\in E_d^{\mathbf{a}}(\mathscr{C}_\infty\cap\square)} |\nabla w_{\widetilde{\Gamma}^x_{K,R,i}}|^2(e)\Theta^2(e)\mathbf{1}_{\{\mathscr{C}_\infty=\mathscr{C}_\infty^\varepsilon\}} \\ &\leqslant \sum_{i=1}^{d}\sum_{\square\in\mathcal{U}}|\square| \underbrace{\left(\frac{1}{|\square|}\sum_{e\in E_d^{\mathbf{a}}(\mathscr{C}_\infty\cap\square)} |\nabla w_{\widetilde{\Gamma}^x_{K,R,i}}|^{2+\varepsilon}(e)\right)^{\frac{2}{2+\varepsilon}}}_{\text{Applying Meyers' inequality}} \underbrace{\left(\frac{1}{|\square|}\sum_{e\in E_d^{\mathbf{a}}(\mathscr{C}_\infty\cap\square)} \Theta^{\frac{2(2+\varepsilon)}{\varepsilon}}(e)\right)^{\frac{\varepsilon}{2+\varepsilon}}}_{\leqslant C \text{ after eq. (3.76)}} \mathbf{1}_{\{\mathscr{C}_\infty=\mathscr{C}_\infty^\varepsilon\}} \\ &\leqslant C\sum_{i=1}^{d}\sum_{\square\in\mathcal{U}}|\square| \left(\frac{1}{|\frac{4}{3}\square|}\sum_{e\in E_d^{\mathbf{a}}(\mathscr{C}_\infty\cap\frac{4}{3}\square)} |\nabla w_{\widetilde{\Gamma}^x_{K,R,i}}|^2(e) + \left(\frac{1}{|\frac{4}{3}\square|}\sum_{e\in E_d^{\mathbf{a}}(\mathscr{C}_\infty\cap\frac{4}{3}\square)} |\widetilde{\Gamma}^x_{K,R,i}|^{2+\varepsilon}(e)\right)^{\frac{2}{2+\varepsilon}}\right).\end{aligned}$$

The first term can be controlled by the $\dot{H}^1$ estimate for $w_{\widetilde{\Gamma}^x_{K,R,i}}$ that

$$\left\|\nabla w_{\widetilde{\Gamma}^x_{K,R,i}}\right\|^2_{L^2(E_d^{\mathbf{a}}(\mathscr{C}_\infty))} \leqslant \Lambda^2 \left\|\Gamma^x_{K,R}\right\|^2_{L^2(E_d^{\mathbf{a}}(\mathscr{C}_\infty))} \leqslant C^2_K R^{-d}.$$

While for the second term, we can now apply eq. (3.23) as $\widetilde{\Gamma}^x_{K,R,i}$ is deterministic

$$\sum_{i=1}^{d}\sum_{\square\in\mathcal{U}}|\square|\left(\frac{1}{|\frac{4}{3}\square|}\sum_{e\in E^{\mathbf{a}}_d(\mathscr{C}_\infty\cap\frac{4}{3}\square)}|\widetilde{\Gamma}^x_{K,R,i}|^{2+\varepsilon}(e)\right)^{\frac{2}{2+\varepsilon}}$$
$$\leqslant\sum_{i=1}^{d}\sum_{\square\in\mathcal{U}}|\square|^{\frac{\varepsilon}{2+\varepsilon}}\sum_{e\in E^{\mathbf{a}}_d(\mathscr{C}_\infty\cap\frac{4}{3}\square)}|\widetilde{\Gamma}^x_{K,R,i}|^2(e)\leqslant\mathcal{O}'_s(C^2_{K,R}R^{-d}).$$

This concludes the proof. □

3.3.2 Construction of the flux conrrectors

In this part, we prove a Helmholtz-Hodge type decomposition for $\mathbf{g}_p$, which is another quantity $\mathbf{S}_p$ used in the further quantification of algorithm. We recall that we use $\mathbf{g}_{p,i}$ to represent the i-th component of the vector field $\mathbf{g}_p:\mathbb{Z}^d\to\mathbb{R}^d$ and the standard heat kernel is defined as $\Phi_R(x):=\frac{1}{(4\pi R^2)^{d/2}}\exp\left(-\frac{x^2}{4R^2}\right)$.

Proposition 3.3.4. *For each $p\in\mathbb{R}^d$, almost surely there exists a vector field $\mathbf{S}_p:\mathbb{Z}^d\to\mathbb{R}^{d\times d}$ called flux corrector of $\mathbf{g}_p$, which takes values in the set of anti-symmetric matrices (that is, $\mathbf{S}_{p,ij}=-\mathbf{S}_{p,ji}$) and satisfying the following equations:*

$$\begin{cases}\mathcal{D}^*\cdot\mathbf{S}_p &= \mathbf{g}_p,\\ -\Delta\mathbf{S}_{p,ij} &= \mathcal{D}_{e_j}\mathbf{g}_{p,i}-\mathcal{D}_{e_i}\mathbf{g}_{p,j},\end{cases}\tag{3.77}$$

where the first equation means that for every $i\in\{1,2\cdots d\}$, $\sum_{j=1}^d\mathcal{D}^*_{e_j}\mathbf{S}_{p,ij}=\mathbf{g}_{p,i}$.

The quantity satisfies similar estimation as eq. (3.45) and eq. (3.46): there exist two positive constants $s:=s(d,\mathfrak{p},\Lambda)$, $C:=C(d,\mathfrak{p},\Lambda,s)$ such that

$$\forall 1\leqslant i,j\leqslant d,\quad \forall x\in\mathbb{Z}^d,\quad |\mathcal{D}_{e_k}\mathbf{S}_{p,ij}|(x)\leqslant\mathcal{O}_s(C|p|),\tag{3.78}$$

and for the heat kernel Φ_R, we have

$$|\Phi_R\star[\mathcal{D}_{e_k}\mathbf{S}_{p,ij}]|(x)\leqslant\mathcal{O}_s(C|p|R^{-\frac{d}{2}}).\tag{3.79}$$

Heuristic analysis

The following discussion gives a little heuristic analysis before a rigorous proof. In fact, if we define a field $H_p:\mathbb{Z}^d\to\mathbb{R}^d$ such that

$$-\Delta H_{p,i}=\mathbf{g}_{p,i},\tag{3.80}$$

where $-\Delta:=-\nabla\cdot\nabla=\mathcal{D}^*\cdot\mathcal{D}$ is the discrete Laplace and then we define $\mathbf{S}_p$ such that

$$\mathbf{S}_{p,ij}=\mathcal{D}_{e_j}H_{p,i}-\mathcal{D}_{e_i}H_{p,j}.\tag{3.81}$$

We see that this definition gives us a solution of eq. (3.77) since

$$-\Delta\mathbf{S}_{p,ij}=-\Delta\left(\mathcal{D}_{e_j}H_{p,i}-\mathcal{D}_{e_i}H_{p,j}\right)=\mathcal{D}_{e_j}(-\Delta H_{p,i})-\mathcal{D}_{e_i}(-\Delta H_{p,j})=\mathcal{D}_{e_j}\mathbf{g}_{p,i}-\mathcal{D}_{e_i}\mathbf{g}_{p,j}.$$
$$(\mathcal{D}^*\cdot\mathbf{S}_p)_i=\sum_{j=1}^d\mathcal{D}^*_{e_j}\left(\mathcal{D}_{e_j}H_{p,i}-\mathcal{D}_{e_i}H_{p,j}\right)=-\Delta H_{p,i}-\mathcal{D}_{e_i}\sum_{j=1}^d\mathcal{D}^*_{e_j}H_{p,j}=\mathbf{g}_{p,i}.$$

Here we use one property that $H_{p,i} = (-\Delta)^{-1}\mathbf{g}_{p,i}$ so that H_p is also divergence free. This idea works on periodic homogenization problem [147, Lemma 3.1], but in our context, one key problem is to well define eq. (3.80). In the present work, we apply an elementary probabilistic approach: Let $(S_k)_{k\geqslant 0}$ defines a lazy discrete time simple random walk on $\mathbb{Z}^d$ with probability $\frac{1}{2}$ to stay unmoved and $\frac{1}{4d}$ to move towards one of the nearest neighbors on $\mathbb{Z}^d$, and we use $(P_t)_{t\in\mathbb{N}}$ to define its semigroup, with the notation

$$\begin{aligned} P_t(x,y) &:= P_t(y-x) = \mathbb{P}[S_t = y-x] \\ (P_t f)(x) &:= \sum_{y\in\mathbb{Z}^d} P_t(x,y)f(y) = ([P_t]\star[f])(x), \qquad \forall f\in L^1(\mathbb{Z}^d), \end{aligned} \tag{3.82}$$

where $[\cdot]$ denotes a constant extension on every $z+\left(-\frac{1}{2},\frac{1}{2}\right)^d$. Using the representation of the solution of harmonic function by a simple random walk

$$H_{p,i}(x) = \frac{1}{4d}\sum_{t=0}^{\infty}(P_t\mathbf{g}_{p,i})(x),$$

and we deduce from the definition of $\mathbf{S}_e$ in eq. (3.81)

$$\mathbf{S}_{p,ij}(x) = \frac{1}{4d}\sum_{t=0}^{\infty}\mathcal{D}_{e_j}(P_t\mathbf{g}_{p,i})(x) - \mathcal{D}_{e_i}(P_t\mathbf{g}_{p,j})(x).$$

If we believe that P_t is close to the heat kernel that $P_t(x,y)\simeq \frac{1}{(\pi t)^{d/2}}\exp\left(-\frac{|y-x|^2}{t}\right)$, and that the operator $\mathcal{D}$ helps to gain another factor of $t^{-\frac{1}{2}}$, then Proposition 3.1.2 would give us that $|\mathcal{D}_{e_j}(P_t\mathbf{g}_{p,i})|(x)\leqslant \mathcal{O}_s(t^{-\frac{1}{2}-\frac{d}{4}})$. We expect that this upper bound is sharp in general, and the fact that $\sum_{t=1}^{\infty}t^{-\frac{1}{2}-\frac{d}{4}} = \infty$ prevents us from being able to define $\mathbf{S}_{p,ij}$ directly in dimension $d=2$. Nevertheless we can make sense of

$$(\mathcal{D}_{e_k}\mathbf{S}_{p,ij})(x) = \frac{1}{4d}\sum_{t=0}^{\infty}\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t\mathbf{g}_{p,i})(x) - \mathcal{D}_{e_k}\mathcal{D}_{e_i}(P_t\mathbf{g}_{p,j})(x), \tag{3.83}$$

because differentiating P_t a second time will allow us to gain an extra factor of $t^{-\frac{1}{2}}$, and thus give us that that $|\mathcal{D}_{e_k}\mathcal{D}_{e_j}P_t\mathbf{g}_{p,i}|\leqslant \mathcal{O}_s(Ct^{-1-\frac{d}{4}})$. Then we can apply eq. (3.23) to say that $\mathcal{D}_{e_k}\mathbf{S}_{p,ij}$ is well-defined and prove other properties.

Rigorous construction of $\mathcal{D}\mathbf{S}$

Proof of Proposition 3.3.4. We will give a rigorous proof that eq. (3.83) gives a well-defined anti-symmetric valued vector field $\mathbf{S}_e$. The proof can be divided into three steps.

Step 1: Stochastic integrability of $\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t\mathbf{g}_{p,i})$. In the first step, we prove that eq. (3.83) makes sense, that is the part $\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t\mathbf{g}_{p,i})(x) - \mathcal{D}_{e_k}\mathcal{D}_{e_i}(P_t\mathbf{g}_{p,j})(x)$ is summable. In the heuristic analysis, we compare P_t with the heat kernel, which can be reformulated carefully by the local central limit theorem.

Lemma 3.3.6 (Page 61, Exercise 2.10 of [170])**.** *We denote by $\bar{P}_t(x,y) = \frac{1}{(\pi t)^{d/2}}\exp\left(-\frac{|y-x|^2}{t}\right)$, then there exists a positive constant $C(d)$, such that for all $t>0$,*

$$\sup_{x\in\mathbb{Z}^d}\left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}P_t - \mathcal{D}_{e_k}\mathcal{D}_{e_j}\bar{P}_t\right|(x)\leqslant C(d)t^{-\frac{d+3}{2}}. \tag{3.84}$$

Proof. The proof follows the idea in [170, Theorem 2.3.5] and also relies on [170, Lemmas 2.3.3 and 2.3.4] where we have

$$P_t(x) = \bar{P}_t(x) + V_t(x,r) + \frac{1}{(2\pi)^d t^{\frac{d}{2}}} \int_{|\theta|\leqslant r} e^{-\frac{ix\cdot\theta}{\sqrt{t}}} e^{-\frac{|\theta|^2}{4}} F_t(\theta)\, d\theta,$$

and there exits $\zeta > 0$ such that for every $0 < \theta, r \leqslant t^{\frac{1}{8}}$, $V_t(x,r), F_t(\theta)$ satisfy

$$|F_t|(\theta) \leqslant \frac{|\theta|^4}{t}, \qquad |V_t(x,r)| \leqslant c(d) t^{-\frac{d}{2}} e^{-\zeta r^2}.$$

We apply $\mathcal{D}_{e_k}\mathcal{D}_{e_j}$ with respect to x and obtain that

$$\left|\mathcal{D}_{e_k}\mathcal{D}_{e_j} P_t - \mathcal{D}_{e_k}\mathcal{D}_{e_j}\bar{P}_t\right|(x) = \left|\mathcal{D}_{e_k}\mathcal{D}_{e_j} V_t(x,r) + \frac{1}{(2\pi)^d t^{\frac{d}{2}}} \int_{|\theta|\leqslant r} \mathcal{D}_{e_k}\mathcal{D}_{e_j} e^{-\frac{ix\cdot\theta}{\sqrt{t}}} e^{-\frac{|\theta|^2}{4}} F_t(\theta)\, d\theta\right|.$$

We take $r = t^{\frac{1}{8}}$, then the term $V_t(x,r)$ has an error of exponential type

$$\left|V_t(x, t^{\frac{1}{8}})\right| \leqslant c(d) t^{-\frac{d}{2}} e^{-\zeta t^{\frac{1}{4}}} \leqslant c'(d) t^{-\frac{d+3}{2}}.$$

So we focus on another part, by a simple finite difference calculus we have$\mathcal{D}_{e_k}\mathcal{D}_{e_j} e^{-\frac{ix\cdot\theta}{\sqrt{t}}} \leqslant \frac{2\theta}{\sqrt{t}}$. Moreover, we apply $|F_t|(\theta) \leqslant \frac{|\theta|^4}{t}$ and have

$$\left|\frac{1}{(2\pi)^d t^{\frac{d}{2}}} \int_{|\theta|\leqslant r} \mathcal{D}_{e_k}\mathcal{D}_{e_j} e^{-\frac{ix\cdot\theta}{\sqrt{t}}} e^{-\frac{|\theta|^2}{4}} F_t(\theta)\, d\theta\right| \leqslant \left|\frac{1}{(2\pi)^d t^{\frac{d+3}{2}}} \int_{|\theta|\leqslant r} |\theta|^5 e^{-\frac{|\theta|^2}{4}}\, d\theta\right| \leqslant C t^{-\frac{d+3}{2}}.$$

This concludes the proof. □

We prove that eq. (3.83) is well defined by showing that

$$\mathbb{P}\text{-a.s} \qquad \forall 1 \leqslant i,j,k \leqslant d, \forall x \in \mathbb{Z}^d, \quad \sum_{t=0}^{\infty} \left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t \mathbf{g}_{p,i})\right|(x) < \infty. \tag{3.85}$$

We break this term into two

$$\sum_{t=0}^{\infty} \left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t \mathbf{g}_{p,i})\right|(x) = \underbrace{\sum_{t=0}^{\infty} \left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}(\bar{P}_t \mathbf{g}_{p,i})\right|(x)}_{\text{eq. (3.86)-a}} + \underbrace{\sum_{t=0}^{\infty} \left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t - \bar{P}_t)\mathbf{g}_{p,i}\right|(x)}_{\text{eq. (3.86)-b}}. \tag{3.86}$$

$\bar{P}_t$ is better than P_t since it is a standard heat kernel and we can do explicit calculation. We observe that

$$\forall t \geqslant 1, \forall y \in \mathbb{Z}^d, \left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}\bar{P}_t\right|(y) \leqslant \frac{C(d)}{t}\bar{P}_{2t}(y) = \frac{C(d)}{t(2\pi t)^{d/2}} \exp\left(-\frac{|y|^2}{2t}\right) \leqslant \frac{C(d)}{t^{\frac{d+2}{2}}\left(|\frac{y}{\sqrt{t}}| \vee 1\right)^{\frac{d+1}{2}}}.$$

Then $K_{\sqrt{t}} := \mathcal{D}_{e_k}\mathcal{D}_{e_j}\bar{P}_t$ is a kernel described in Proposition 3.1.2 with the constant $C_{K,\sqrt{t}} := \frac{C(d)}{t}$, so we have

$$\left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}(\bar{P}\mathbf{g}_{p,i})\right|(x) = \left|K_{\sqrt{t}} \star [\mathbf{g}_{p,i}]\right|(x) \leqslant \mathcal{O}_s(C t^{-1-\frac{d}{4}}).$$

We put these estimates with Proposition 3.3.1 in the eq. (3.86)-a and get

$$\text{eq. (3.86)-a} \leqslant |\mathbf{g}_{p,i}|(x) + \sum_{t=1}^{\infty} \left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t \mathbf{g}_{p,i})\right|(x) \leqslant \mathcal{O}_s(C) + \sum_{t=1}^{\infty} \mathcal{O}_s(Ct^{-1-\frac{d}{4}}) \leqslant \mathcal{O}_s(C).$$

On the other hand, to handle eq. (3.86)-b, we choose $\varepsilon > 0$ and study at first

$$\begin{aligned}
\left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t - \bar{P}_t)\mathbf{g}_{p,i}\right|(x) &\leqslant \left| \int_{|y|\leqslant t^{\frac{1}{2}+\varepsilon}} \left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t - \bar{P}_t)\right](y) \left[\mathbf{g}_{p,i}\right](x-y)\, dy \right| \\
&\quad + \left| 4 \int_{|y|\geqslant t^{\frac{1}{2}+\varepsilon}} \left([P_t] + [\bar{P}_t]\right)(y) \left[\mathbf{g}_{p,i}\right](x-y)\, dy \right| \\
&\overset{\text{Lemma 3.3.6}}{\leqslant} \int_{|y|\leqslant t^{\frac{1}{2}+\varepsilon}} \mathcal{O}_s\left(\frac{C}{t^{\frac{d+3}{2}}}\right) dy + 4 \int_{|y|\geqslant t^{\frac{1}{2}+\varepsilon}} \left([P_t] + [\bar{P}_t]\right)(y) \mathcal{O}_s(C)\, dy \\
&\overset{\text{eq. (3.23)}}{\leqslant} \mathcal{O}_s\left(Ct^{-(\frac{3}{2}-d\varepsilon)}\right) + \mathcal{O}_s\left(\int_{|y|\geqslant t^{\frac{1}{2}+\varepsilon}} \left([P_t] + [\bar{P}_t]\right)(y)\, dy\right).
\end{aligned}$$

We divide the estimation into two terms since Lemma 3.3.6 is uniform but not optimal for the tail probability, which is of type sub-Gaussian so that the mass outside $|t|^{\frac{1}{2}+\varepsilon}$ should be very small. By direct calculation, we have that $\int_{|y|\geqslant t^{\frac{1}{2}+\varepsilon}} \left[\bar{P}_t\right](y)\, dy \leqslant C(d)e^{-t^{2\varepsilon}}$ and by Hoeffding's inequality for the lazy simple random walk $(S_t)_{t\geqslant 0}$

$$\int_{|y|\geqslant t^{\frac{1}{2}+\varepsilon}} [P_t](y)\, dy = \mathbb{P}\left[|S_t| \geqslant t^{\frac{1}{2}+\varepsilon}\right] \leqslant 2\exp\left(-\frac{2t^{1+2\varepsilon}}{\mathbb{V}\mathrm{ar}[S_t]}\right) \leqslant 2e^{-4t^{2\varepsilon}}.$$

Combining these tail event estimates and by choosing $\varepsilon = \frac{1}{4d}$, we obtain that

$$\left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t - \bar{P}_t)\mathbf{g}_{p,i}\right|(x) \leqslant \mathcal{O}_s(Ct^{-\frac{5}{4}}),$$

and this concludes that eq. (3.86)-b $\leqslant \mathcal{O}_s(C)$, so eq. (3.85) holds, eq. (3.78) holds and that $\mathcal{D}_{e_k}\mathbf{S}_{p,ij}$ is well defined.

Remark. In the proof, we also obtained one quantitative estimate of the following type: There exist two constants $s := s(d, \mathfrak{p}, \Lambda)$, $C := C(d, \mathfrak{p}, \Lambda, s)$ such that for every random field $X : \mathbb{Z}^d \to \mathbb{R}$ satisfying for every $z \in \mathbb{Z}^d, |X(z)| \leqslant \mathcal{O}_s(\theta)$, we have

$$\forall 1 \leqslant i,j,k \leqslant d, x \in \mathbb{Z}^d, \left|\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t - \bar{P}_t)X\right|(x) \leqslant \mathcal{O}_s(C\theta t^{-\frac{5}{4}}). \tag{3.87}$$

By a similar approach with the classical local central limit theorem [170, Theorem 2.3.5], we can also prove that

$$\forall x \in \mathbb{Z}^d, |P_t \mathbf{g}_{p,i}|(x) \leqslant \mathcal{O}_s(Ct^{-\frac{1}{4}}). \tag{3.88}$$

Step 2: Verification of eq. (3.77). The verification of eq. (3.77) is direct thanks to eq. (3.85). We will also use the semigroup property that

$$P_t(x) - P_{t-1}(x) = \frac{1}{4d}\Delta P_{t-1}(x). \tag{3.89}$$

$$\begin{aligned}
(\mathcal{D}^* \cdot \mathbf{S}_p)_i(x) &= \frac{1}{4d}\sum_{t=0}^{\infty}\sum_{j=1}^{d} \mathcal{D}^*_{e_j}\mathcal{D}_{e_j}P_t\mathbf{g}_{p,i}(x) - \mathcal{D}^*_{e_j}\mathcal{D}_{e_i}P_t\mathbf{g}_{p,j}(x) \\
&= \frac{1}{4d}\sum_{t=0}^{\infty}(\underbrace{-\Delta P_t}_{\text{eq. (3.89)}})\mathbf{g}_{p,i}(x) - \mathcal{D}_{e_i}P_t\underbrace{(\mathcal{D}^* \cdot \mathbf{g}_p)}_{=0}(x) \\
&= \sum_{t=0}^{\infty}(P_t - P_{t+1})\mathbf{g}_{p,i}(x) \\
&= \mathbf{g}_{p,i}(x).
\end{aligned}$$

In the last step, we use implicitly that $\lim_{t\to\infty} P_t \mathbf{g}_{p,i}(x) = 0$ almost surely. This is true by Borel-Cantelli lemma and the estimation $|P_t \mathbf{g}_{p,i}|(x) \leqslant \mathcal{O}_s(Ct^{-\frac{1}{4}})$ (see eq. (3.88)):

$$\sum_{t=1}^{\infty} \mathbb{P}[|P_t \mathbf{g}_{p,i}|(x) \geqslant \varepsilon] \leqslant \sum_{t=1}^{\infty} \exp\left(-\left(C\varepsilon t^{\frac{1}{4}}\right)^s\right) < \infty.$$

The second part of eq. (3.77), concerning $-\Delta \mathbf{S}_{p,ij}$, is easy to verify by a similar calculation

$$\begin{aligned}
-\Delta \mathbf{S}_{p,ij}(x) &= \frac{1}{4d}\sum_{t=0}^{\infty} \mathcal{D}^*_{e_k}\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t \mathbf{g}_{p,i})(x) - \mathcal{D}^*_{e_k}\mathcal{D}_{e_k}\mathcal{D}_{e_i}(P_t \mathbf{g}_{p,j})(x) \\
&= \frac{1}{4d}\sum_{t=0}^{\infty} (-\Delta P_t \mathcal{D}_{e_j}\mathbf{g}_{p,i})(x) - (-\Delta P_t \mathcal{D}_{e_i}\mathbf{g}_{p,j})(x) \\
&= (\mathcal{D}_{e_j}\mathbf{g}_{p,i})(x) - (\mathcal{D}_{e_i}\mathbf{g}_{p,j})(x).
\end{aligned}$$

Finally, by the definition, we can define $\mathbf{S}_{p,ij}$ just by integration of $\mathcal{D}\mathbf{S}_{p,ij}$ along a path. This construction does not depend on the choice of path since $\mathcal{D}\mathbf{S}_{p,ij}$ is a potential field.

Step 3: Estimation of $|\Phi_R \star [\mathcal{D}_{e_k}\mathbf{S}_{p,ij}]|(x)$. This is a result of the convolution. Thanks to the eq. (3.85), we can apply Fubini lemma to $|\Phi_R \star [\mathcal{D}_{e_k}\mathbf{S}_{p,ij}]|(x)$ that

$$\begin{aligned}
|\Phi_R \star [\mathcal{D}_{e_k}\mathbf{S}_{p,ij}]|(x) &= \left|\frac{1}{4d}\sum_{t=0}^{\infty} \Phi_R \star \left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}P_t\right] \star [\mathbf{g}_{p,i}] - \Phi_R \star \left[\mathcal{D}_{e_k}\mathcal{D}_{e_i}P_t\right] \star [\mathbf{g}_{p,j}]\right|(x) \\
&= \left|\frac{1}{4d}\sum_{t=0}^{\infty} \left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}P_t\right] \star (\Phi_R \star [\mathbf{g}_{p,i}]) - \left[\mathcal{D}_{e_k}\mathcal{D}_{e_i}P_t\right] \star (\Phi_R \star [\mathbf{g}_{p,j}])\right|(x)
\end{aligned}$$

The main idea is that $\Phi_R \star [\mathbf{g}_{p,j}] \leqslant \mathcal{O}_s\left(CR^{-\frac{d}{2}}\right)$ by Proposition 3.1.2 , then we repeat the main argument of stochastic integrability of $\mathcal{D}_{e_k}\mathbf{S}_{p,ij}$ to get a better estimate. We focus on just one term:

$$\begin{aligned}
\left|\frac{1}{4d}\sum_{t=0}^{\infty} \left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}P_t\right] \star (\Phi_R \star [\mathbf{g}_{p,i}])\right| \leqslant& \underbrace{\left|\frac{1}{4d}\sum_{t=0}^{\infty} \left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}P_t - \mathcal{D}_{e_k}\mathcal{D}_{e_j}\bar{P}_t\right] \star (\Phi_R \star [\mathbf{g}_{p,i}])\right|}_{\text{eq. (3.90)-a}} \\
&+ \underbrace{\left|\frac{1}{4d}\sum_{t=0}^{\infty} \left(\left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}\bar{P}_t\right] - \mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}}}\right) \star (\Phi_R \star [\mathbf{g}_{p,i}])\right|}_{\text{eq. (3.90)-b}} \\
&+ \underbrace{\left|\frac{1}{4d}\sum_{t=0}^{\infty} \mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}}} \star (\Phi_R \star [\mathbf{g}_{p,i}])\right|}_{\text{eq. (3.90)-c}}.
\end{aligned} \tag{3.90}$$

We treat the three terms one by one. For eq. (3.90)-a, we apply eq. (3.87) with $X := \Phi_R \star [\mathbf{g}_{p,i}]$ and we use also eq. (3.23)

$$\begin{aligned}
\text{eq. (3.90)-a} \leqslant& \frac{1}{d}|\Phi_R \star [\mathbf{g}_{p,i}]|(x) + \frac{1}{4d}\sum_{t=1}^{\infty} \left|\left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}(P_t - \bar{P}_t)\right] \star (\Phi_R \star [\mathbf{g}_{p,i}])\right|(x) \\
\leqslant& \mathcal{O}_s(CR^{-\frac{d}{2}}) + \sum_{t=1}^{\infty} \mathcal{O}_s(Ct^{-\frac{5}{4}}R^{-\frac{d}{2}}) \\
\leqslant& \mathcal{O}_s(CR^{-\frac{d}{2}}).
\end{aligned} \tag{3.91}$$

For the term eq. (3.90)-b, we observe that for every $y \in \mathbb{R}^d$,

$$\left|\left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}\bar{P}_t\right](y) - \mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}}}(y)\right| = \left|\left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}}}\right](y) - \mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}}}(y)\right| \leqslant \frac{C(d)}{t^{\frac{3}{2}}}\Phi_{\sqrt{t}}(y).$$

We apply this estimate and use eq. (3.23) to obtain that

$$\begin{aligned}
&\left|\frac{1}{4d}\sum_{t=0}^{\infty}\left(\left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}\bar{P}_t\right] - \mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}}}\right) \star (\Phi_R \star [\mathbf{g}_{p,i}])\right|(x) \\
\leqslant &\frac{1}{4d}\sum_{t=0}^{\infty}\left|\left[\mathcal{D}_{e_k}\mathcal{D}_{e_j}\bar{P}_t\right] - \mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}}}\right| \star |\Phi_R \star [\mathbf{g}_{p,i}]|(x) \\
\leqslant &\frac{1}{d}|\Phi_R \star [\mathbf{g}_{p,i}]|(x) + \frac{1}{4d}\sum_{t=1}^{\infty}\frac{C(d)}{t^{\frac{3}{2}}}\Phi_{\sqrt{t}} \star |\Phi_R \star [\mathbf{g}_{p,i}]|(x) \\
\leqslant &\mathcal{O}_s(CR^{-\frac{d}{2}}) + \sum_{t=1}^{\infty}\mathcal{O}_s(Ct^{-\frac{3}{2}}R^{-\frac{d}{2}}) \\
\leqslant &\mathcal{O}_s(CR^{-\frac{d}{2}}).
\end{aligned} \tag{3.92}$$

For the last term $\left|\frac{1}{4d}\sum_{t=0}^{\infty}\mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}}} \star (\Phi_R \star [\mathbf{g}_{p,i}])\right|(x)$, we use the property of semi-group, the linearity of the finite difference operator and we apply Proposition 3.1.2 to the kernel $\mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}}}$

$$\begin{aligned}
\left|\frac{1}{4d}\sum_{t=0}^{\infty}\mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}}} \star (\Phi_R \star [\mathbf{g}_{p,i}])\right|(x) &= \left|\frac{1}{4d}\sum_{t=0}^{\infty}\mathcal{D}_{e_k}\mathcal{D}_{e_j}\left(\Phi_{\sqrt{\frac{t}{2}+R^2}} \star [\mathbf{g}_{p,i}]\right)\right|(x) \\
&= \left|\frac{1}{4d}\sum_{t=0}^{\infty}\left(\mathcal{D}_{e_k}\mathcal{D}_{e_j}\Phi_{\sqrt{\frac{t}{2}+R^2}}\right) \star [\mathbf{g}_{p,i}]\right|(x) \\
&\leqslant \sum_{t=0}^{\infty}\mathcal{O}_s\left(C\left(\frac{t}{2}+R^2\right)^{-1-\frac{d}{2}}\right) \\
&\leqslant \mathcal{O}_s(CR^{-\frac{d}{2}}).
\end{aligned} \tag{3.93}$$

This concludes the proof as we put the three estimates eq. (3.91),eq. (3.92) and eq. (3.93) in eq. (3.90) and eq. (3.79). □

L^q, L^∞ estimate of $\mathbf{S}_p$

Once we establish the spatial average estimate for $\mathcal{D}\mathbf{S}_p$, we also have its L^q and L^∞ estimate.

Proposition 3.3.5. *There exist three finite positive constants $s := s(d,\mathfrak{p},\Lambda), k := k(d,\mathfrak{p},\Lambda)$ and $C := C(d,\mathfrak{p},\Lambda,s)$ such that for each $1 \leqslant i,j \leqslant d, p \in \mathbb{R}^d$ and $q \in [1,\infty)$,*

$$\left(R^{-d}\int_{B_R}|\mathbf{S}_{p,ij} - (\mathbf{S}_{p,ij})_{B_R}|^q\right)^{\frac{1}{q}} \leqslant \begin{cases} \mathcal{O}_s(C|p|q^k\log^{\frac{1}{2}}(R)) & d=2, \\ \mathcal{O}_s(C|p|q^k) & d=3, \end{cases} \tag{3.94}$$

and for each $x,y \in \mathbb{Z}^d$,

$$|\mathbf{S}_{p,ij}(x) - \mathbf{S}_{p,ij}(y)| \leqslant \begin{cases} \mathcal{O}_s(C|p|\log^{\frac{1}{2}}|x-y|) & d=2, \\ \mathcal{O}_s(C|p|) & d=3. \end{cases} \tag{3.95}$$

Proof. Similar to [83, Theorems 1 and 2], these estimates are the results of local estimate and spatial average estimates proved in eq. (3.49), eq. (3.19), eq. (3.78) and eq. (3.79) by applying a heat kernel type multi-scale Poincaré's inequality. We refer to [83, Sections 4 and 5]. □

3.4 Two-scale expansion on the cluster

In this part, we prove Theorem 3.1.2 which is the heart of all the analysis of our algorithm as stated in Section 3.1.3. Here we prove a more detailed version of the theorem.

Proposition 3.4.1 (Two-scale expansion on percolation)**.** *Under the same context of Theorem 3.1.2, there exist three random variables* $\mathcal{X}, \mathcal{Y}_1, \mathcal{Y}_2$ *satisfying*

$$\mathcal{X} \leqslant \mathcal{O}_1(C(d,\mathfrak{p},\Lambda)m), \qquad \mathcal{Y}_1 \leqslant \mathcal{O}_s\left(C(d,\mathfrak{p},\Lambda,s)\ell(\lambda)m^{\frac{1}{s}}\right), \qquad \mathcal{Y}_2 \leqslant \mathcal{O}_s\left(C(d,\mathfrak{p},\Lambda,s)\lambda^{\frac{d}{2}}m^{\frac{1}{s}}\right),$$

and we have the estimate

$$\begin{aligned}\left\|\nabla(w-v)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} \leqslant C(d,\Lambda)\Big(&\|\mathcal{D}\bar{v}\|_{L^2(\square_m)}\left(3^{-\frac{m}{2}}\ell^{\frac{1}{2}}(\lambda)\mathcal{X}^d + 3^{-\frac{m}{2}}\ell^{-\frac{1}{2}}(\lambda)\mathcal{Y}_1\mathcal{X}^d + \mu\mathcal{Y}_1 + \mathcal{Y}_2\mathcal{X}^d\right)\\ &+ \|\mathcal{D}\bar{v}\|^{\frac{1}{2}}_{L^2(\square_m)}\|\mathcal{D}^*\mathcal{D}\bar{v}\|^{\frac{1}{2}}_{L^2(\mathrm{int}(\square_m))}\left(\ell^{\frac{1}{2}}(\lambda)\mathcal{X}^d + \ell^{-\frac{1}{2}}(\lambda)\mathcal{Y}_1\mathcal{X}^d\right)\\ &+ \|\mathcal{D}^*\mathcal{D}\bar{v}\|_{L^2(\mathrm{int}(\square_m))}\mathcal{Y}_1\mathcal{X}^d\Big).\end{aligned}$$

3.4.1 Main part of the proof

The main idea of the proof is to use the quantities $\{\phi_{e_k}\}_{k=1,\ldots,d}$ and $\{\mathbf{S}_{e_k,ij}\}_{i,j,k=1,\ldots,d}$ analyzed in previous work and in Section 3.3, under the condition $\square_m \in \mathcal{P}_*$. We do some simple manipulations at first. Throughout the proof, we use the notation $h := v - w$.

Proof. Step 1: Setting up. We define a modified coarsened function $\widetilde{h}$

$$\widetilde{h}(x) = \begin{cases} h(x) & x \in \mathscr{C}_*(\square_m),\\ [h]_{\mathcal{P}}(x) & x \in \square_m \backslash \mathscr{C}_*(\square_m), \mathrm{dist}(\square_{\mathcal{P}}(x), \partial\square_m) \geqslant 1,\\ 0 & x \in \square_m \backslash \mathscr{C}_*(\square_m), \mathrm{dist}(\square_{\mathcal{P}}(x), \partial\square_m) = 0.\end{cases} \tag{3.96}$$

We put it as a test function in eq. (3.15)

$$\left\langle \widetilde{h}, (\mu^2_{\mathscr{C},m} - \nabla\cdot\mathbf{a}_{\mathscr{C},m}\nabla)v\right\rangle_{\mathrm{int}(\square_m)} = \left\langle \widetilde{h}, (\mu^2_{\mathscr{C},m} - \nabla\cdot\bar{\mathbf{a}}\nabla)\bar{v}\right\rangle_{\mathrm{int}(\square_m)}.$$

Since $\widetilde{h} \in C_0(\square_m)$, we can apply the formula eq. (3.28) and get

$$\left\langle \mu_{\mathscr{C},m}\widetilde{h}, \mu_{\mathscr{C},m}v\right\rangle_{\square_m} + \left\langle \nabla\widetilde{h}, \mathbf{a}_{\mathscr{C},m}\nabla v\right\rangle_{\square_m} = \left\langle \mu_{\mathscr{C},m}\widetilde{h}, \mu_{\mathscr{C},m}\bar{v}\right\rangle_{\square_m} + \left\langle \nabla\widetilde{h}, \bar{\mathbf{a}}\nabla\bar{v}\right\rangle_{\square_m}. \tag{3.97}$$

We subtract a term of w on the two sides to get

$$\begin{aligned}&\left\langle \mu_{\mathscr{C},m}\widetilde{h}, \mu_{\mathscr{C},m}(v-w)\right\rangle_{\square_m} + \left\langle \nabla\widetilde{h}, \mathbf{a}_{\mathscr{C},m}\nabla(v-w)\right\rangle_{\square_m}\\ &\qquad = \left\langle \mu_{\mathscr{C},m}\widetilde{h}, \mu_{\mathscr{C},m}(\bar{v}-w)\right\rangle_{\square_m} + \left\langle \nabla\widetilde{h}, \bar{\mathbf{a}}\nabla\bar{v} - \mathbf{a}_{\mathscr{C},m}\nabla w\right\rangle_{\square_m}.\end{aligned}$$

We put $v - w = h$ into the identity and obtain that

$$\begin{aligned}&\left\langle \mu_{\mathscr{C},m}\widetilde{h}, \mu_{\mathscr{C},m}h\right\rangle_{\square_m} + \left\langle \nabla\widetilde{h}, \mathbf{a}_{\mathscr{C},m}\nabla h\right\rangle_{\square_m}\\ &\qquad = \left\langle \mu_{\mathscr{C},m}\widetilde{h}, \mu_{\mathscr{C},m}(\bar{v}-w)\right\rangle_{\square_m} + \left\langle \nabla\widetilde{h}, \bar{\mathbf{a}}\nabla\bar{v} - \mathbf{a}_{\mathscr{C},m}\nabla w\right\rangle_{\square_m}.\end{aligned} \tag{3.98}$$

Step 2: Restriction tricks. There are three observations:

- Observation 1. The effect of $\mu_{\mathscr{C},m}$ restricts the inner product to $\mathscr{C}_*(\square_m)$, and on $\mathscr{C}_*(\square_m)$ we have $\widetilde{h} = h$ by eq. (3.96). Thus we have

$$\left\langle \mu_{\mathscr{C},m}\widetilde{h}, \mu_{\mathscr{C},m}h\right\rangle_{\square_m} = \mu^2\left\langle h, h\right\rangle_{\mathscr{C}_*(\square_m)}, \qquad \left\langle \mu_{\mathscr{C},m}\widetilde{h}, \mu_{\mathscr{C},m}(\bar{v}-w)\right\rangle_{\square_m} = \mu^2\left\langle h, \bar{v}-w\right\rangle_{\mathscr{C}_*(\square_m)}.$$

- Observation 2. The definition of $\mathbf{a}_{\mathscr{C},m}$ also restricts the inner product on $E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))$ and we have
$$\left\langle \nabla \widetilde{h}, \mathbf{a}_{\mathscr{C},m} \nabla h \right\rangle_{\square_m} = \langle \nabla h, \mathbf{a} \nabla h \rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))},$$
as $\mathbf{a}_{\mathscr{C},m} = 0$ outside $E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))$ by eq. (3.39).

- Observation 3. This step is the key where we gain much in the estimate and where we use the condition $\square_m \in \mathcal{P}_*$. We apply the formula eq. (3.27) to $\left\langle \nabla \widetilde{h}, \bar{\mathbf{a}} \nabla \bar{v} - \mathbf{a}_{\mathscr{C},m} \nabla w \right\rangle_{\square_m}$ to obtain that
$$\begin{aligned}\left\langle \nabla \widetilde{h}, \bar{\mathbf{a}} \nabla \bar{v} - \mathbf{a}_{\mathscr{C},m} \nabla w \right\rangle_{\square_m} &= \left\langle \widetilde{h}, -\nabla \cdot (\bar{\mathbf{a}} \nabla \bar{v} - \mathbf{a}_{\mathscr{C},m} \nabla w) \right\rangle_{\mathrm{int}(\square_m)} \\ &= \left\langle \widetilde{h}, \mathcal{D}^* \cdot (\bar{\mathbf{a}} \mathcal{D} \bar{v} - \mathbf{a}_{\mathscr{C},m} \mathcal{D} w) \right\rangle_{\mathrm{int}(\square_m)} \\ &= \left\langle \widetilde{h}, \mathcal{D}^* \cdot (\bar{\mathbf{a}} \mathcal{D} \bar{v} - \mathbf{a}_{\mathscr{C}} \mathcal{D} w) \right\rangle_{\mathrm{int}(\square_m)} \\ &\quad + \left\langle \widetilde{h}, \mathcal{D}^* \cdot (\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C},m}) \mathcal{D} w \right\rangle_{\mathrm{int}(\square_m)}.\end{aligned}$$
We use the condition $\square_m \in \mathcal{P}_*$, which implies that $\mathscr{C}_*(\square_m) \subseteq \mathscr{C}_\infty$ and
$$\operatorname{supp}\left(\mathcal{D}^* \cdot (\mathbf{a}_{\mathscr{C}} - \mathbf{a}_{\mathscr{C},m}) \mathcal{D} w\right) \subseteq (\mathscr{C}_\infty \cap \square_m) \backslash \mathscr{C}_*(\square_m).$$
In Definition 3.B.1 and Lemma 3.B.2, we prove that $(\mathscr{C}_\infty \cap \square_m) \backslash \mathscr{C}_*(\square_m)$ is the union of small clusters contained in the partition cubes $\square_{\mathcal{P}}$ with distance 1 to $\partial \square_m$, where $\widetilde{h}$ equals 0. Therefore, we obtain that
$$\left\langle \widetilde{h}, \mathcal{D}^* \cdot (\bar{\mathbf{a}} \mathcal{D} \bar{v} - \mathbf{a}_{\mathscr{C},m} \mathcal{D} w) \right\rangle_{\mathrm{int}(\square_m)} = \left\langle \widetilde{h}, \mathcal{D}^* \cdot (\bar{\mathbf{a}} \mathcal{D} \bar{v} - \mathbf{a}_{\mathscr{C}} \mathcal{D} w) \right\rangle_{\mathrm{int}(\square_m)}.$$
Using an identity
$$\mathcal{D}^* \cdot (\mathbf{a}_{\mathscr{C}} \mathcal{D} w - \bar{\mathbf{a}} \mathcal{D} \bar{v}) = \mathcal{D}^* \cdot \mathbf{F} \text{ on } \mathbb{Z}^d,$$
which will be proved later in Lemma 3.4.1 and $\mathbf{F}$ is a vector field $\mathbf{F} : \mathbb{Z}^d \to \mathbb{R}^d$, we conclude
$$\begin{aligned}\left\langle \nabla \widetilde{h}, \bar{\mathbf{a}} \nabla \bar{v} - \mathbf{a}_{\mathscr{C},m} \nabla w \right\rangle_{\square_m} \mathbf{1}_{\{\square_m \in \mathcal{P}_*\}} &= \left\langle \widetilde{h}, -\mathcal{D}^* \cdot \mathbf{F} \right\rangle_{\mathrm{int}(\square_m)} \mathbf{1}_{\{\square_m \in \mathcal{P}_*\}} \\ &= -\left\langle \mathcal{D} \widetilde{h}, \mathbf{F} \right\rangle_{\square_m} \mathbf{1}_{\{\square_m \in \mathcal{P}_*\}}.\end{aligned}$$

Combining all these observations, we transform eq. (3.98) to
$$\begin{aligned}\left(\mu^2 \langle h, h \rangle_{\mathscr{C}_*(\square_m)} + \langle \nabla h, \mathbf{a} \nabla h \rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))}\right) \mathbf{1}_{\{\square_m \in \mathcal{P}_*\}} \\ = \left(\mu^2 \langle h, \bar{v} - w \rangle_{\mathscr{C}_*(\square_m)} - \left\langle \mathcal{D} \widetilde{h}, \mathbf{F} \right\rangle_{\square_m}\right) \mathbf{1}_{\{\square_m \in \mathcal{P}_*\}}.\end{aligned} \tag{3.99}$$

Using Hölder's inequality and Young's inequality, we obtain that
$$\langle \nabla h, \mathbf{a} \nabla h \rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))} \mathbf{1}_{\{\square_m \in \mathcal{P}_*\}} \leqslant \left(\frac{\mu^2}{4} \|\bar{v} - w\|^2_{L^2(\mathscr{C}_*(\square_m))} + \left\|\mathcal{D} \widetilde{h}\right\|_{L^2(\square_m)} \|\mathbf{F}\|_{L^2(\square_m)}\right) \mathbf{1}_{\{\square_m \in \mathcal{P}_*\}}. \tag{3.100}$$

Step 3: Study of $\left\|\mathcal{D} \widetilde{h}\right\|_{L^2(\square_m)}$. The next step is to estimate the size of $\left\|\mathcal{D} \widetilde{h}\right\|_{L^2(\square_m)}$. Since $\widetilde{h} \in C_0(\square_m)$, we have that $\left\|\mathcal{D} \widetilde{h}\right\|_{L^2(\square_m)} = \left\|\nabla \widetilde{h}\right\|_{L^2(\square_m)}$. We use the function $[h]_{\mathcal{P},\square_m}$ defined in eq. (3.42)
$$[h]_{\mathcal{P},\square_m}(x) = \begin{cases} [h]_{\mathcal{P}}(x) & \operatorname{dist}(\square_{\mathcal{P}}(x), \partial \square_m) \geqslant 1, \\ 0 & \operatorname{dist}(\square_{\mathcal{P}}(x), \partial \square_m) = 0. \end{cases} \tag{3.101}$$

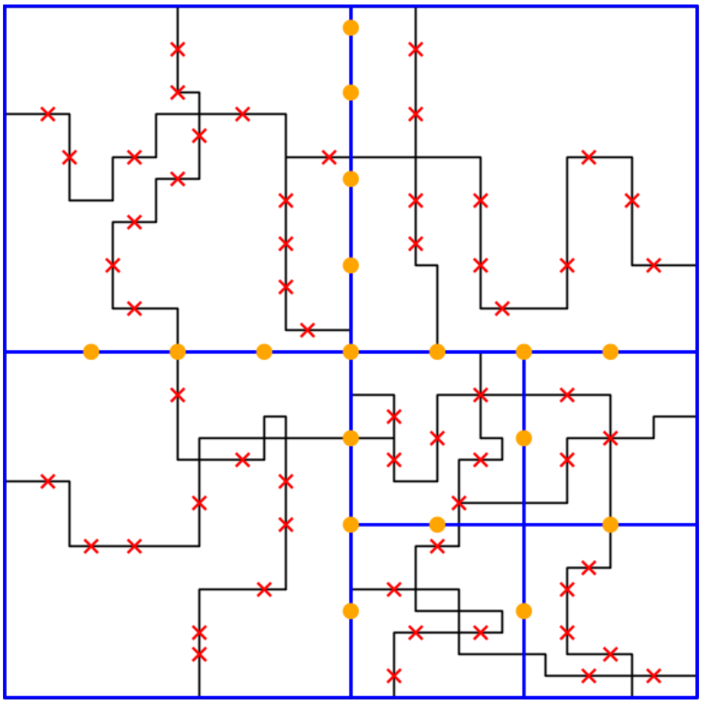

Figure 3.4: The figure shows the sources contributing to $\|\nabla\widetilde{h}\|_{L^2(\square_m)}$. The black segments represent the cluster $\mathscr{C}_*(\square_m)$ while the blue segments represent the partition of good cubes. Using the coarsened function, we see that the quantity can be controlled by the sum of three terms: the difference between $\nabla\widetilde{h}$ and $[h]_{\mathcal{P},\square_m}$ near the cluster $\mathscr{C}_*(\square_m)$, marked with red cross in the image; the gradient $[h]_{\mathcal{P},\square_m}$ at the interface of different partition cubes $\square_{\mathcal{P}}$, marked with orange disk.

as a function to do comparison and apply eq. (3.29) (see Figure 3.4 for the errors from the two terms)

$$\begin{aligned}
\|\nabla\widetilde{h}\|_{L^2(\square_m)} &= \left\|\nabla(\widetilde{h} - [h]_{\mathcal{P},\square_m})\right\|_{L^2(\square_m)} + \left\|\nabla\,[h]_{\mathcal{P},\square_m}\right\|_{L^2(\square_m)} \\
&\leqslant 2d\left\|\widetilde{h} - [h]_{\mathcal{P},\square_m}\right\|_{L^2(\square_m)} + \left\|\nabla\,[h]_{\mathcal{P},\square_m}\right\|_{L^2(\square_m)} \\
&\leqslant 2d\left\|h - [h]_{\mathcal{P},\square_m}\right\|_{L^2(\mathrm{int}(\square_m)\cap\mathscr{C}_*(\square_m))} + \left\|\nabla\,[h]_{\mathcal{P},\square_m}\right\|_{L^2(\square_m)}
\end{aligned}$$

The last step is correct since $\widetilde{h}$ and $[h]_{\mathcal{P},\square_m}$ coincide at the boundary and also on the part $\mathrm{int}(\square_m)\backslash\mathscr{C}_*(\square_m)$. We define

$$\mathcal{X} := \max_{x\in\square_m} \mathrm{size}(\square_{\mathcal{P}}(x)), \tag{3.102}$$

which can be estimated using eq. (3.25) and eq. (3.76) as

$$\mathcal{X} \leqslant \mathcal{O}_1(C(d,\mathfrak{p},\Lambda)m), \tag{3.103}$$

and apply Proposition 3.2.2

$$\begin{aligned}\left\|\nabla\widetilde{h}\right\|_{L^2(\square_m)} &\leqslant 2d\left\|h-[h]_{\mathcal{P},\square_m}\right\|_{L^2(\mathrm{int}(\square_m)\cap\mathscr{C}_*(\square_m))}+\left\|\nabla[h]_{\mathcal{P},\square_m}\right\|_{L^2(\square_m)}\\ &\leqslant C\left(\sum_{\{x,y\}\in E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))}\mathrm{size}(\square_{\mathcal{P}}(x))^{2d}|\nabla h|^2(x,y)\right)^{\frac{1}{2}}\\ &\leqslant C\mathcal{X}^d\left\|\nabla h\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))}.\end{aligned}$$

We put it back to eq. (3.100)

$$\begin{aligned}&\langle\nabla h,\mathbf{a}\nabla h\rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))}\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}\\ &\qquad\leqslant\left(\frac{\mu^2}{4}\|\bar{v}-w\|^2_{L^2(\mathscr{C}_*(\square_m))}+C\mathcal{X}^d\left\|\nabla h\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))}\|\mathbf{F}\|_{L^2(\square_m)}\right)\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}.\end{aligned}$$

and use Young's inequality, finally we get

$$\left\|\nabla h\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))}\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}\leqslant C(d)\Lambda\left(\mu\|\bar{v}-w\|_{L^2(\mathscr{C}_*(\square_m))}+\mathcal{X}^d\|\mathbf{F}\|_{L^2(\square_m)}\right)\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}, \tag{3.104}$$

Step 4: Quantification. It remains to estimate two quantities $\|\bar{v}-w\|_{L^2(\mathscr{C}_*(\square_m))}\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}$ and $\|\mathbf{F}\|_{L^2(\square_m)}\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}$.

The two random variables used in the estimation are defined as

$$\begin{aligned}\mathcal{Y}_1 &:= \max_{1\leqslant i,j,k\leqslant d,\mathrm{dist}(x,\square_m)\leqslant 1}\left|\phi^{(\lambda)}_{e_k}(x)\right|+\left|\mathbf{S}^{(\lambda)}_{e_k,ij}(x)\right|,\\ \mathcal{Y}_2 &:= \max_{1\leqslant i,j,k\leqslant d,\mathrm{dist}(x,\square_m)\leqslant 1}\left|\Phi_{\lambda^{-1}}\star\mathcal{D}_{e_i}[\phi_{e_k}]^{\eta}_{\mathcal{P}}(x)\right|+\left|\Phi_{\lambda^{-1}}\star[\mathcal{D}^*_{e_j}\mathbf{S}_{e_k,ij}](x)\right|,\end{aligned} \tag{3.105}$$

where they involved the spatial average of corrector and flux corrector, the modified corrector defined in eq. (3.14) and modified flux corrector defined in eq. (3.113). They have estimates following eq. (3.25), eq. (3.46), eq. (3.79) and also Proposition 3.3.5, Proposition 3.2.4 that there exists $0<s(d,\mathfrak{p},\Lambda)<\infty$ and $0<C(d,\mathfrak{p},\Lambda)<\infty$ such that

$$\mathcal{Y}_1\leqslant\mathcal{O}_s\left(C(d,\mathfrak{p},\Lambda,s)\ell(\lambda)m^{\frac{1}{s}}\right)\qquad\mathcal{Y}_2\leqslant\mathcal{O}_s\left(C(d,\mathfrak{p},\Lambda,s)\lambda^{\frac{d}{2}}m^{\frac{1}{s}}\right). \tag{3.106}$$

For $\|\bar{v}-w\|_{L^2(\mathscr{C}_*(\square_m))}\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}$, we have

$$\begin{aligned}\|w-\bar{v}\|^2_{L^2(\square_m)}\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}} &= \left\|\sum_{j=1}^d(\Upsilon\mathcal{D}_{e_j}\bar{v})\phi^{(\lambda)}_{e_j}\right\|^2_{L^2(\square_m)}\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}\\ &\leqslant d\underbrace{\left(\max_{1\leqslant j\leqslant d,x\in\square_m}\phi^{(\lambda)}_{e_j}(x)\right)^2}_{\leqslant\mathcal{Y}_1^2}\sum_{j=1}^d\sum_{x\in\square_m}\left((\Upsilon\mathcal{D}_{e_j}\bar{v})(x)\right)^2\\ &\leqslant d\mathcal{Y}_1^2\|\mathcal{D}\bar{v}\|^2_{L^2(\square_m)}.\end{aligned} \tag{3.107}$$

For $\|\mathbf{F}\|_{L^2(\square_m)}\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}$, we use the formula eq. (3.114) that

$\|\mathbf{F}\|^2_{L^2(\square_m)}\leqslant C(d)$ (eq. (3.108)-a + eq. (3.108)-b + eq. (3.108)-c + eq. (3.108)-d + eq. (3.108)-e) , where the five terms are respectively

$$\begin{aligned}
\text{eq. (3.108)-a} &= \sum_{i=1}^{d} \left\|(1-\Upsilon)\left(\mathbf{a}_{\mathscr{C}} - \bar{\mathbf{a}}\right)\left(\mathcal{D}_{e_i}\bar{v}\right)\right\|_{L^2(\square_m)}^2, \\
\text{eq. (3.108)-b} &= \sum_{i,k=1}^{d} \left\|\phi_{e_k}^{(\lambda)}(\cdot+e_k)\mathbf{a}_{\mathscr{C}}\mathcal{D}_{e_i}\left(\Upsilon\mathcal{D}_{e_k}\bar{v}\right)\right\|_{L^2(\square_m)}^2, \\
\text{eq. (3.108)-c} &= \sum_{i,j,k=1}^{d} \left\|\mathbf{S}_{e_k,ij}^{(\lambda)}(\cdot-e_j)\mathcal{D}_{e_j}^{*}\left(\Upsilon\mathcal{D}_{e_k}\bar{v}\right)\right\|_{L^2(\square_m)}^2, \\
\text{eq. (3.108)-d} &= \sum_{i,j,k=1}^{d} \left\|\mathcal{D}_{e_j}^{*}\left(\left[\mathbf{S}_{e_k,ij}\right]\star\Phi_{\lambda^{-1}}\right)\left(\Upsilon\mathcal{D}_{e_k}\bar{v}\right)\right\|_{L^2(\square_m)}^2, \\
\text{eq. (3.108)-e} &= \sum_{i,k=1}^{d} \left\|\mathbf{a}_{\mathscr{C}}\mathcal{D}_{e_i}\left(\left[\phi_{e_k}\right]_{\mathcal{P}}^{\eta}\star\Phi_{\lambda^{-1}}\right)\left(\Upsilon\mathcal{D}_{e_k}\bar{v}\right)\right\|_{L^2(\square_m)}^2.
\end{aligned} \tag{3.108}$$

We treat them term by term. For *eq.* (3.108)-a, noticing that $(1-\Upsilon) \leqslant \mathbf{1}_{\{\mathrm{dist}(\cdot,\partial)\leqslant 2\ell(\lambda)\}}$, we apply the trace formula eq. (3.33)

$$\begin{aligned}
\text{eq. (3.108)-a} &= \sum_{i=1}^{d} \left\|(1-\Upsilon)\left(\mathbf{a}_{\mathscr{C}} - \bar{\mathbf{a}}\right)\left(\mathcal{D}_{e_i}\bar{v}\right)\right\|_{L^2(\square_m)}^2 \\
&\leqslant 2\sum_{i=1}^{d} \left\|\left(\mathcal{D}_{e_i}\bar{v}\right)\mathbf{1}_{\{\mathrm{dist}(\cdot,\partial)\leqslant 2\ell(\lambda)\}}\right\|_{L^2(\square_m)}^2 \\
&\leqslant C(d)\ell(\lambda)\left(3^{-m}\left\|\mathcal{D}\bar{v}\right\|_{L^2(\square_m)}^2 + \left\|\mathcal{D}\bar{v}\right\|_{L^2(\square_m)}\left\|\mathcal{D}^{*}\mathcal{D}\bar{v}\right\|_{L^2(\mathrm{int}(\square_m))}\right).
\end{aligned} \tag{3.109}$$

For the term eq. (3.108)-b, we notice that

$$\mathcal{D}_{e_i}\left(\Upsilon\mathcal{D}_{e_k}\bar{v}\right) = (\mathcal{D}_{e_i}\Upsilon)(\mathcal{D}_{e_k}\bar{v}) + \Upsilon(\cdot+e_i)(\mathcal{D}_{e_i}\mathcal{D}_{e_k}\bar{v}),$$

and the support of $\mathcal{D}_{e_i}\Upsilon$ is contained in the region of distance between $\ell(\lambda)$ and $2\ell(\lambda)$ from $\partial\square_m$ i.e.

$$\mathcal{D}_{e_i}\Upsilon \leqslant \frac{1}{\ell(\lambda)}\mathbf{1}_{\{\cdot\in\square_m,\frac{1}{2}\ell(\lambda)\leqslant\mathrm{dist}(\cdot,\partial)\leqslant 3\ell(\lambda)\}},$$

then we apply these in eq. (3.108)-b and also eq. (3.33) and obtain that

$$\begin{aligned}
&\sum_{i,k=1}^{d} \left\|\phi_{e_k}^{(\lambda)}(\cdot+e_k)\mathbf{a}_{\mathscr{C}}\mathcal{D}_{e_i}\left(\Upsilon\mathcal{D}_{e_k}\bar{v}\right)\right\|_{L^2(\square_m)}^2 \\
&\leqslant \sum_{i,k=1}^{d} \left\|\underbrace{\phi_{e_k}^{(\lambda)}(\cdot+e_k)}_{\leqslant\mathcal{Y}_1}(\mathcal{D}_{e_i}\Upsilon)(\mathcal{D}_{e_k}\bar{v})\right\|_{L^2(\square_m)}^2 + \sum_{i,k=1}^{d} \left\|\underbrace{\phi_{e_k}^{(\lambda)}(\cdot+e_k)}_{\leqslant\mathcal{Y}_1}\Upsilon(\cdot+e_i)(\mathcal{D}_{e_i}\mathcal{D}_{e_k}\bar{v})\right\|_{L^2(\square_m)}^2 \\
&\leqslant \sum_{i,k=1}^{d} \mathcal{Y}_1^2\left\|\frac{1}{\ell(\lambda)}\mathbf{1}_{\{\frac{1}{2}\ell(\lambda)\leqslant\mathrm{dist}(\cdot,\partial)\leqslant 3\ell(\lambda)\}}(\mathcal{D}_{e_k}\bar{v})\right\|_{L^2(\square_m)}^2 + \mathcal{Y}_1^2\sum_{i,k=1}^{d}\left\|\mathbf{1}_{\{\mathrm{dist}(\cdot,\partial\square_m)\geqslant\ell(\lambda)\}}(\mathcal{D}_{e_i}\mathcal{D}_{e_k}\bar{v})\right\|_{L^2(\square_m)}^2 \\
&\leqslant C(d)\mathcal{Y}_1^2\left(\frac{1}{\ell(\lambda)}\left(3^{-m}\left\|\mathcal{D}\bar{v}\right\|_{L^2(\square_m)}^2 + \left\|\mathcal{D}\bar{v}\right\|_{L^2(\square_m)}\left\|\mathcal{D}^{*}\mathcal{D}\bar{v}\right\|_{L^2(\mathrm{int}(\square_m))}\right) + \left\|\mathcal{D}^{*}\mathcal{D}\bar{v}\right\|_{L^2(\mathrm{int}(\square_m))}^2\right).
\end{aligned} \tag{3.110}$$

In the last step, we apply eq. (3.33) and use the interior H^2 norm of $\bar{v}$ since the function Υ is supported just in the interior with distance $\ell(\lambda)$ from $\partial\square_m$. eq. (3.108)-c follows the similar estimate.

For the term eq. (3.108)-d, we use the quantity $\mathcal{Y}_2$ to estimate it

$$\begin{aligned}\text{eq. (3.108)-d} &= \sum_{i,j,k=1}^{d} \Bigg\| \underbrace{\mathcal{D}_{e_j}^{*}\left([\mathbf{S}_{e_k,ij}] \star \Phi_{\lambda^{-1}}\right)}_{\leqslant \mathcal{Y}_2} (\Upsilon \mathcal{D}_{e_k} \bar{v}) \Bigg\|_{L^2(\square_m)}^{2} \\ &\leqslant C(d) \mathcal{Y}_2^2 \left\| \mathcal{D}\bar{v} \right\|_{L^2(\square_m)}^2 .\end{aligned} \tag{3.111}$$

The term eq. (3.108)-e follows the similar estimate.

We combine eq. (3.109),eq. (3.110) and eq. (3.111) together and obtain that

$$\begin{aligned}\|\mathbf{F}\|_{L^2(\square_m)} \leqslant C(d) \Big(\|\mathcal{D}\bar{v}\|_{L^2(\square_m)} \Big(3^{-\frac{m}{2}} \ell^{\frac{1}{2}}(\lambda) + 3^{-\frac{m}{2}} \ell^{-\frac{1}{2}}(\lambda) \mathcal{Y}_1 + \mathcal{Y}_2 \Big) \\ \|\mathcal{D}\bar{v}\|_{L^2(\square_m)}^{\frac{1}{2}} \|\mathcal{D}^*\mathcal{D}\bar{v}\|_{L^2(\mathrm{int}(\square_m))}^{\frac{1}{2}} \Big(\ell^{\frac{1}{2}}(\lambda) + \ell^{-\frac{1}{2}}(\lambda)\mathcal{Y}_1 \Big) + \|\mathcal{D}^*\mathcal{D}\bar{v}\|_{L^2(\mathrm{int}(\square_m))} \mathcal{Y}_1 \Big).\end{aligned} \tag{3.112}$$

We put the two estimates eq. (3.107) and eq. (3.108) into eq. (3.104) and get the desired result. □

3.4.2 Construction of a vector field

In this part we calculate the vector field **F** used in the last paragraph. We define at first the *modified flux corrector* $\{\mathbf{S}^{(\lambda)}_{e_k,ij}\}_{1 \leqslant i,j,k \leqslant d}$ similar to eq. (3.14) that

$$\mathbf{S}^{(\lambda)}_{e_k,ij} := \mathbf{S}_{e_k,ij} - [\mathbf{S}_{e_k,ij}] \star \Phi_{\lambda^{-1}} . \tag{3.113}$$

Lemma 3.4.1. *There exists a vector field* $\mathbf{F} : \mathbb{Z}^d \to \mathbb{R}^d$ *such that* $\mathcal{D}^* \cdot (\mathbf{a}_{\mathscr{C}} \mathcal{D}w - \bar{\mathbf{a}}\mathcal{D}\bar{v}) = \mathcal{D}^* \cdot \mathbf{F}$, *with the formula*

$$\begin{aligned}\mathbf{F}_i = &(1-\Upsilon)(\mathbf{a}_{\mathscr{C}} - \bar{\mathbf{a}})(\mathcal{D}_{e_i}\bar{v}) + \sum_{k=1}^{d} \phi_{e_k}^{(\lambda)}(\cdot + e_k) \mathbf{a}_{\mathscr{C}} \mathcal{D}_{e_i}(\Upsilon \mathcal{D}_{e_k}\bar{v}) \\ &- \sum_{j,k=1}^{d} \mathbf{S}^{(\lambda)}_{e_k,ij}(\cdot - e_j) \mathcal{D}^*_{e_j}(\Upsilon \mathcal{D}_{e_k}\bar{v}) + \sum_{j,k=1}^{d} \mathcal{D}^*_{e_j}\left([\mathbf{S}_{e_k,ij}] \star \Phi_{\lambda^{-1}}\right)(\Upsilon \mathcal{D}_{e_k}\bar{v}) \\ &- \sum_{k=1}^{d} \mathbf{a}_{\mathscr{C}} \mathcal{D}_{e_i}\left([\phi_{e_k}]^{\eta}_{\mathcal{P}} \star \Phi_{\lambda^{-1}}\right)(\Upsilon \mathcal{D}_{e_k}\bar{v}) .\end{aligned} \tag{3.114}$$

Proof. We write

$$\begin{aligned}[\mathbf{a}_{\mathscr{C}}\mathcal{D}w - \bar{\mathbf{a}}\mathcal{D}\bar{v}]_i(x) &= \left[(\mathbf{a}_{\mathscr{C}} - \bar{\mathbf{a}})\mathcal{D}\bar{v} + \sum_{k=1}^{d} \mathbf{a}_{\mathscr{C}} \underbrace{\mathcal{D}\left((\Upsilon \mathcal{D}_{e_k}\bar{v}) \phi_{e_k}^{(\lambda)} \right)}_{\text{Using } eq. \text{ (3.26)}} \right](x) \\ &= \underbrace{[(1-\Upsilon)(\mathbf{a}_{\mathscr{C}} - \bar{\mathbf{a}})\mathcal{D}\bar{v}]_i(x)}_{\text{eq. (3.115)-a}} + \underbrace{\sum_{k=1}^{d} \phi_{e_k}^{(\lambda)}(x+e_i)\mathbf{a}_{\mathscr{C}}(x, x+e_i)\mathcal{D}_{e_i}(\Upsilon\mathcal{D}_{e_k}\bar{v})(x)}_{\text{eq. (3.115)-b}} \\ &\quad + \sum_{k=1}^{d} \left[\left(\mathbf{a}_{\mathscr{C}}\mathcal{D}\phi_{e_k}^{(\lambda)} + (\mathbf{a}_{\mathscr{C}} - \bar{\mathbf{a}})\mathcal{D}l_{e_k} \right)(\Upsilon\mathcal{D}_{e_k}\bar{v}) \right]_i(x)\end{aligned} \tag{3.115}$$

The terms eq. (3.115)-a and eq. (3.115)-b appear in the eq. (3.114) as the first and second term on the right hand side, so it suffices to treat the remaining terms in eq. (3.115), where we apply the definition of $\mathbf{S}_{e_k}$ eq. (3.77)

$$\begin{aligned}
&\sum_{k=1}^{d} \left[\left(\mathbf{a}_{\mathscr{C}} \mathcal{D}\phi^{(\lambda)}_{e_k} + (\mathbf{a}_{\mathscr{C}} - \bar{\mathbf{a}})\mathcal{D}l_{e_k} \right) (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right]_i (x) \\
&= \sum_{k=1}^{d} \left[\underbrace{\left(\mathbf{a}_{\mathscr{C}} \left(\mathcal{D}\phi_{e_k} + \mathcal{D}l_{e_k} \right) - \bar{\mathbf{a}}\mathcal{D}l_{e_k} \right)}_{=\mathcal{D}^* \cdot \mathbf{S}_{e_k}} (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right]_i (x) - \sum_{k=1}^{d} \left[\mathbf{a}_{\mathscr{C}} \left(\mathcal{D} \left[\phi_{e_k} \right]^{\eta}_{\mathcal{P}} \star \Phi_{\lambda^{-1}} \right) (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right]_i (x) \\
&= \underbrace{\sum_{k=1}^{d} \left[\mathcal{D}^* \cdot \mathbf{S}^{(\lambda)}_{e_k} (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right]_i (x)}_{\text{eq. (3.116)-a}} + \underbrace{\sum_{k=1}^{d} \left[\mathcal{D}^* \cdot \left([\mathbf{S}_{e_k}] \star \Phi_{\lambda^{-1}} \right) (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right]_i (x)}_{\text{eq. (3.116)-b}} \\
&\qquad - \underbrace{\sum_{k=1}^{d} \left[\mathbf{a}_{\mathscr{C}} \left(\mathcal{D} \left[\phi_{e_k} \right]^{\eta}_{\mathcal{P}} \star \Phi_{\lambda^{-1}} \right) (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right]_i (x)}_{\text{eq. (3.116)-c}} .
\end{aligned} \tag{3.116}$$

The terms eq. (3.116)-b and eq. (3.116)-c also appear in the definition of $\mathbf{F}_i$ eq. (3.114) as the forth and fifth term. We study the term eq. (3.116)-a and use the anti-symmetry that $\mathbf{S}_{e_k,ij} = -\mathbf{S}_{e_k,ji}$

$$\begin{aligned}
\mathcal{D}^* \cdot \text{eq. (3.116)-a} &= \sum_{i,k=1}^{d} \mathcal{D}^*_{e_i} \left[\mathcal{D}^* \cdot \mathbf{S}^{(\lambda)}_{e_k} (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right]_i (x) \\
&= \sum_{i,j,k=1}^{d} \mathcal{D}^*_{e_i} \left(\left(\mathcal{D}^*_{e_j} \mathbf{S}^{(\lambda)}_{e_k,ij} \right) (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right) (x) \\
&= \underbrace{\sum_{i,j,k=1}^{d} \mathcal{D}^*_{e_i} \mathcal{D}^*_{e_j} \left(\mathbf{S}^{(\lambda)}_{e_k,ij} (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right) (x)}_{=0 \text{ by anti-symmetry}} - \sum_{i,j,k=1}^{d} \mathcal{D}^*_{e_i} \left(\mathbf{S}^{(\lambda)}_{e_k,ij} (\cdot - e_j) \mathcal{D}^*_{e_j} (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right) (x) \\
&= \sum_{i=1}^{d} \mathcal{D}^*_{e_i} \left(- \sum_{j,k=1}^{d} \left(\mathbf{S}^{(\lambda)}_{e_k,ij} (\cdot - e_j) \mathcal{D}^*_{e_j} (\Upsilon \mathcal{D}_{e_k} \bar{v}) \right) \right) (x).
\end{aligned}$$

This gives the formula in eq. (3.114). □

3.5 Analysis of the algorithm

We are now ready to complete the proof of Theorem 3.1.1, and we start by analyzing our algorithm with standard H^1 and H^2 estimates for $\bar{u}$ in eq. (3.5).

Lemma 3.5.1 (H^1 and H^2 estimates)**.** *In the iteration eq. (3.5) we have the following estimates*

$$\|\nabla \bar{u}\|_{L^2(\square_m)} \leqslant |\bar{\mathbf{a}}|^{-1}(1+\Lambda) \left\| \nabla (u - u_0) \mathbf{1}_{\{\mathbf{a} \neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))}, \tag{3.117}$$

$$\|\mathcal{D}^* \mathcal{D} \bar{u}\|_{L^2(\mathrm{int}(\square_m))} \leqslant C(d,\Lambda) |\bar{\mathbf{a}}|^{-1} \lambda \left\| \nabla (u - u_0) \mathbf{1}_{\{\mathbf{a} \neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))}, \tag{3.118}$$

$$\left\| \nabla (\hat{u} - u) \mathbf{1}_{\{\mathbf{a} \neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))} \leqslant 2 |\bar{\mathbf{a}}|^{-1} (1+\Lambda)^2 \left\| \nabla (u - u_0) \mathbf{1}_{\{\mathbf{a} \neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))}. \tag{3.119}$$

Proof. We start by testing eq. (3.9) and eq. (3.10) with the function u_1, and we also use the trick that $\lambda_{\mathscr{C},m}$ and $\mathbf{a}_{\mathscr{C},m}$ restrict the problem on $(\mathscr{C}_*(\square_m), E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m)))$

$$\begin{aligned}&\left\langle \lambda_{\mathscr{C},m} u_1, \lambda_{\mathscr{C},m} u_1 \right\rangle_{\square_m} + \left\langle \nabla u_1, \mathbf{a}_{\mathscr{C},m} \nabla u_1 \right\rangle_{\square_m} = \left\langle \nabla u_1, \mathbf{a}_{\mathscr{C},m} (u - u_0) \right\rangle_{\square_m}, \\ \Longrightarrow\ & \lambda^2 \|u_1\|^2_{L^2(\mathscr{C}_*(\square_m))} + \Lambda^{-1} \left\| \nabla u_1 \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|^2_{L^2(\mathscr{C}_*(\square_m))} \\ &\qquad \leqslant \left\| \nabla (u - u_0) \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))} \left\| \nabla u_1 \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))}.\end{aligned}$$

We obtain that

$$\lambda \|u_1\|_{L^2(\mathscr{C}_*(\square_m))} \leqslant \Lambda \left\| \nabla (u - u_0) \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))}, \tag{3.120}$$

$$\left\| \nabla u_1 \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))} \leqslant \Lambda \left\| \nabla (u - u_0) \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))}. \tag{3.121}$$

Combining the first equation and the second equation in eq. (3.9) and eq. (3.10), we obtain that

$$-\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} = -\nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla (u - u_0 - u_1) \qquad \text{in } \operatorname{int}(\square_m),$$

then we test it by the function $\bar{u}$ and use Cauchy-Schwarz inequality to obtain that

$$\begin{aligned}\langle \nabla \bar{u}, \bar{\mathbf{a}} \nabla \bar{u} \rangle_{\square_m} &= \left\langle \nabla \bar{u}, \mathbf{a}_{\mathscr{C},m} \nabla (u - u_0 - u_1) \right\rangle_{L^2(\square_m)} \\ &\leqslant \|\nabla \bar{u}\|_{L^2(\square_m)} \left\| \nabla (u - u_0 - u_1) \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{\mathscr{C}_*(\square_m)} \\ \Longrightarrow \|\nabla \bar{u}\|_{L^2(\square_m)} &\leqslant |\bar{\mathbf{a}}|^{-1} \left\| \nabla (u - u_0 - u_1) \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{\mathscr{C}_*(\square_m)}.\end{aligned}$$

Using eq. (3.121) we obtain that

$$\begin{aligned}\|\nabla \bar{u}\|_{L^2(\square_m)} &\leqslant |\bar{\mathbf{a}}|^{-1} \left(\left\| \nabla (u - u_0 - u_1) \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{\mathscr{C}_*(\square_m)} \right) \\ &\leqslant |\bar{\mathbf{a}}|^{-1} \left(\left\| \nabla (u - u_0) \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{\mathscr{C}_*(\square_m)} + \left\| \nabla u_1 \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{\mathscr{C}_*(\square_m)} \right) \\ &\leqslant |\bar{\mathbf{a}}|^{-1} (1 + \Lambda) \left\| \nabla (u - u_0) \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{\mathscr{C}_*(\square_m)}.\end{aligned}$$

This proves the formula eq. (3.117).

Concerning eq. (3.118), we use the estimation of H^2 regularity eq. (3.32) for $-\nabla \cdot \bar{\mathbf{a}} \nabla \bar{u} = \lambda^2_{\mathscr{C},m} u_1$ since $\bar{\mathbf{a}}$ is constant and obtain that

$$\sum_{i,j=1}^{d} \left\| \mathcal{D}^*_{e_i} \mathcal{D}_{e_j} \bar{u} \right\|^2_{L^2(\operatorname{int}(\square_m))} \leqslant C(d) |\bar{\mathbf{a}}|^{-2} \left\| \lambda^2_{\mathscr{C},m} u_1 \right\|^2_{L^2(\square_m)}.$$

We put the result from eq. (3.120) and eq. (3.117) and obtain that

$$\|\mathcal{D}^* \mathcal{D} \bar{u}\|_{L^2(\operatorname{int}(\square_m))} \leqslant C(d, \Lambda) |\bar{\mathbf{a}}|^{-1} \lambda \left\| \nabla (u - u_0) \mathbf{1}_{\{\mathbf{a}\neq 0\}} \right\|_{L^2(\mathscr{C}_*(\square_m))}.$$

To prove eq. (3.119), we put eq. (3.10), the first equation and the second equation of eq. (3.9) into the right hand side of the third equation and obtain that

$$(\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla) u_2 = \lambda^2_{\mathscr{C},m} \bar{u}^2 - \nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla (u - u_0 - u_1) \qquad \text{in } \operatorname{int}(\square_m).$$

We subtract $(\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla) \bar{u}$ on the two sides to obtain

$$(\lambda^2_{\mathscr{C},m} - \nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla)(u_2 - \bar{u}) = -\nabla \cdot \mathbf{a}_{\mathscr{C},m} \nabla (u - u_0 - u_1 - \bar{u}) \qquad \text{in } \operatorname{int}(\square_m),$$

and then we test it by $(u_2 - \bar{u})$ to obtain that

$$\left\|\nabla(u_2 - \bar{u})\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} \leqslant \Lambda \left\|\nabla(u - u_0 - u_1 - \bar{u})\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))}. \tag{3.122}$$

Therefore, combining eq. (3.122), eq. (3.120) and eq. (3.117) we can obtain a trivial bound for our algorithm

$$\begin{aligned}\left\|\nabla(\hat{u} - u)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} &\leqslant \left\|\nabla(u - u_0 - u_1 - \bar{u})\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} + \left\|\nabla(u_2 - \bar{u})\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} \\ &\leqslant 2|\bar{\mathbf{a}}|^{-1}(1+\Lambda)^2 \left\|\nabla(u - u_0)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))}.\end{aligned}$$

□

The trivial bound eq. (3.119) is not optimal. In the typical case $\square_m \in \mathcal{P}_*$ in large scale, we can use Theorem 3.1.2 to help us get a better bound, and this help use conclude the performance of our algorithm.

Proof of Theorem 3.1.1. We analyze the algorithm in two cases: $\square_m \in \mathcal{P}_*$ and $\square_m \notin \mathcal{P}_*$. In the case $\square_m \notin \mathcal{P}_*$, we use eq. (3.119) that

$$\left\|\nabla(\hat{u} - u)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} \mathbf{1}_{\{\square_m \notin \mathcal{P}_*\}} \leqslant 2|\bar{\mathbf{a}}|^{-1}(1+\Lambda)^2 \left\|\nabla(u - u_0)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} \mathbf{1}_{\{\square_m \notin \mathcal{P}_*\}}.$$

In the case $\square_m \in \mathcal{P}_*$, we combine the first equation and the second equation of eq. (3.9) and eq. (3.10), together with the third term they give

$$\begin{aligned} -\nabla\cdot\mathbf{a}_{\mathscr{C},m}\nabla(u - u_0 - u_1) &= -\nabla\cdot\bar{\mathbf{a}}\nabla\bar{u} && \text{in } \operatorname{int}(\square_m), \\ (\lambda^2_{\mathscr{C},m} - \nabla\cdot\mathbf{a}_{\mathscr{C},m}\nabla)u_2 &= (\lambda^2_{\mathscr{C},m} - \nabla\cdot\bar{\mathbf{a}}\nabla)\bar{u} && \text{in } \operatorname{int}(\square_m). \end{aligned}$$

This gives us two equations of two-scale expansion. As in eq. (3.12), we define $w := \bar{u} + \sum_{k=1}^d (\Upsilon\mathcal{D}_{e_k}\bar{u})\phi_{e_k}^{(\lambda)}$ and apply Theorem 3.1.2

$$\begin{aligned}&\left\|\nabla(\hat{u} - u)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} \mathbf{1}_{\{\square_m \in \mathcal{P}_*\}} \\ &\leqslant \left(\left\|\nabla(w - (u - u_0 - u_1))\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} + \left\|\nabla(u_2 - w)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))}\right) \mathbf{1}_{\{\square_m \in \mathcal{P}_*\}}.\end{aligned} \tag{3.123}$$

The last equation gives a bound of type proposition 3.4.1. Together with the Lemma 3.5.1 and the estimate for case $\square_m \notin \mathcal{P}_*$, we obtain that

$$\left\|\nabla(\hat{u} - u)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))} \leqslant \mathcal{Z} \left\|\nabla(u_0 - u)\mathbf{1}_{\{\mathbf{a}\neq 0\}}\right\|_{L^2(\mathscr{C}_*(\square_m))}$$

where $\mathcal{Z}$ is given by

$$\mathcal{Z} = C(d,\Lambda)\left(3^{-\frac{m}{2}}\ell^{-\frac{1}{2}}(\lambda)\mathcal{Y}_1\mathcal{X}^d + \mathcal{Y}_2\mathcal{X}^d + \lambda^{\frac{1}{2}}\ell^{-\frac{1}{2}}(\lambda)\mathcal{Y}_1\mathcal{X}^d + \lambda\mathcal{Y}_1\mathcal{X}^d + \mathbf{1}_{\{\square_m \notin \mathcal{P}_*\}}\right). \tag{3.124}$$

This gives the exact expression of the quantity $\mathcal{Z}$. To conclude, we have to quantify $\mathcal{Z}$ and we use eq. (3.103), eq. (3.106), eq. (3.37) and eq. (3.23) that there exist two positive constants $s(d,\mathfrak{p},\Lambda)$ and $C(d,\mathfrak{p},\Lambda,s)$ such that

$$\mathcal{Z} \leqslant \mathcal{O}_s\left(C(d,\mathfrak{p},\Lambda,s)\left(\left(3^{-\frac{m}{2}} + \lambda + \lambda^{\frac{d}{2}} + \lambda^{\frac{1}{2}}\right)m^{\frac{1}{s}+d}\ell^{\frac{1}{2}}(\lambda) + \lambda m^{\frac{1}{s}}\ell(\lambda) + 3^{-m}\right)\right). \tag{3.125}$$

Observing that $3^{-m} < \lambda$, then the dominating order writes $\mathcal{Z} \leqslant \mathcal{O}_s\left(C\lambda^{\frac{1}{2}}\ell^{\frac{1}{2}}(\lambda)m^{\frac{1}{s}+d}\right)$ and this concludes the proof of Theorem 3.1.1. □

3.6 Numerical experiments

We report on numerical experiments corresponding to our algorithm. In a cube $\square$ of size L, we try to solve a localized corrector problem, that is, we look for the function $\phi_{L,p} \in C_0(\square)$ such that

$$-\nabla \cdot \mathbf{a}\nabla(\phi_{L,p} + l_p) = 0 \quad \text{in } \mathscr{C}_*(\square). \tag{3.126}$$

The quantity $\phi_{L,p}$ is very similar to the corrector ϕ_p and has sublinear growth. This is a good example for illustrating the usefulness of our algorithm, since the homogenized approximation to this function is simply the null function, which is not very informative.

In our example, we take $d = 2$, $p = e_1$ and $L = 243$. We implement the algorithm to get a series of approximated solutions $\hat{u}_n$ where $\hat{u}_0 = 0$. Moreover, we use the residual error to see the convergence

$$\operatorname{res}(\hat{u}_n) := \frac{1}{|\square|} \|-\nabla \cdot \mathbf{a}\nabla(\hat{u}_n + l_p)\|^2_{L^2(\mathscr{C}_*(\square))} = \frac{1}{|\square|} \|-\nabla \cdot \mathbf{a}\nabla(\hat{u}_n - \phi_{L,p})\|^2_{L^2(\mathscr{C}_*(\square))}.$$

See the Figure 3.5 for a simulation of the corrector $\phi_{L,p}$ with high resolution, and Figure 3.6 for its residual errors.

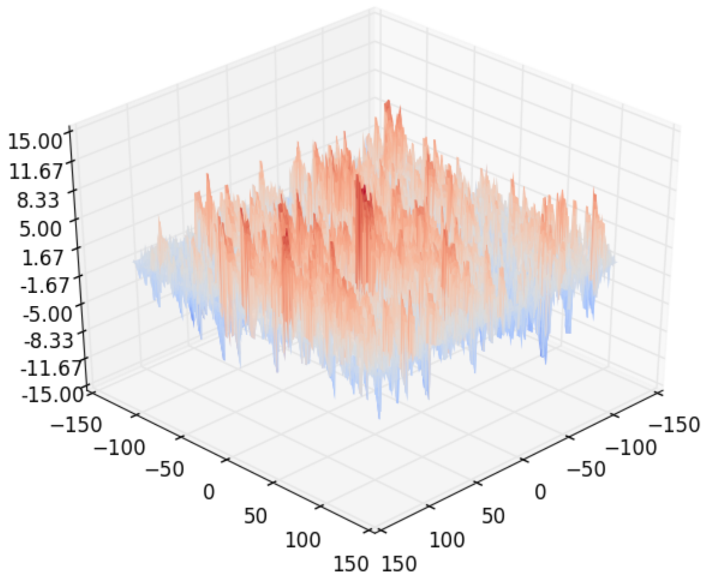

Figure 3.5: A simulation for the corrector on the maximal cluster in a cube 243 × 243.

3.A Proof of some discrete functional inequality

Lemma 3.A.1 (H^2 interior estimate for elliptic equation). *Given two functions $v, f \in C_0(\square_m)$ satisfying the discrete elliptic equation*

$$-\Delta v = f, \qquad \textit{in } \operatorname{int}(\square_m), \tag{3.127}$$

we have an interior estimate

$$\|\mathcal{D}^*\mathcal{D}v\|^2_{L^2(\operatorname{int}(\square_m))} := \sum_{i,j=1}^{d} \left\|\mathcal{D}^*_{e_i}\mathcal{D}_{e_j} v\right\|^2_{L^2(\operatorname{int}(\square_m))} \leqslant d\, \|f\|^2_{L^2(\operatorname{int}(\square_m))}. \tag{3.128}$$

round	errors
1	0.0282597982969
2	0.0126490361046
3	0.00707540548365
4	0.00435201077274
5	0.00282913420116
6	0.00190945842802
7	0.00132483912845
8	0.000939101476657

Figure 3.6: A table of errors $\mathrm{res}(\hat{u}_n)$.

Proof. We extend the elliptic equation to the whole space at first. The function v, f have a natural null extension on $\mathbb{Z}^d$ satisfying

$$\mathcal{D}^* \cdot \mathcal{D}v = f + (\mathcal{D}^* \cdot \mathcal{D}v)\mathbf{1}_{\{\partial \square_m\}}, \qquad \text{in } \mathbb{Z}^d.$$

To simplify the notation, we denote by $\bar{f}$ the term on the right hand side. Then, by one step difference of direction e_j, we have

$$\mathcal{D}^* \cdot \mathcal{D}(\mathcal{D}_{e_j} v(x)) = \mathcal{D}_{e_j} \bar{f}(x).$$

We test this equation with a function ϕ of compact support, then by eq. (3.28) we obtain

$$\left\langle \mathcal{D}\phi, \mathcal{D}(\mathcal{D}_{e_j} v)\right\rangle_{\mathbb{Z}^d} = \left\langle \mathcal{D}^*_{e_j}\phi, \bar{f}\right\rangle_{\mathbb{Z}^d}.$$

Putting $\phi = (\mathcal{D}_{e_j} v)$ in this formula, we obtain that

$$\begin{aligned}\left\langle \mathcal{D}(\mathcal{D}_{e_j} v), \mathcal{D}(\mathcal{D}_{e_j} v)\right\rangle_{\mathbb{Z}^d} &= \left\langle \mathcal{D}^*_{e_j}\mathcal{D}_{e_j} v, \bar{f}\right\rangle_{\mathbb{Z}^d} \\ &= \left\langle \mathcal{D}^*_{e_j}\mathcal{D}_{e_j} v, f\right\rangle_{\mathbb{Z}^d} + \left\langle \mathcal{D}^*_{e_j}\mathcal{D}_{e_j} v, (\mathcal{D}^* \cdot \mathcal{D}v)\mathbf{1}_{\{\partial \square_m\}}\right\rangle_{\mathbb{Z}^d}.\end{aligned}$$

We do the sum over the d canonical directions and get

$$\begin{aligned}\sum_{i,j=1}^{d} \left\langle \mathcal{D}_{e_i}\mathcal{D}_{e_j} v, \mathcal{D}_{e_i}\mathcal{D}_{e_j} v\right\rangle_{\mathbb{Z}^d} &= \sum_{j=1}^{d} \left\langle \mathcal{D}^*_{e_j}\mathcal{D}_{e_j} v, f\right\rangle_{\mathbb{Z}^d} + \sum_{j=1}^{d} \left\langle \mathcal{D}^*_{e_j}\mathcal{D}_{e_j} v, (\mathcal{D}^* \cdot \mathcal{D}v)\mathbf{1}_{\{\partial \square_m\}}\right\rangle_{\mathbb{Z}^d} \\ &= \sum_{j=1}^{d} \left\langle \mathcal{D}^*_{e_j}\mathcal{D}_{e_j} v, f\right\rangle_{\mathbb{Z}^d} + \left\langle \mathcal{D}^* \cdot \mathcal{D}v, (\mathcal{D}^* \cdot \mathcal{D}v)\mathbf{1}_{\{\partial \square_m\}}\right\rangle_{\mathbb{Z}^d}.\end{aligned}$$

Since $\mathcal{D}^*_{e_j} v(x) = -\mathcal{D}_{e_j} v(x - e_j)$, we have

$$\sum_{i,j=1}^{d} \left\langle \mathcal{D}^*_{e_i}\mathcal{D}_{e_j} v, \mathcal{D}^*_{e_i}\mathcal{D}_{e_j} v\right\rangle_{\mathbb{Z}^d} = \sum_{j=1}^{d} \left\langle \mathcal{D}^*_{e_j}\mathcal{D}_{e_j} v, f\right\rangle_{\mathbb{Z}^d} + \left\langle \mathcal{D}^* \cdot \mathcal{D}v, (\mathcal{D}^* \cdot \mathcal{D}v)\mathbf{1}_{\{\partial \square_m\}}\right\rangle_{\mathbb{Z}^d}.$$

There are three observations for this equation.

- $\mathrm{supp}(\mathcal{D}^*_{e_i}\mathcal{D}_{e_j} v) \subseteq \square_m$.

- For any $x \in \partial \square_m$,

$$(\mathcal{D}^* \cdot \mathcal{D}v)^2 = \left(\sum_{j=1}^d \mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v(x) \right)^2 = \sum_{j=1}^d \left(\mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v(x) \right)^2,$$

since $v = 0$ on the boundary and only one term of $\left\{ \mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v(x) \right\}_{j=1\cdots d}$ is not null.

- On the boundary $\partial \square_m$, $f = 0$ since $f \in C_0(\square_m)$.

Combining the three observations, we get

$$\sum_{i,j=1}^d \left\langle \mathcal{D}^*_{e_i} \mathcal{D}_{e_j} v, \mathcal{D}^*_{e_i} \mathcal{D}_{e_j} v \right\rangle_{\square_m} = \sum_{j=1}^d \left\langle \mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v, f \right\rangle_{\mathrm{int}(\square_m)} + \sum_{j=1}^d \left\langle \mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v, \mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v \right\rangle_{\partial \square_m}.$$

Thus, all the terms in the last sum on the right hand side can be found on the left hand side. We use Cauchy-Schwarz inequality and Young's inequality

$$\begin{aligned}
\sum_{i,j=1}^d \left\langle \mathcal{D}^*_{e_i} \mathcal{D}_{e_j} v, \mathcal{D}^*_{e_i} \mathcal{D}_{e_j} v \right\rangle_{\mathrm{int}(\square_m)} &\leqslant \sum_{j=1}^d \left\langle \mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v, f \right\rangle_{\mathrm{int}(\square_m)} \\
&\leqslant \sum_{j=1}^d \left\langle \mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v, \mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v \right\rangle^{\frac{1}{2}}_{\mathrm{int}(\square_m)} \langle f, f \rangle^{\frac{1}{2}}_{\mathrm{int}(\square_m)} \\
&\leqslant \sum_{j=1}^d \left(\frac{1}{2} \left\langle \mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v, \mathcal{D}^*_{e_j} \mathcal{D}_{e_j} v \right\rangle_{\mathrm{int}(\square_m)} + \frac{1}{2} \langle f, f \rangle_{\mathrm{int}(\square_m)} \right) \\
\Longrightarrow \sum_{i,j=1}^d \left\langle \mathcal{D}^*_{e_i} \mathcal{D}_{e_j} v, \mathcal{D}^*_{e_i} \mathcal{D}_{e_j} v \right\rangle_{\mathrm{int}(\square_m)} &\leqslant d \langle f, f \rangle_{\mathrm{int}(\square_m)},
\end{aligned}$$

which concludes the proof. □

The same technique to do an integration along the path helps us to get an estimate of trace.

Lemma 3.A.2 (Trace inequality). *For every $u : \square_m \to \mathbb{R}$ and $0 \leqslant K \leqslant \frac{3^m}{4}$, we have the following inequality*

$$\left\| u \mathbf{1}_{\{\mathrm{dist}(\cdot, \partial \square_m) \leqslant K\}} \right\|^2_{L^2(\square_m)} \leqslant C(d)(K+1) \left(3^{-m} \|u\|^2_{L^2(\square_m)} + \|u\|_{L^2(\square_m)} \|\nabla u\|_{L^2(\square_m)} \right). \quad (3.129)$$

Proof. We use the notation $L_{m,t}$ to define the level set in $\square_m$ with distance t to the boundary

$$L_{m,t} := \{ x \in \square_m : \mathrm{dist}(x, \partial \square_m) = t \}.$$

Then, we observe that $L_{m,0} = \partial \square_m$ and we have the partition

$$\square_m = \bigsqcup_{t=0}^{\left\lfloor \frac{3^m}{2} \right\rfloor} L_{m,t}.$$

Using the pigeonhole principle, it is easy to prove that there exists a $t^* \in \left[0, \left\lfloor \frac{3^m}{4} \right\rfloor - 1 \right]$ such that

$$\|u\|^2_{L^2(L_{m,t^*})} \leqslant \frac{4}{3^m} \|u\|^2_{L^2(\square_m)}, \quad (3.130)$$

and we define $t^* := \arg\min_{t\in[0,\lfloor\frac{3m}{4}\rfloor]} \|u\|^2_{L^2(L_{m,t})}$. We call L_{m,t^*} the *pivot level* and it plays the same role as the null boundary in the proof of Poincaré's inequality. In the following, we will apply the trick of integration along the path to prove the eq. (3.33) for one lever $L_{m,t}$. For every $x \in L_{m,t}$, we denote by $\mathbf{r}(x,t^*)$ a *root* on the pivot level L_{m,t^*}, and choose a path $\gamma_{x,t^*} = \{\gamma_k^{x,t^*}\}_{0\leqslant k\leqslant n}$ such that

$$\gamma_0^{x,t^*} = \mathbf{r}(x,t^*), \quad \gamma_k^{x,t^*} \sim \gamma_{k+1}^{x,t^*}, \quad \gamma_n^{x,t^*} = x.$$

Moreover, we use $|\gamma^{x,t^*}|$ to represent the number of steps of the path, for example here $|\gamma^{x,t^*}| = n$. We apply a discrete Newton-Leibniz formula to get

$$u^2(x) - u^2(\mathbf{r}(x,t^*)) = \sum_{k=0}^{|\gamma^{x,t^*}|}\left(u^2(\gamma_{k+1}^{x,t^*}) - u^2(\gamma_k^{x,t^*})\right) = \sum_{k=0}^{|\gamma^{x,t^*}|} \nabla u(\gamma_k^{x,t^*}, \gamma_{k+1}^{x,t^*})\left(u(\gamma_{k+1}^{x,t^*}) + u(\gamma_k^{x,t^*})\right).$$

We put this formula into the norm of $\|u\|_{L^2(L_{m,t})}$, $t \in \left[0, \lfloor\frac{3m}{4}\rfloor\right]$ and and apply Cauchy-Schwarz inequality to obtain that

$$\begin{aligned}
\|u\|^2_{L^2(L_{m,t})} &= \sum_{x\in L_{m,t}}\left(u^2(\mathbf{r}(x,t^*)) + \sum_{k=0}^{|\gamma^{x,t^*}|}\nabla u(\gamma_k^{x,t^*}, \gamma_{k+1}^{x,t^*})\left(u(\gamma_{k+1}^{x,t^*}) + u(\gamma_k^{x,t^*})\right)\right)\\
&\leqslant \sum_{x\in L_{m,t}} u^2(\mathbf{r}(x,t^*))\\
&\quad + 2\left(\sum_{x\in L_{m,t}}\sum_{k=0}^{|\gamma^{x,t^*}|}\left(\nabla u(\gamma_k^{x,t^*}, \gamma_{k+1}^{x,t^*})\right)^2\right)^{\frac{1}{2}}\left(\sum_{x\in L_{m,t}}\sum_{k=0}^{|\gamma^{x,t^*}|}\left(u^2(\gamma_{k+1}^{x,t^*}) + u^2(\gamma_k^{x,t^*})\right)\right)^{\frac{1}{2}}\\
&\leqslant \sum_{y\in L_{m,t^*}} u^2(y)\left(\sum_{x\in L_{m,t}}\mathbf{1}_{\{y=\mathbf{r}(x,t^*)\}}\right)\\
&\quad + 4\left(\sum_{\{y_1,y_2\}\in E_d(\square_m)}(\nabla u(y_1,y_2))^2\left(\sum_{x\in L_{m,t}}\mathbf{1}_{\{\{y_1,y_2\}\in\gamma^{x,t^*}\}}\right)\right)^{\frac{1}{2}}\\
&\qquad \times\left(\sum_{y\in\square_m}u^2(y)\left(\sum_{x\in L_{m,t}}\mathbf{1}_{\{y\in\gamma^{x,t^*}\}}\right)\right)^{\frac{1}{2}}.
\end{aligned}$$

The next step is to decide how to choose the root $\mathbf{r}(x,t^*)$ and the path. The main idea is to make every edge and every vertex as root is passed by $\{\gamma^{x,t^*}\}_{x\in L_{m,t}}$ a finite number of times bounded by a constant $C(d)$. One possible plan is to choose the root $\mathbf{r}(x,t^*)$ and the path γ^{x,t^*} a discrete path in $(\mathbb{Z}^d, E_d)$ which is the closest to the vector $\overrightarrow{Ox}$, then it is a simple exercise to see that it gives us a bound $C(d)$. See Figure 3.7 as a visualization. Then we get that

$$\|u\|^2_{L^2(L_{m,t})} \leqslant C(d)\left(\|u\|^2_{L^2(L_{m,t^*})} + \|u\|_{L^2(\square_m)}\|\nabla u\|_{L^2(\square_m)}\right).$$

Then we put the eq. (3.130) and get

$$\|u\|^2_{L^2(L_{m,t})} \leqslant 4C(d)\left(3^{-m}\|u\|^2_{L^2(\square_m)} + \|u\|_{L^2(\square_m)}\|\nabla u\|_{L^2(\square_m)}\right).$$

eq. (3.33) is just a result by summing all the levels of distance less than K. □

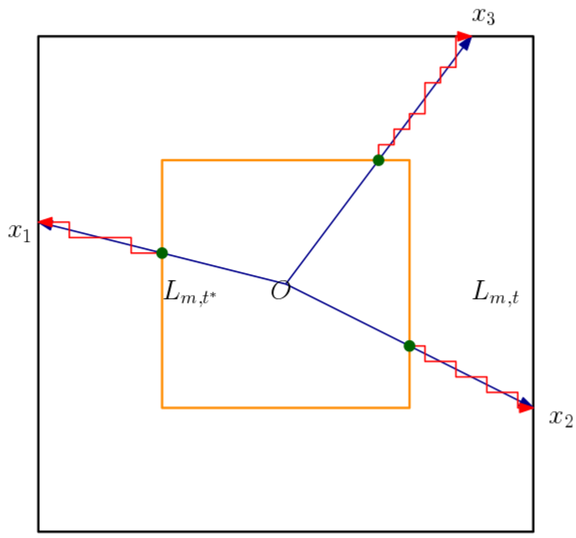

Figure 3.7: To construct the path γ^{x,t^*} for every $x \in L_{m,t}$, one can find at first the pivot level L_{m,t^*}. Then we connect O and x and find one of its closest discrete path in $(\mathbb{Z}^d, E_d)$ and denote by γ^{x,t^*} the segment from L_{m,t^*} to x. By this consturction, every edge and vertex is passed by the paths $\{\gamma^{x,t^*}\}_{x\in L_{m,t}}$ at most $C(d)$ times. In this picture, the arrows in blud indicate the vectors Ox_1, Ox_2, Ox_3 and the segments with arrow in red are the paths $\gamma^{x_1,t^*}, \gamma^{x_2,t^*}, \gamma^{x_3,t^*}$.

3.B Small clusters

This part is devoted to studying the small clusters in the percolation. Many of the arguments presented here have appeared in the previous work [19]. We extract those results from [19] and expand upon certain points that are useful for our purposes. The motivation to state these results comes from the technique of partition of good cubes:

Question 3.B.1. In a cube $\square \in \mathcal{T}$ and its enlarged domain $\mathrm{cl}_{\mathcal{P}}(\square)$, besides the maximal cluster $\mathscr{C}_*(\square)$, what is the behavior of the other finite connected clusters ?

Question 3.B.2. When we apply Lemma 3.2.3, since $\mathscr{C}_*(\square)$ and $\bigcup_{z\in\square} \mathscr{C}_*(\square_{\mathcal{P}}(z))$ are not necessarily equal, how can we describe the difference between the two ?

Question 3.B.3. What is the difference between $\mathscr{C}_\infty \cap \square$ and $\mathscr{C}_*(\square)$?

We start with a first very elementary lemma:

Lemma 3.B.1. *For any $\square \in \mathcal{T}$ and $z \in (\mathscr{C}_\infty \cap \mathrm{cl}_{\mathcal{P}}(\square)) \backslash \mathscr{C}_*(\square)$, there exists a cluster $\mathscr{C}'$ such that $z \in \mathscr{C}'$ and $\mathscr{C}' \overset{\mathbf{a}}{\leftrightarrow} \partial\, \mathrm{cl}_{\mathcal{P}}(\square)$.*

Proof. For a cube $\square \in \mathcal{T}$ and its enlarged domain $\mathrm{cl}_{\mathcal{P}}(\square)$, there exist three types of clusters:

1. One unique maximal cluster $\mathscr{C}_*(\square)$;
2. The isolated clusters which connect neither to $\mathscr{C}_*(\square)$ nor to the boundary $\partial\, \mathrm{cl}_{\mathcal{P}}(\square)$;
3. The clusters which do not connect to $\mathscr{C}_*(\square)$ but connect to the boundary.

Then it is clear the cluster $\mathscr{C}'$ containing $z \in (\mathscr{C}_\infty \cap \mathrm{cl}_{\mathcal{P}}(\square)) \backslash \mathscr{C}_*(\square)$ can only be of the third type and this proves the lemma. $\square$

We define the third class above as small clusters (see Figure 3.9). For any $z \in \mathbb{Z}^d$, we denote by $\mathscr{C}'(z)$ the clusters containing z.

Definition 3.B.1 (Small clusters)**.** For any $\square \in \mathcal{T}$, we define *small clusters* in $\square$ as the union of clusters , restricted to $\mathrm{cl}_{\mathcal{P}}(\square)$, different from $\mathscr{C}_*(\square)$ but connecting to $\partial\,\mathrm{cl}_{\mathcal{P}}(\square)$, and we denote it by $\mathscr{C}_s(\square)$, i.e.

$$\mathscr{C}_s(\square) := \bigcup_{z \in \partial\,\mathrm{cl}_{\mathcal{P}}(\square) \setminus \mathscr{C}_*(\square)} \mathscr{C}'(z).$$

Intuitively, these small clusters should be of order $\mathrm{size}(\square)^{d-1}$ when the cube $\square$ is large. This is indeed true, as we prove the following lemma:

Lemma 3.B.2. *For any $\square \in \mathcal{T}$, the set $\mathscr{C}_s(\square)$ has the following decomposition*

$$\mathscr{C}_s(\square) \subseteq \bigcup_{z \in \partial\,\mathrm{cl}_{\mathcal{P}}(\square)} \square_{\mathcal{P}}(z), \tag{3.131}$$

and has the estimate

$$|\mathscr{C}_s(\square)|\mathbf{1}_{\{\square \in \mathcal{P}_*\}} \leqslant \mathcal{O}_1(C\,\mathrm{size}(\square)^{d-1}). \tag{3.132}$$

Proof. We prove at first eq. (3.131). In the case that $\square \notin \mathcal{P}_*$, it is obvious since it has to enlarge to $\mathrm{cl}_{\mathcal{P}}(\square)$ which is a larger cube, and all the terms on the right hand side of eq. (3.131) refer to $\mathrm{cl}_{\mathcal{P}}(\square)$.

In the case that $\square \in \mathcal{P}_*$, we consider one cluster $\mathscr{C}'$ connecting to $x \in \partial\square$. We suppose that it is not contained in the union of the elements of $\mathcal{P}$ lying on $\partial\square$, then it has to cross into the interior. As illustration in Figure 3.8, it has several situations:

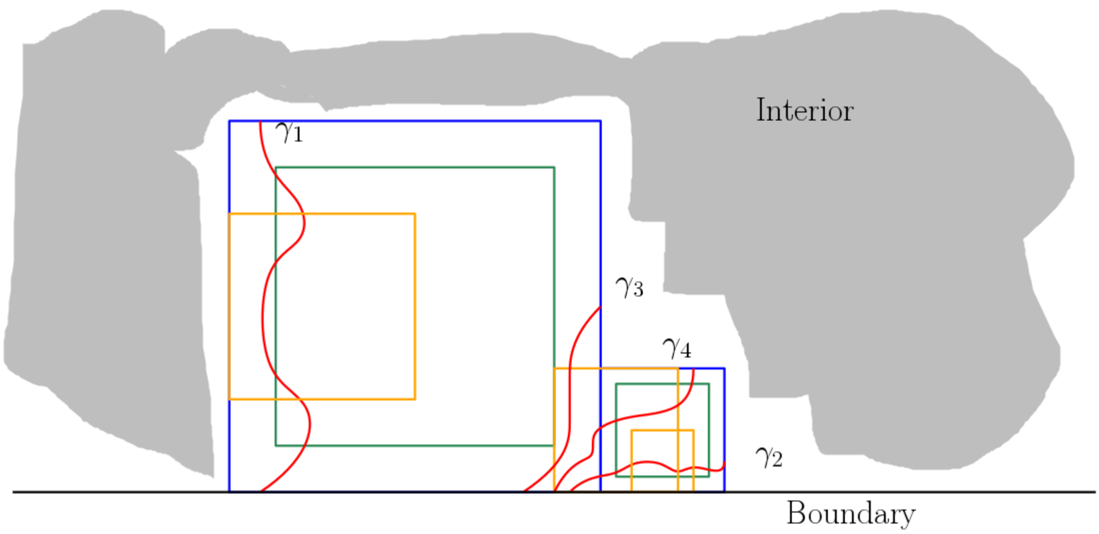

Figure 3.8: The image explain why $\mathscr{C}'$ is contained in the union of partition cubes lying at the $\partial\square$. Without loss of generality, we suppose the big blue cube is $\square_{\mathcal{P}}(x)$ and the small one is its neighbor good cube. The cube in color of green is the part of size $\frac{3}{4}$ of the good cube. The paths in red represent different typical situations that if a finite cluster connects to the boundary $\partial\square$ and wants to cross $\bigcup_{z \in \partial\,\mathrm{cl}_{\mathcal{P}}(\square)} \square_{\mathcal{P}}(z)$.

1. The first case is that $\mathscr{C}'$ cross at least one pair of $(d-1)$-dimensional opposite face of partition cube, as showed as γ_1 or γ_2. For the case γ_1, we have $|\mathrm{diam}(\mathscr{C}')| > \mathrm{size}(\square_{\mathcal{P}}(x))$;

for the case γ_2, we have $|\operatorname{diam}(\mathscr{C}')| > \frac{1}{3}\operatorname{size}(\square_{\mathcal{P}}(x))$. Then by the definition of partition cube, we can find a cube $\square'$ of $\frac{1}{2}\operatorname{size}(\square_{\mathcal{P}}(x))$ to contain parts of γ_1 and $\square'$ intersects $\frac{3}{4}\square_{\mathcal{P}}(x)$, so by the definition of good cube we have necessarily $\mathscr{C}' \overset{\mathbf{a}}{\leftrightarrow} \mathscr{C}_*(\square_{\mathcal{P}}(x))$. Same discussion can be also applied to the case γ_2. This gives a contradiction.

2. The second case is that $\mathscr{C}'$ does not cross any pair of $(d-1)$-dimensional opposite face of partition cube, but also enter the interior of $\square$ by $\partial\square_{\mathcal{P}}(x)$ or the boundary of its neighbor, so $|\operatorname{diam}(\mathscr{C}')| > \frac{1}{3}\operatorname{size}(\square_{\mathcal{P}}(x))$. One can always find a cube $\square'$ of size $\frac{1}{3}\operatorname{size}(\square_{\mathcal{P}}(x))$ crossed by $\mathscr{C}'$. If it is the case in γ_3 that $\square'$ intersects $\frac{3}{4}\square_{\mathcal{P}}(x)$, then we apply the definition of good cubes and $\mathscr{C}' \overset{\mathbf{a}}{\leftrightarrow} \mathscr{C}_*(\square_{\mathcal{P}}(x))$. Otherwise, in the case γ_4, $\mathscr{C}'$ must cross a cube $\square''$ of size $\frac{1}{6}\operatorname{size}(\square_{\mathcal{P}}(x)$ in its neighbor and we apply the same discussion, which also gives a contradiction.

To estimate the upper bound eq. (3.132), we use the decomposition above and calculate the volume of $\bigcup_{z\in\partial\square}\square_{\mathcal{P}}(z)$ by doing a contour integration along $\partial\square$ of height function $\operatorname{size}(\square_{\mathcal{P}}(z))$ and then applying eq. (3.23),

$$|\mathscr{C}_s(\square)|\mathbf{1}_{\{\square\in\mathcal{P}_*\}} \leqslant \left|\bigcup_{z\in\partial\square}\square_{\mathcal{P}}(z)\right| \leqslant \sum_{z\in\partial\square}\operatorname{size}(\square_{\mathcal{P}}(z)) \leqslant \mathcal{O}_1(C\operatorname{size}(\square)^{d-1}).$$

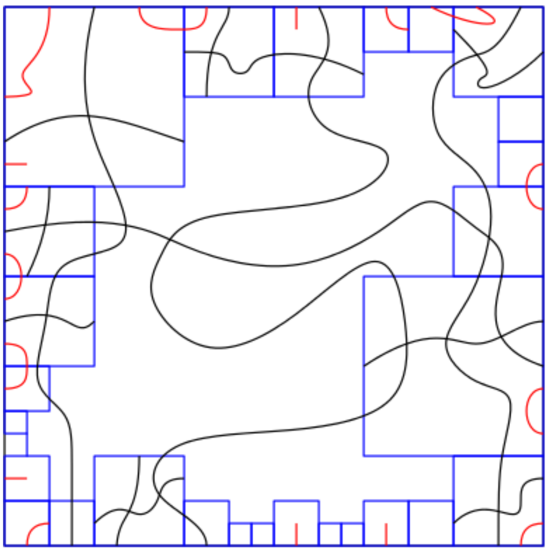

Figure 3.9: The black cluster is $\mathscr{C}_*(\square)$. The cubes in blue are the good cubes at the boundary, which contains the small cluster $\mathscr{C}_s(\square)$ in color red. Its volume can be controlled by the integration along $\partial\square$ of the size of the partition cubes.

□

Thus, Lemmas 3.B.1 and 3.B.2 answer Question 3.B.1, and the notation of $\mathscr{C}_s(\square)$ also helps us to solve Question 3.B.2:

Lemma 3.B.3. *For $\square\in\mathcal{T}$ such that $\operatorname{size}(\square) > n > 0$, we have the estimate*

$$\left|\mathscr{C}_*(\square)\backslash\left(\bigcup_{z\in 3^n\mathbb{Z}^d\cap\square}\mathscr{C}_*(z+\square_n)\right)\right|\mathbf{1}_{\{\square\in\mathcal{P}_*\}} \leqslant \mathcal{O}_1(C|\square|3^{-n}). \tag{3.133}$$

Proof. We decompose this difference in every cube of size 3^n

$$\begin{aligned}
&\left|\mathscr{C}_*(\square)\setminus\left(\bigcup_{z\in 3^n\mathbb{Z}^d\cap\square}\mathscr{C}_*(z+\square_n)\right)\right|\mathbf{1}_{\{\square\in\mathcal{P}_*\}}\\
\leqslant&\sum_{z\in 3^n\mathbb{Z}^d\cap\square}\left|(\mathscr{C}_*(\square)\cap(z+\square_n))\setminus\left(\bigcup_{z\in 3^n\mathbb{Z}^d\cap\square}\mathscr{C}_*(z+\square_n)\right)\right|\mathbf{1}_{\{\square\in\mathcal{P}_*\}}\\
\leqslant&\sum_{z\in 3^n\mathbb{Z}^d\cap\square}|(\mathscr{C}_*(\square)\cap(z+\square_n))\setminus\mathscr{C}_*(z+\square_n)|\,\mathbf{1}_{\{\square\in\mathcal{P}_*\}}\\
\leqslant&\sum_{z\in 3^n\mathbb{Z}^d\cap\square}|(\mathscr{C}_*(\square)\cap(z+\square_n))\setminus\mathscr{C}_*(z+\square_n)|\,\mathbf{1}_{\{z+\square_n\in\mathcal{P}_*\}}\\
&\quad+\sum_{z\in 3^n\mathbb{Z}^d\cap\square}|(\mathscr{C}_*(\square)\cap(z+\square_n))\setminus\mathscr{C}_*(z+\square_n)|\,\mathbf{1}_{\{z+\square_n\notin\mathcal{P}_*\}}.
\end{aligned}$$

The two terms can be treated separately. For the case $z+\square_n\in\mathcal{P}_*$, as we have mentioned, we have $\mathrm{cl}_{\mathcal{P}}(z+\square_n)=z+\square_n$ and $|(\mathscr{C}_*(\square)\cap(z+\square_n))\setminus\mathscr{C}_*(z+\square_n)|$ can be counted at the boundary $\partial(z+\square_n)$.

We turn this argument into the estimate using eq. (3.34) and eq. (3.23)

$$\begin{aligned}
\sum_{z\in 3^n\mathbb{Z}^d\cap\square}|(\mathscr{C}_*(\square)\cap(z+\square_n))\setminus\mathscr{C}_*(z+\square_n)|\,\mathbf{1}_{\{z+\square_n\in\mathcal{P}_*\}}\leqslant&\sum_{z\in 3^n\mathbb{Z}^d\cap\square}|\mathscr{C}_s(z+\square_n)|\mathbf{1}_{\{z+\square_n\in\mathcal{P}_*\}}\\
\leqslant&\mathcal{O}_1(C|\square|3^{-n}).
\end{aligned}$$

For another part, we use eq. (3.34) and eq. (3.23) directly that

$$\begin{aligned}
\sum_{z\in 3^n\mathbb{Z}^d\cap\square}\left|(\mathscr{C}_*(\square)\cap(z+\square_n))\setminus\mathscr{C}_*(z+\square_n)\mathbf{1}_{\{z+\square_n\notin\mathcal{P}_*\}}\right|\leqslant&\sum_{z\in 3^n\mathbb{Z}^d\cap\square}|z+\square_n|\mathbf{1}_{\{z+\square_n\notin\mathcal{P}_*\}}\\
\leqslant&\mathcal{O}_1(C|\square|3^{-n}).
\end{aligned}$$

We combine all these estimates and conclude the result. □

Finally, we study Question 3.B.3 on $(\mathscr{C}_\infty(\square)\cap\square)\setminus\mathscr{C}_*(\square)$:

Lemma 3.B.4. *Under the condition $\square\in\mathcal{P}_*$, and we use $\widetilde{\square}$ to represent its predecessor, then we have $(\mathscr{C}_\infty\cap\square)=(\mathscr{C}_*(\widetilde{\square})\cap\square)$, and we have the estimate that*

$$|(\mathscr{C}_\infty\cap\square)\setminus\mathscr{C}_*(\square)|\,\mathbf{1}_{\{\square\in\mathcal{P}_*\}}\leqslant\mathcal{O}_1(C|\square|^{\frac{d-1}{d}}).$$

Proof. The lemma says when the cube $\square_m$ is even better than a good cube, $\mathscr{C}_*(\widetilde{\square})\cap\square$ can contain all the part of $\mathscr{C}_\infty\cap\square$. One direction $(\mathscr{C}_*(\widetilde{\square})\cap\square)\subseteq(\mathscr{C}_\infty\cap\square)$ is obvious. We prove the other direction $(\mathscr{C}_\infty\cap\square)\subseteq(\mathscr{C}_*(\widetilde{\square})\cap\square)$ by contradiction. We suppose that this direction is not correct so that there exists $z\in(\mathscr{C}_\infty\cap\square)$ but $z\notin(\mathscr{C}_*(\widetilde{\square})\cap\square)$. By Lemma 3.B.1, there exists a cluster $\mathscr{C}'$ different from $\mathscr{C}_*(\widetilde{\square})$ and $z\in\mathscr{C}'$ and $\mathscr{C}'$ connects to $\partial\widetilde{\square}$.($\square\in\mathcal{P}_*\Rightarrow\widetilde{\square}\in\mathcal{P}_*$.) Since $\mathscr{C}_*(\square)$ is part of $\mathscr{C}_*(\widetilde{\square})$, $\mathscr{C}'$ cannot connect to $\mathscr{C}_*(\square)$. Thus, there exists an open path γ such that $z\in\gamma\subseteq\mathscr{C}'$ intersecting $\partial\square$ and we have $|\gamma|>\frac{1}{3}\,\mathrm{size}(\widetilde{\square})$. This violate the second term in Proposition 3.2.1 that a large path should belong to part of $\mathscr{C}_*(\widetilde{\square})$. We suppose that $\mathrm{size}(\square)=3^n$ and then apply Lemma 3.B.3 to obtain that

$$\begin{aligned}
|(\mathscr{C}_\infty\cap\square)\setminus\mathscr{C}_*(\square)|\,\mathbf{1}_{\{\square\in\mathcal{P}_*\}}=&\left|(\mathscr{C}_*(\widetilde{\square})\cap\square)\setminus\mathscr{C}_*(\square)\right|\mathbf{1}_{\{\square\in\mathcal{P}_*\}}\\
\leqslant&\left|\mathscr{C}_*(\widetilde{\square})\setminus\left(\bigcup_{z\in 3^n\mathbb{Z}^d\cap\widetilde{\square}}\mathscr{C}_*(z+\square_n)\right)\right|\mathbf{1}_{\{\square\in\mathcal{P}_*\}}\\
\leqslant&\,\mathcal{O}_1(C|\square|3^{-n}).
\end{aligned}$$

□

Remark. The same argument can prove even a stronger result that $\left(\mathscr{C}_\infty \cap \frac{3}{4}\widetilde{\square}\right) = \left(\mathscr{C}_*(\widetilde{\square}) \cap \frac{3}{4}\widetilde{\square}\right)$.

3.C Characterization of the effective conductance

In the literature, there are several approaches to define the effective conductance $\bar{\mathbf{a}}$, and the object of this section is to give a proof of the equivalence of these definitions in the context of percolation.

Let us at first recall the definition and some useful propositions in the previous work [19, Definition 5.1]: we define the energy in the domain $U \subseteq \mathbb{Z}^d$ with $l_p(x) := p \cdot x$ boundary condition

$$\nu(U,p) := \inf_{v \in l_p + C_0(\mathrm{cl}_{\mathcal{P}}(U))} \frac{1}{2|\mathrm{cl}_{\mathcal{P}}(U)|} \langle \nabla v \cdot \mathbf{a} \nabla v \rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(U))}, \tag{3.134}$$

and we denote by $v(\cdot, U, p)$ its minimiser. The effective conductance $\bar{\mathbf{a}}$ is a deterministic positive scalar defined by

$$\frac{1}{2} p \cdot \bar{\mathbf{a}} p := \lim_{m \to \infty} \mathbb{E}[\nu(\square_m, p)], \tag{3.135}$$

with the rate of convergence [19, Lemma 4.8]: there exists $s(d) > 0, \alpha(d, \mathfrak{p}, \Lambda) \in \left(0, \frac{1}{4}\right]$ and $C(d, p, \Lambda) < \infty$ such that for every $\square \in \mathcal{T}$

$$\left| \frac{1}{2} p \cdot \bar{\mathbf{a}} p - \nu(\square, p) \right| \leqslant \mathcal{O}_s\left(C|p|^2 \,\mathrm{size}(\square)^{-\alpha}\right). \tag{3.136}$$

We will also use the following trivial bound several times in the proof

$$\nu(U,p) \leqslant \frac{1}{2|\mathrm{cl}_{\mathcal{P}}(U)|} \langle \nabla l_p \cdot \mathbf{a} \nabla l_p \rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(U))} \leqslant d|p|^2. \tag{3.137}$$

The main theorem in this part is to prove the following characterization.

Theorem 3.C.1 (Characterization of the effective conductance)**.** *In the context of homogenization in supercritical percolation, the effective conductance $\bar{\mathbf{a}}$ is a positive scalar constant and the following definitions are equivalent:*

$$p \cdot \bar{\mathbf{a}} p \overset{a.s}{=} \lim_{m \to \infty} \frac{1}{|\mathrm{cl}_{\mathcal{P}}(\square_m)|} \langle \nabla v(\cdot, \square_m, p) \cdot \mathbf{a} \nabla v(\cdot, \square_m, p) \rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))}. \tag{3.138}$$

$$p \cdot \bar{\mathbf{a}} p \overset{a.s}{=} \lim_{m \to \infty} \frac{1}{|\square_m|} \inf_{v \in l_p + C_0(\square_m)} \langle \nabla v \cdot \mathbf{a} \nabla v \rangle_{E_d^{\mathbf{a}}(\square_m)}. \tag{3.139}$$

$$p \cdot \bar{\mathbf{a}} p = \mathbb{E}[\mathcal{D}(\phi_p + l_p) \cdot \mathbf{a}_{\mathscr{C}} \mathcal{D}(\phi_p + l_p)]. \tag{3.140}$$

$$\bar{\mathbf{a}} p = \mathbb{E}[\mathbf{a}_{\mathscr{C}} \mathcal{D}(\phi_p + l_p)]. \tag{3.141}$$

Before starting the proof, we give some remarks on these definitions. Equation (3.138) is just a variant of eq. (3.135). Equation (3.139) differs from the first one in that just it does the minimization but does not enlarge the domain to $\mathrm{cl}_{\mathcal{P}}(\square_m)$ nor restricts the problem to $\mathscr{C}_*(\square_m)$. Equation (3.140) uses the linear $\mathbf{a}$-harmonic function in the whole space instead of that in $\square_m$, so it is stationary. The last one is a little different from the previous three ones, but we need it in Proposition 3.1.2, thus we add it to the list of equivalent definitions.

Proof. Equation (3.138) is a direct consequence of eq. (3.135) and eq. (3.136), Markov's inequality and the lemma of Borel-Cantelli to transform it to an "almost sure" version.

Equation (3.139) is a variant from the first one, especially when $\square_m \in \mathcal{P}_*$ they are very close. So we do the decomposition

$$\begin{aligned}&\left|\frac{1}{|\square_m|}\inf_{v\in l_p+C_0(\square_m)}\langle\nabla v\cdot\mathbf{a}\nabla v\rangle_{E_d^{\mathbf{a}}(\square_m)}-p\cdot\bar{\mathbf{a}}p\right|\\ \leqslant&\left|\frac{1}{|\square_m|}\inf_{v\in l_p+C_0(\square_m)}\langle\nabla v\cdot\mathbf{a}\nabla v\rangle_{E_d^{\mathbf{a}}(\square_m)}-p\cdot\bar{\mathbf{a}}p\right|\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}\\ &\quad+\left|\frac{1}{|\square_m|}\inf_{v\in l_p+C_0(\square_m)}\langle\nabla v\cdot\mathbf{a}\nabla v\rangle_{E_d^{\mathbf{a}}(\square_m)}-p\cdot\bar{\mathbf{a}}p\right|\mathbf{1}_{\{\square_m\notin\mathcal{P}_*\}}\end{aligned}$$

and the second one can be handled easily by a trivial bound by comparing with l_p as in eq. (3.137)

$$\left|\frac{1}{|\square_m|}\inf_{v\in l_p+C_0(\square_m)}\langle\nabla v\cdot\mathbf{a}\nabla v\rangle_{E_d^{\mathbf{a}}(\square_m)}-p\cdot\bar{\mathbf{a}}p\right|\mathbf{1}_{\{\square_m\notin\mathcal{P}_*\}}\leqslant\mathcal{O}_1(C(d,\mathfrak{p},\Lambda)|p|^2 3^{-m}).$$

By an argument of Borel-Cantelli, we prove this term converges almost surely to 0. Then we focus on the case $\square_m \in \mathcal{P}_*$, In fact, in this case the minimiser on $E_d^{\mathbf{a}}(\square_m)$ is the sum of the one on each clusters. Observing that the one on isolated cluster from $\partial\square_m$ can be null since it has no boundary condition, so we have to deal with the one on $\mathscr{C}_*(\square_m)$ and the one on the small clusters $\mathscr{C}_s(\square_m)$. We apply eq. (3.132), the estimate eq. (3.34) and a trivial bound eq. (3.137) to get

$$\begin{aligned}&\left|\frac{1}{|\square_m|}\inf_{v\in l_p+C_0(\square_m)}\langle\nabla v\cdot\mathbf{a}\nabla v\rangle_{E_d^{\mathbf{a}}(\square_m)}-p\cdot\bar{\mathbf{a}}p\right|\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}\\ \leqslant&\left|\frac{1}{|\operatorname{cl}_{\mathcal{P}}(\square_m)|}\langle\nabla v(\cdot,\square_m,p)\cdot\mathbf{a}\nabla v(\cdot,\square_m,p)\rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))}-p\cdot\bar{\mathbf{a}}p\right|\\ &\quad+\left|\frac{1}{|\operatorname{cl}_{\mathcal{P}}(\square_m)|}\langle\nabla v(\cdot,\square_m,p)\cdot\mathbf{a}\nabla v(\cdot,\square_m,p)\rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))}-p\cdot\bar{\mathbf{a}}p\right|\mathbf{1}_{\{\square_m\notin\mathcal{P}_*\}}\\ &\quad+\left|\frac{1}{|\square_m|}\inf_{v\in l_p+C_0(\square_m)}\langle\nabla v\cdot\mathbf{a}\nabla v\rangle_{E_d^{\mathbf{a}}(\mathscr{C}_s(\square_m))}\right|\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}\\ \leqslant&\left|\frac{1}{|\operatorname{cl}_{\mathcal{P}}(\square_m)|}\langle\nabla v(\cdot,\square_m,p)\cdot\mathbf{a}\nabla v(\cdot,\square_m,p)\rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))}-p\cdot\bar{\mathbf{a}}p\right|\\ &\quad+C(d)|p|^2\mathbf{1}_{\{\operatorname{size}(\square_{\mathcal{P}}(0))>3^m\}}+C(d)|p|^2\frac{|\mathscr{C}_s(\square_m)|\mathbf{1}_{\{\square_m\in\mathcal{P}_*\}}}{|\square_m|}\\ \leqslant&\left|\frac{1}{|\operatorname{cl}_{\mathcal{P}}(\square_m)|}\langle\nabla v(\cdot,\square_m,p)\cdot\mathbf{a}\nabla v(\cdot,\square_m,p)\rangle_{E_d^{\mathbf{a}}(\mathscr{C}_*(\square_m))}-p\cdot\bar{\mathbf{a}}p\right|+\mathcal{O}_1(C(d,\mathfrak{p},\Lambda)|p|^2 3^{-m}).\end{aligned}$$

So its almost sure limit is the same as the first one when $m\to\infty$.

By a similar calculation, one can prove a variant of eq. (3.139) that reads

$$p\cdot\bar{\mathbf{a}}p\overset{a.s}{=}\lim_{m\to\infty}\frac{1}{|\square_m|}\inf_{v\in l_p+C_0(\square_m)}\langle\nabla v\cdot\mathbf{a}_{\mathscr{C}}\nabla v\rangle_{E_d^{\mathbf{a}}(\square_m)},\tag{3.142}$$

and we recall, see [145, Theorem 9.1], that this definition coincides with eq. (3.140). By a calculus of variation argument, we have

$$\forall p,q\in\mathbb{R}^d,\qquad q\cdot\bar{\mathbf{a}}p=\mathbb{E}[\mathcal{D}(\phi_q+l_q)\cdot\mathbf{a}_{\mathscr{C}}\mathcal{D}(\phi_p+l_p)].$$

Moreover, observing that $\mathbf{1}_{\{\mathbf{a}\neq 0\}}\mathcal{D}\phi_q + q$ and $\mathbf{a}_{\mathscr{C}}(\mathcal{D}\phi_p + p)$ are stationary, and the former is a the gradient and the latter is divergence free, we can use the Div-Curl and Birkhoff theorems

$$q \cdot \bar{\mathbf{a}}p = \mathbb{E}[\mathbf{1}_{\{\mathbf{a}\neq 0\}}\mathcal{D}\phi_q + q]\mathbb{E}[\mathbf{a}_{\mathscr{C}}(\mathcal{D}\phi_p + p)] = q \cdot \mathbb{E}[\mathbf{a}_{\mathscr{C}}(\mathcal{D}\phi_p + p)].$$

This concludes the equivalence with eq. (3.141). □

Chapter 4

Convergence rate of the heat kernel on the infinite percolation cluster

We study the heat kernel and the Green's function on the infinite supercritical percolation cluster in dimension $d \geqslant 2$ and prove a quantitative homogenization theorem for these functions with an almost optimal rate of convergence. These results are a quantitative version of the local central limit theorem proved by Barlow and Hambly in [41]. The proof relies on a structure of renormalization for the infinite percolation cluster introduced in [19], Gaussian bounds on the heat kernel established by Barlow in [39] and tools of the theory of quantitative stochastic homogenization. An important step in the proof is to establish a $C^{0,1}$-large-scale regularity theory for caloric functions on the infinite cluster and is of independent interest.

This chapter corresponds to the article [85], which is joint work with Paul Dario and is published in Annals of Probability.

4.1 Introduction

4.1.1 General introduction and main results

In this article, we study the continuous-time random walk on the infinite cluster of the supercritical Bernoulli bond percolation of the Euclidean lattice $\mathbb{Z}^d$, in dimension $d \geqslant 2$. The model considered is a specific case of the general random conductance model and can be described as follows. We let E_d be the set of bonds of $\mathbb{Z}^d$, i.e., the set of unordered pairs of nearest neighbors of $\mathbb{Z}^d$. We denote by Ω the set of functions from E_d to the set of non-negative real numbers $[0, \infty)$. A generic element of Ω is denoted by $\mathbf{a}$ and called *an environment*.

For a given environment $\mathbf{a} \in \Omega$ and a given bond $e \in E_d$, we call the value $\mathbf{a}(e)$ *the conductance* of the bond e. We fix an ellipticity parameter $\lambda \in (0, 1]$ and add some randomness to the model by assuming that the collection of conductances $\{\mathbf{a}(e)\}_{e \in E_d}$ is an i.i.d. family of random variables whose law is supported in the set $\{0\} \cup [\lambda, 1]$. We define $\mathfrak{p} := \mathbb{P}[\mathbf{a}(e) \neq 0]$ and assume that

$$\mathfrak{p} > \mathfrak{p}_c(d),$$

where $\mathfrak{p}_c(d)$ is the bond percolation threshold for the lattice $\mathbb{Z}^d$. This assumption ensures that, almost surely, there exists a unique infinite connected component of edges with non-zero conductances (or cluster) which we denote by $\mathscr{C}_\infty$ (see [67]). This cluster has a non-zero density which is given by the probability $\theta(\mathfrak{p}) := \mathbb{P}[0 \in \mathscr{C}_\infty]$. The model of continuous-time random walk considered in this article is *the variable speed random walk* (or VSRW) and is defined as follows. Given an environment $\mathbf{a} \in \Omega$ and a starting point $y \in \mathscr{C}_\infty$, we endow each edge $e \in E_d$ with a random clock whose law is exponential of parameter $\mathbf{a}(e)$ and assume that they are mutually independent. We then let $(X_t)_{t \geqslant 0}$ be the random walk which starts from y, i.e., $X_0 = y$, and, when $X(t) = x$, the random walker waits at x until one of the clocks at an adjacent edge to x rings, and moves across the edge to the neighboring point instantly. We then restart the clocks. This construction gives rise to a continuous-time Markov process on the infinite cluster $\mathscr{C}_\infty$ whose generator is the elliptic operator $\nabla \cdot \mathbf{a} \nabla$ defined by, for each function $u : \mathscr{C}_\infty \to \mathbb{R}$ and each point $x \in \mathscr{C}_\infty$,

$$\nabla \cdot \mathbf{a} \nabla u(x) := \sum_{z \sim x} \mathbf{a}(\{x, z\}) \left(u(z) - u(x)\right). \tag{4.1}$$

We denote the transition density of the random walk by

$$p(t, x, y) = p^{\mathbf{a}}(t, x, y) := \mathbb{P}_y^{\mathbf{a}}(X_t = x),$$

and often omit the dependence in the environment $\mathbf{a}$ in the notation. The transition density can be equivalently defined as the solution of the parabolic equation

$$\begin{cases} \partial_t p(\cdot, \cdot, y) - \nabla \cdot \mathbf{a} \nabla p(\cdot, \cdot, y) = 0 & \text{in } (0, \infty) \times \mathscr{C}_\infty, \\ p(0, \cdot, y) = \delta_y & \text{in } \mathscr{C}_\infty. \end{cases} \tag{4.2}$$

Due to this characterization, we often refer to the transition density p as the *heat kernel* or the *parabolic Green's function*.

There are other related models of random walk on supercritical percolation clusters which have been studied in the literature, two of the most common ones are:

(i) *The constant speed random walk* (or CSRW), the random walker starts from a point $y \in \mathscr{C}_\infty$. When $X(t) = x$, it waits for an exponential time of parameter 1 and then jumps to a neighboring point z according to the transition probability

$$P(x,z) = \frac{\mathbf{a}(\{x,z\})}{\sum_{w\sim x}\mathbf{a}(\{x,w\})}. \tag{4.3}$$

This construction also gives rise to a continuous-time Markov process whose generator is given by, for each function $u : \mathscr{C}_\infty \to \mathbb{R}$ and each point $y \in \mathscr{C}_\infty$,

$$\frac{1}{\sum_{z\sim x}\mathbf{a}(\{x,z\})}\sum_{z\sim x}\mathbf{a}(\{x,z\})\left(u(z)-u(x)\right).$$

(ii) *The simple random walk* (or SRW), the random walk $(X_n)_{n\in\mathbb{N}}$ is indexed on the integers, it starts from a point $y \in \mathscr{C}_\infty$, when $X_n = x$, the value of X_{n+1} is chosen randomly among all the neighbors of x following the transition probability (4.3).

These processes have similar, although not identical, properties and have been the subject of interest in the literature. In the case of the percolation cluster, i.e., when the environment $\mathbf{a}$ is only allowed to take the values 0 or 1, an annealed invariance principle was proved in [89] by De Masi, Ferrari, Goldstein and Wick. In [211], Sidoravicius and Sznitman proved a quenched invariance principle for the simple random walk in dimension $d \geqslant 4$. This result was extended to every dimension $d \geqslant 2$ by Berger and Biskup in [49] (for the SRW) and by Mathieu and Piatnitski in [180] (for the CSRW).

For the VSRW, a similar quenched invariance principle holds: there exists a deterministic diffusivity constant $\bar\sigma > 0$ such that, for almost every environment, the following convergence holds in the Skorokhod topology

$$\varepsilon X_{\frac{\cdot}{\varepsilon^2}} \underset{\varepsilon\to 0}{\overset{(\text{law})}{\longrightarrow}} \bar\sigma B_\cdot, \tag{4.4}$$

where $B_\cdot$ is a standard Brownian motion. From a homogenization perspective, the diffusivity $\bar\sigma^2$ of the limiting Brownian motion is related to the homogenized coefficient $\bar{\mathbf{a}}$ associated to the elliptic and parabolic problems on the percolation cluster by the identity $\bar{\mathbf{a}} = \frac{1}{2}\theta(\mathfrak{p})\bar\sigma^2$ (see the formula (4.182) of Appendix 4.B).

The properties of the heat kernel p on the infinite cluster have been investigated in the literature. In [181], Mathieu and Remy proved that, almost surely, the heat kernel decays as fast as $t^{-d/2}$. These bounds were extended in [39] by Barlow who established Gaussian lower and upper bounds for this function; we will recall his precise result in Theorem 4.3.1 below.

In the article [41], Barlow and Hambly proved a parabolic Harnack inequality, a local central limit theorem for the CSRW, and bounds on the elliptic Green's function on the infinite cluster. Their main result can be adapted to the case of the VSRW, and reads as follows: if we define, for each $t \geqslant 0$ and $x \in \mathbb{R}^d$,

$$\bar p(t,x) := \frac{1}{(2\pi\bar\sigma^2 t)^{d/2}}\exp\left(-\frac{|x|^2}{2\,\bar\sigma^2 t}\right), \tag{4.5}$$

the heat kernel with diffusivity $\bar\sigma$, then, for each time $T > 0$, the following convergence holds, $\mathbb{P}$-almost surely on the event $\{0 \in \mathscr{C}_\infty\}$,

$$\lim_{n\to\infty}\left|n^{d/2}p(nt, g_n^\omega(x), 0) - \theta(\mathfrak{p})^{-1}\bar p(t,x)\right| = 0, \tag{4.6}$$

uniformly in the spatial variable $x \in \mathbb{R}^d$ and in the time variable $t \geqslant T$, where the notation $g_n^\omega(x)$ means the closest point to $\sqrt{n}x$ in the infinite cluster under the environment ω.

The main result of this article is a quantitative version of the local central limit theorem for the VSRW and is stated below.

Theorem 4.1.1. *For each exponent $\delta > 0$, there exist a positive constant $C < \infty$ and an exponent $s > 0$, depending only on the parameters $d, \lambda, \mathfrak{p}$ and δ, such that for every $y \in \mathbb{Z}^d$, there exists a non-negative random time $\mathcal{T}_{\mathrm{par},\delta}(y)$ satisfying the stochastic integrability estimate*

$$\forall T \geqslant 0,\ \mathbb{P}\left(\mathcal{T}_{\mathrm{par},\delta}(y) \geqslant T\right) \leqslant C \exp\left(-\frac{T^s}{C}\right),$$

such that, on the event $\{y \in \mathscr{C}_\infty\}$, for every $x \in \mathscr{C}_\infty$ and every $t \geqslant \max\left(\mathcal{T}_{\mathrm{par},\delta}(y), |x-y|\right)$,

$$\left|p(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y)\right| \leqslant Ct^{-\frac{d}{2}-\left(\frac{1}{2}-\delta\right)} \exp\left(-\frac{|x-y|^2}{Ct}\right). \tag{4.7}$$

Remark. The heat kernel p does not exactly converge to the heat kernel $\bar{p}$ and there is an additional normalization constant $\theta(\mathfrak{p})^{-1}$ in (4.7). A heuristic reason explaining why such a term is necessary is the following: since $p(t,\cdot,y)$ is a probability measure on the infinite cluster, one has

$$\sum_{x \in \mathscr{C}_\infty} p\left(t,x,y\right) = 1.$$

One also has, by definition of the heat kernel $\bar{p}$,

$$\int_{\mathbb{R}^d} \bar{p}(t,x-y)\,dx = 1.$$

Since the infinite cluster has density $\theta(\mathfrak{p})$, we expect that

$$\sum_{x \in \mathscr{C}_\infty} \bar{p}(t,x-y) \simeq \theta(\mathfrak{p}) \int_{\mathbb{R}^d} \bar{p}(t,x-y)\,dx = \theta(\mathfrak{p}),$$

and we refer to Proposition 4.A.3 for a precise statement. As a consequence, we cannot expect the maps p and $\bar{p}$ to be close since they have different mass on the infinite cluster; adding the normalization term $\theta(\mathfrak{p})^{-1}$ ensures that the mass of $\theta(\mathfrak{p})^{-1}\bar{p}$ on the infinite cluster is approximately equal to 1.

As an application of this result, we deduce a quantitative homogenization theorem for the elliptic Green's function on the infinite cluster. In dimension $d \geqslant 3$, given an environment $\mathbf{a} \in \Omega$ and a point $y \in \mathscr{C}_\infty$, we define the Green's function $g(\cdot,y)$ as the solution of the equation

$$-\nabla \cdot \mathbf{a}\nabla g(\cdot,y) = \delta_y \text{ in } \mathscr{C}_\infty \text{ such that } g(x,y) \underset{x\to\infty}{\longrightarrow} 0.$$

This function exists, is unique almost surely and is related to the transition probability p through the identity

$$g(x,y) = \int_0^\infty p(t,x,y)\,dt. \tag{4.8}$$

In dimension 2, the situation is different since the Green's function is not bounded at infinity, and we define $g\left(\cdot,y\right)$ as the unique function which satisfies

$$-\nabla \cdot \mathbf{a}\nabla g(\cdot,y) = \delta_y \text{ in } \mathscr{C}_\infty,\ \frac{1}{|x|} g(x,y) \underset{x\to\infty}{\longrightarrow} 0 \text{ and } g(y,y) = 0.$$

This function is related to the transition probability p through the identity

$$g(x,y) = \int_0^\infty (p(t,x,y) - p(t,y,y))\, dt.$$

In the statement below, we denote by $\bar{g}$ the homogenized Green's function defined by the formula, for each point $x \in \mathbb{R}^d \setminus \{0\}$,

$$\bar{g}(x) := \begin{cases} -\frac{1}{\pi \bar{\sigma}^2 \theta(\mathfrak{p})} \ln |x| & \text{if } d = 2, \\ \frac{\Gamma(d/2-1)}{\left(2\pi^{d/2}\bar{\sigma}^2\theta(\mathfrak{p})\right)} \frac{1}{|x|^{d-2}} & \text{if } d \geqslant 3, \end{cases} \tag{4.9}$$

where the symbol Γ denotes the standard Gamma function. Theorem 4.1.2 describes the asymptotic behavior of the Green's function g.

Theorem 4.1.2. *For each exponent $\delta > 0$, there exist a positive constant $C < \infty$ and an exponent $s > 0$, depending only on the parameters $d, \lambda, \mathfrak{p}$ and δ, such that for every $y \in \mathbb{Z}^d$, there exists a non-negative random variable $\mathcal{M}_{\mathrm{ell},\delta}(y)$ satisfying*

$$\forall R \geqslant 0,\ \mathbb{P}\left(\mathcal{M}_{\mathrm{ell},\delta}(y) \geqslant R\right) \leqslant C \exp\left(-\frac{R^s}{C}\right),$$

such that, on the event $\{y \in \mathscr{C}_\infty\}$:

1. *In dimension $d \geqslant 3$, for every point $x \in \mathscr{C}_\infty$ satisfying $|x-y| \geqslant \mathcal{M}_{\mathrm{ell},\delta}(y)$,*

$$|g(x,y) - \bar{g}(x-y)| \leqslant \frac{1}{|x-y|^{1-\delta}} \frac{C}{|x-y|^{d-2}}. \tag{4.10}$$

2. *In dimension 2, the limit*

$$K(y) := \lim_{x \to \infty} \left(g(x,y) - \bar{g}(x-y)\right),$$

exists, is finite almost surely and satisfies the stochastic integrability estimate

$$\forall R \geqslant 0,\ \mathbb{P}\left(|K(y)| \geqslant R\right) \leqslant C \exp\left(-\frac{R^s}{C}\right).$$

Moreover, for every point $x \in \mathscr{C}_\infty$ satisfying $|x-y| \geqslant \mathcal{M}_{\mathrm{ell},\delta}(y)$,

$$|g(x,y) - \bar{g}(x-y) - K(y)| \leqslant \frac{C}{|x-y|^{1-\delta}}. \tag{4.11}$$

Remark. In dimension 2, the situation is specific due to the unbounded behavior of the Green's function, and the theorem identifies the first-order term. The second term in the asymptotic development is of constant order and is random: with the normalization chosen for the Green's function, the constant K depends on the geometry of the infinite cluster and cannot be deterministic. We nevertheless expect it not to be too large and prove that it satisfies a stretched exponential stochastic integrability estimate.

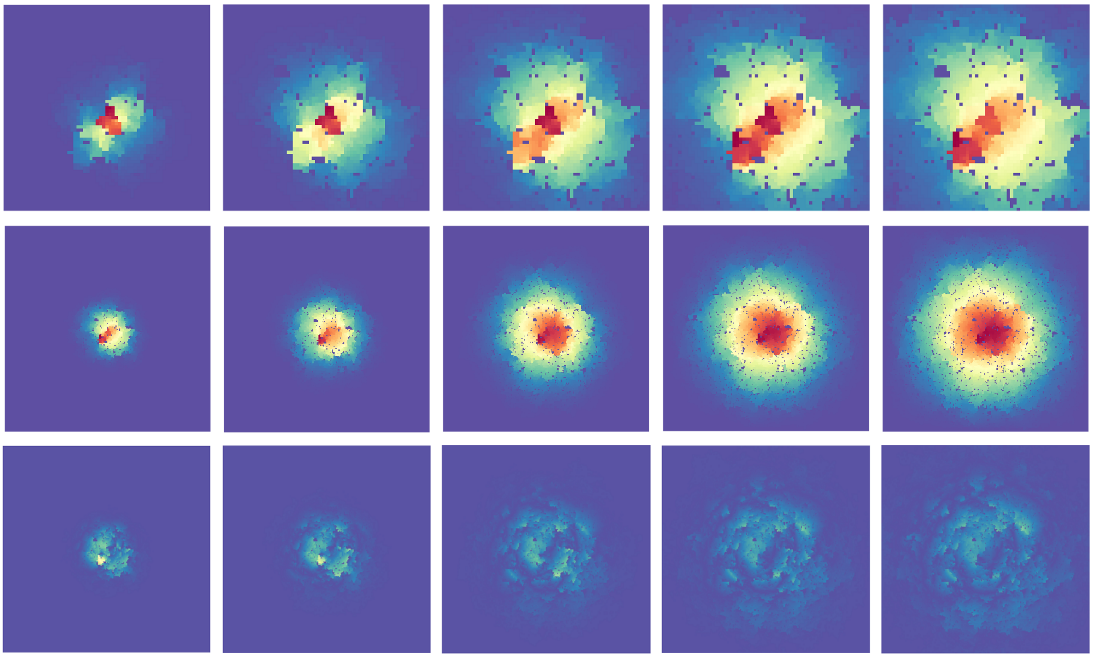

Figure 4.1: A simulation to illustrate the convergence of the parabolic Green's function for the VSRW on the infinite cluster with $\mathfrak{p} = 0.6$. We use different colors to represent the level sets of the map $t^{\frac{d}{2}} p(t, \cdot, y)$ in the first two rows. The figures in first row are drawn for the short times $t = 100, 200, 300, 400, 500$ in a cube of size 64×64 and the level sets of the heat kernel are perturbed by the geometry of the infinite cluster. In the second row, the figures are drawn for the long times $t = 500, 1000, 2000, 3000, 4000$ and in a cube of size 256×256; in this case, homogenization happens and the geometry of the level sets of the heat kernel is similar to the one of a Gaussian heat kernel. In the third row, we simulate the function $t^{\frac{d}{2}} \left| p(t, \cdot, y) - \theta(\mathfrak{p})^{-1} \bar{p}(t, \cdot - y) \right| \mathbf{1}_{\{\{x \in \mathscr{C}_\infty\}\}}$ associated to the figures in the second line and we observe that the errors decay to 0 as the time tends to infinity.

We complete this section by mentioning a potential application of these theorems. Theorem 4.1.1 shows that the law of the VSRW on the infinite percolation cluster converges quantitatively to the one of the Brownian motion $(\bar{\sigma} B_t)_{t \geqslant 0}$. To go one step further in the analysis, one can try to construct a coupling between the random walk $(X_t)_{t \geqslant 0}$ and the Brownian motion $(\bar{\sigma} B_t)_{t \geqslant 0}$ such that their trajectories are close, i.e., such that $\sup_{0 \leqslant s \leqslant t} |X_s - \bar{\sigma} B_s|$ is small. This question is known as the embedding problem: a good error should be at least of order $o(\sqrt{t})$. In the case of the simple random walk on $\mathbb{Z}^d$, the optimal result is given by the Komlós-Major-Tusnády Approximation (see [155, 156]) and gives an error of order $O(\log t)$. Adapting this result to the setting considered here requires to take into account the degenerate geometry of the percolation cluster; we believe that Theorem 4.1.1 can be useful in this regard.

4.1.2 Strategy of the proof

On the supercritical percolation cluster, a qualitative version of Theorem 4.1.1 is established by Barlow and Hambly in [41], where the strategy implemented is to first prove a parabolic Harnack inequality for the heat equation. From the Harnack inequality, one derives a $C^{0,\alpha}$-Hölder regularity estimate (for some small exponent $\alpha > 0$) on the heat kernel. It is then possible to combine this additional regularity with the quenched invariance principle, established on the percolation cluster in [211, 181, 49], to obtain the local central limit theorem.

In the present article, the strategy adopted is different and follows ideas from the theory of stochastic homogenization, more specifically the ones of [25, Chapter 8]. A first crucial ingredient in the proof is the first-order corrector, which can be characterized as follows: given a slope $p \in \mathbb{R}^d$, the corrector ϕ_p is defined as the unique function (up to a constant) which is a solution of the elliptic equation

$$-\nabla \cdot \mathbf{a} \left(p + \nabla \phi_p \right) = 0 \text{ in } \mathscr{C}_\infty,$$

and which has sublinear oscillation, i.e.,

$$\frac{1}{r} \operatorname*{osc}_{x \in \mathscr{C}_\infty \cap B_r} \phi_p := \frac{1}{r} \left(\sup_{x \in \mathscr{C}_\infty \cap B_r} \phi_p - \inf_{x \in \mathscr{C}_\infty \cap B_r} \phi_p \right) \underset{r \to \infty}{\longrightarrow} 0.$$

The corrector is defined and some of its important properties are presented in Section 4.2.3. We note that the use of the corrector to study random walk on supercritical percolation cluster is not new: it is a key ingredient in the proofs of the quenched invariance principle (see [211, 181, 49]). Once equipped with this function, the analysis relies on a classical strategy in stochastic homogenization: the two-scale expansion. The general approach relies on the definition of the function

$$h(t,x,y) := \theta(\mathfrak{p})^{-1} \left(\bar{p}\,(t, x-y) + \sum_{k=1}^{d} \partial_k \bar{p}(t, x-y) \phi_{e_k}(x) \right), \tag{4.12}$$

where $(e_k)_{k = \{1, \cdots, d\}}$ denotes the canonical basis of $\mathbb{R}^d$ and $\bar{p}$ is the continuous heat kernel defined in (4.5). The strategy is then to compute the value of

$$\partial_t h - \nabla \cdot \mathbf{a} \nabla h, \tag{4.13}$$

by using the explicit formula on h stated in (4.12) and to prove that it is quantitatively small in the correct functional space (precisely, the parabolic $\underline{H}^{-1}$ space introduced in (4.33)). Obtaining this result requires two types of quantitative information on the corrector:

- One needs to have quantitative sublinearity of the corrector, i.e.,

$$\frac{1}{r^\alpha} \operatorname*{osc}_{x \in \mathscr{C}_\infty \cap B_r} \phi_p \underset{r \to \infty}{\longrightarrow} 0, \tag{4.14}$$

for every exponent $\alpha > 0$.

- One needs to have a quantitative control on the flux of the corrector in the weak $\underline{H}^{-1}$ norm,

$$\frac{1}{r^\alpha} \left\| \mathbf{a}\left(p + \nabla \phi_p\right) - \frac{1}{2}\bar{\sigma}^2 p \right\|_{\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r)} \underset{r \to \infty}{\longrightarrow} 0, \tag{4.15}$$

for every exponent $\alpha > 0$, where $\bar{\sigma}^2$ is the same diffusivity constant as in the definition (4.5) of the heat kernel $\bar{p}$.

The sublinearity of the corrector in the setting of the percolation cluster is established qualitatively in [211, 181, 49] and quantitatively in [19, 83, 134]. The second property (4.15) cannot be directly deduced from the results of [19, 83, 134] and Appendix 4.B is devoted to the proof of this result.

Once one has good quantitative control over the $\underline{H}^{-1}$-norm of $\partial_t h - \nabla \cdot \mathbf{a} \nabla h$, the proof of the result follows from the following two arguments:

(i) First, one shows that the function h is (quantitatively) close to the function $\theta^{-1}(\mathfrak{p})\bar{p}$. This is achieved by proving that the second term in the right side of (4.12) is small and relies on the quantitative sublinearity of the corrector stated in (4.14).

(ii) Second, one needs to show that the function h is (quantitatively) close to the heat kernel p. To prove this, the strategy is to use that the map p solves the parabolic equation

$$\partial_t p - \nabla \cdot \mathbf{a} \nabla p = 0,$$

and subtract it from (4.13) to obtain that $\partial_t (p-h) - \nabla \cdot \mathbf{a} \nabla (p-h)$ is small in the $\underline{H}^{-1}$ norm. We then use the function $(p-h)$ as a test function in the previous equation, to deduce that $(p-h)$ has to be small in the H^1-norm.

This strategy is essentially carried out in Section 4.4.2. Nevertheless, a number of difficulties have to be treated in order to implement it. They are mainly due to three distinct causes which are listed below.

First, the heat kernel p has an initial condition at time $t = 0$ which is a Dirac (see the equation (4.2)). It is rather singular and causes serious troubles in the analysis. To fix this issue, one replaces the initial condition in (4.2) by a function which is smoother, but which is still a good approximation of the Dirac function. The argument is sketched in the following paragraph. We fix a large time $t > 0$ and want to prove the main estimate (4.7) for this particular time t. To this end, we replace the initial condition δ_y by the function $\bar{p}(\tau, \cdot - y)$ for some time $\tau \ll t$, and we define

$$\begin{cases} \partial_t q - \nabla \cdot (\mathbf{a} \nabla q) = 0 & \text{in } (\tau, \infty) \times \mathscr{C}_\infty, \\ q(\tau, \cdot, y) = \theta(\mathfrak{p})^{-1} \bar{p}(\tau, \cdot - y) & \text{on } \mathscr{C}_\infty. \end{cases} \tag{4.16}$$

The strategy is then to make the following compromise: we want to choose the coefficient τ small enough (in particular, much smaller than t) so that the initial data $\bar{p}(\tau, \cdot - y)$ is close to the Dirac function δ_y, the objective being that the function $q(t, \cdot, y)$ is close to $p(t, \cdot, y)$ (see Lemma 4.4.1); we also want to choose τ large enough so that the initial data $\bar{p}(\tau, \cdot)$ is smooth enough. Our choice will be $\tau = t^{1-\kappa}$ for some small exponent $\kappa > 0$. This approach is essentially the subject of Section 4.4.1.

The second difficulty is that the two-scale expansion described at the beginning of the section only yields the result for a small exponent, i.e., we obtain a result of the form

$$\left|p(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y)\right| \leqslant Ct^{-\kappa}t^{-\frac{d}{2}}\exp\left(-\frac{|x-y|^2}{Ct}\right), \tag{4.17}$$

for a small exponent $\kappa > 0$. This result is much weaker than the near-optimal exponent $\frac{1}{2} - \delta$ stated in Theorem 4.1.1. The strategy is thus to improve the value of the exponent by a bootstrap argument: by redoing the two-scale expansion and by using the estimate (4.17) in the proof, we obtain an improved estimate of the form

$$\left|p(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y)\right| \leqslant Ct^{-\kappa_1}t^{-\frac{d}{2}}\exp\left(-\frac{|x-y|^2}{Ct}\right), \tag{4.18}$$

where κ_1 is a new exponent which is strictly larger than the original exponent κ. We can then redo the proof a second time and use the estimate (4.18) to obtain the inequality with an exponent κ_2 strictly larger than κ_1. An iteration of the argument shows that there exists an increasing sequence κ_n such that, for each $n \in \mathbb{N}$, the following estimate holds

$$\left|p(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y)\right| \leqslant Ct^{-\kappa_n}t^{-\frac{d}{2}}\exp\left(-\frac{|x-y|^2}{Ct}\right). \tag{4.19}$$

The sequence κ_n is defined inductively (see the formula (4.149)) and we can prove that it converges toward the value $\frac{1}{2}$; this is sufficient to prove the near optimal estimate stated in Theorem 4.1.1.

The third difficulty is the degenerate structure of the environment. It is treated by defining a renormalization structure for the infinite cluster which was first introduced in [19]: building upon standard results in supercritical percolation, we construct a partition of the lattice $\mathbb{Z}^d$ into cubes of different random sizes which are well-connected in the sense of Antal, Penrose and Pisztora (see [17, 200]), using a Calderón-Zygmund type stopping time argument. The sizes of the cubes of the partition are random variables which measure how close the geometry of the cluster is from the geometry of the lattice: in the regions where the sizes of the cubes are small, the cluster is well-behaved and its geometry is similar to the one of the Euclidean lattice, while in the regions where the sizes of the cubes are large, the geometry of the cluster is ill-behaved (see Figure 4.3). The probability to have a large cube in the partition is small and stretched exponential integrability estimates are available for these random variables (see Proposition 4.2.2 (iii) or [200]).

This partition provides a random scale above which the geometry of the infinite cluster is similar to the one of the Euclidean lattice and it allows to adapt the tools of functional analysis needed to perform the two-scale expansion to the percolation cluster. Similar strategies using renormalization techniques where used to study random walk on the supercritical percolation cluster and we refer for instance to the work of Barlow in [39], who established a Poincaré inequality on the percolation cluster, or to the one of Mathieu and Remy in [181].

The general strategy to study the random walk on the infinite cluster is thus to prove that there exists a random scale above which the geometry of the infinite cluster $\mathscr{C}_\infty$ is similar to the geometry of the lattice $\mathbb{Z}^d$, and to deduce from it that, above a random time which is related to the aforementioned random scale, the random walk has a behavior which is similar to the one of the random walk on $\mathbb{Z}^d$. As a consequence, most of the results described in this article only hold above a random scale (or random time) above which the infinite cluster has renormalized. Moreover, we need to appeal to a number of random scales (or random times) in the proofs, above which some analytical tools are available: the scale $\mathcal{M}_{\mathrm{reg}}$ above which a $C^{0,1}$-regularity theory is valid (see Theorem 4.1.3), the time $\mathcal{T}_{\mathrm{NA}}$ above which a Nash-Aronson estimate for the heat kernel is available (see Theorem 4.3.1) etc. For all these random scales and times, stretched exponential integrability estimates are valid.

This strategy describes the proof of Theorem 4.1.1. Once this result is established, Theorem 4.1.2, pertaining to the elliptic Green's function, can be deduced from it thanks to the Duhamel principle stated in (4.8). This is the subject of Section 4.5.

We complete this section by describing the content and purposes of Section 4.3. To perform the analysis described in the previous paragraphs, and in particular to prove that the function q defined (4.16) is a good approximation of the heat kernel p, one needs to have some control over the quantities at stake. In particular, it is useful to have a good control on the heat kernel p and its gradient ∇p. The first one is given by the article of Barlow [39], which provides Gaussian upper and lower bounds for the heat kernel p (see Theorem 4.3.1). For the gradient of the heat kernel, we expect to have a behavior similar to the one of the gradient of the heat kernel on $\mathbb{R}^d$, i.e., a $C^{0,1}$-regularity estimate of the form

$$|\nabla_x p(t,x,y)| \leqslant C t^{-\frac{d}{2}-\frac{1}{2}} \exp\left(-\frac{|x-y|^2}{Ct}\right).$$

Section 4.3 is devoted to proving a large-scale version of this estimate and is independent of Section 4.4 and Section 4.5. The precise statement established in this section is the following.

Theorem 4.1.3. *There exist an exponent $s(d,\mathfrak{p},\lambda) > 0$, a positive constant $C(d,\mathfrak{p},\lambda) < \infty$ such that for each point $x \in \mathbb{Z}^d$, there exists a non-negative random variable $\mathcal{M}_{\mathrm{reg}}(x)$ satisfying the stochastic integrability estimate*

$$\forall R \geqslant 0,\ \mathbb{P}\left(\mathcal{M}_{\mathrm{reg}}(x) \geqslant R\right) \leqslant C \exp\left(-\frac{R^s}{C}\right), \tag{4.20}$$

such that the following statement is valid: for every radius $r \geqslant \mathcal{M}_{\mathrm{reg}}(x)$, every point $y \in \mathscr{C}_\infty$ and every time $t \geqslant \max\left(4r^2, |x-y|\right)$, the following estimate holds,

$$\left\|\nabla_x p(t,\cdot,y)\right\|_{\underline{L}^2(B_r(x)\cap\mathscr{C}_\infty)} \leqslant C t^{-\frac{d}{2}-\frac{1}{2}} \exp\left(-\frac{|x-y|^2}{Ct}\right),$$

where the notation $\underline{L}^2\left(B_r(x)\cap\mathscr{C}_\infty\right)$ denotes the average L^2-norm over the set $B_r(x)\cap\mathscr{C}_\infty$ and is defined in (4.31).

Remark. By using the symmetry of the heat kernel, a similar regularity estimate holds for the gradient in the second variable: for each point $y \in \mathbb{Z}^d$, there exists a non-negative random variable $\mathcal{M}_{\mathrm{reg}}(y)$ satisfying the stochastic integrability estimate (4.20) such that for every radius $r \geqslant \mathcal{M}_{\mathrm{reg}}(y)$, every point $x \in \mathscr{C}_\infty$ and every time $t \geqslant \max\left(4r^2, |x-y|\right)$,

$$\left\|\nabla_y p(t,x,\cdot)\right\|_{\underline{L}^2(B_r(y)\cap\mathscr{C}_\infty)} \leqslant C t^{-\frac{d}{2}-\frac{1}{2}} \exp\left(-\frac{|x-y|^2}{Ct}\right).$$

The strategy of the proof of this result relies on tools from homogenization theory, in particular the two-scale expansion and the large-scale regularity theory. It is described at the beginning of Section 4.3.

4.1.3 Related results

Related results about the random conductance model

The random conductance model has been the subject of active research over the recent years, by various authors and under different assumptions over the law of the environment. In the case of uniform ellipticity, i.e., when the environment is allowed to take values in $[\lambda, 1]$, a quenched invariance principle is proved by Osada in [195] (in the continuous setting) and by Sidoravicius and Sznitman in [211] (in the discrete setting). Gaussian bounds on the heat kernel follow from [92]. This framework is the one of the theory of stochastic homogenization and we refer to Section 4.1.3 for further information.

In the setting when the conductances are only bounded from above, a quenched invariance principle was proved by Mathieu in [179] and by Biskup and Prescott in [55]. In the case when the conductances are bounded from below, a quenched invariance principle and heat kernel bounds are proved in [40] by Barlow and Deuschel. In [10], Andres, Barlow, Deuschel and Hambly established a quenched invariance principle in the general case when the conductances are allowed to take values in $[0, \infty)$.

The i.i.d. assumption on the environment can be relaxed: in [12], Andres, Deuschel and Slowik proved a quenched invariance principle for the random walk for general ergodic environment under the moment condition

$$\mathbb{E}\left[\mathbf{a}(e)^p\right] + \mathbb{E}\left[\mathbf{a}(e)^{-q}\right] < \infty \text{ for } p, q \in (1, \infty) \text{ satisfying } \frac{1}{p} + \frac{1}{q} < \frac{2}{d}. \tag{4.21}$$

We also refer to the works of Chiarni, Deuschel [77], Deuschel, Nguyen, Slowik [94] and Bella and Schäffner [44] for additional quenched invariance principles in degenerate ergodic environment. The case of ergodic, time-dependent, degenerate environment is investigated by Andres, Chiarini, Deuschel, and Slowik in [11] where they establish a quenched invariance principle under some moment conditions on the environment. More general models of random walks on percolation clusters with long range correlation, including random interlacements and level sets of the Gaussian free field, are studied by Procaccia, Rosenthal and Sapozhnikov in [203], where a quenched invariance principle is established.

The heat kernel has been studied under various assumptions on the environment: a first important property that needs to be investigated is the question of the existence of Gaussian lower and upper bounds. Such estimates are valid in the case of the percolation cluster presented in this article and were originally proved by Barlow in [39]. This result also holds when the conductances are bounded from below and we refer to the works of Mourrat [182] (Theorem 10.1 of the second arxiv version) and of Barlow, Deuschel [40]. It is also known that it cannot hold in full generality: in [50], Berger, Biskup, Hoffman and Kozma established that, when the law of the conductances has a fat tail at 0, the heat kernel can behave anomalistically due to trapping phenomenon (even though a quenched invariance principle still holds by [10]). We refer to the works of Barlow, Boukhadra [54] and Boukhadra [61, 62] for additional results in this direction. Gaussian estimates on the heat kernel for more general graphs were studied by Andres, Deuschel and Slowik in [14] and [15].

The question of Gaussian upper and lower bounds on the heat kernel is related, and in many situations equivalent, to the existence of a parabolic Harnack inequality (see for instance Delmotte [92]). On the percolation cluster, the parabolic Harnack inequality is established in [41]. We refer to the article of Andres, Deuschel, Slowik [13] for a proof of elliptic and parabolic Harnack inequalities on general graphs with unbounded weights, to the work of Sapozhnikov [209] for a proof of quenched heat kernel bounds and parabolic Harnack inequality for a general class of percolation models with long-range correlations on $\mathbb{Z}^d$ and to the articles of Chang [72] and Alves and Sapozhnikov [9] for similar results on loop soup models.

Results on the elliptic Green's function usually follow from the ones established on the parabolic Green's function, by an application of the formula (4.8) in dimension larger than 3. In dimension 2 the situation is different and requires separate considerations; in [16], Andres, Deuschel and Slowik characterize the asymptotics of the Green's function associated to the random walk killed upon exiting a ball under general assumptions on the environment.

Finally, we refer to [53] for a general review on the random conductance model.

Related result about stochastic homogenization

The theory of qualitative stochastic homogenization was developed in the 80's, with the works of Kozlov [160], Papanicolaou and Varadhan [198] and Yurinskiĭ [225] in the uniformly elliptic setting. Still in the uniformly elliptic setting, a quantitative theory of stochastic homogenization has been developed in the recent years up to the point that it is now well-understood thanks to the works of Gloria and Otto in [123, 124, 126, 125] and Gloria, Neukamm, Otto [121, 122], building upon the ideas of Naddaf and Spencer in [188]. These results have applications to random walks in random environment, as is explained in [102]. Another approach was initiated by Armstrong and Smart in [31], who extended the techniques of Avellaneda and Lin [33, 35] and the ones of Dal Maso and Modica [80, 81]. These results were then improved in [23, 24], and we refer to the monograph [25] for a detailed review of this approach.

The aforementioned works treated the case of uniformly elliptic environments and the question of the extension of the theory to degenerate environments has drawn some attention over the past few years. A number of results have been achieved and some of them are closely related to the works on the random conductance model presented in the previous section. In [193], Neukamm, Schäffner and Schlömerkemper proved Γ-convergence of the Dirichlet energy associated to some nonconvex energy functionals with degenerate growth. In [164], Lamacz, Neukamm and Otto studied a model of Bernoulli bond percolation, which is modified such that every bond in a fixed direction is declared open. In [108], Fleger, Heida and Slowik proved homogenization results for a degenerate random conductance model with long range jumps. In [43], Bella, Fehrman and Otto studied homogenization of degenerate environment under the moment condition (4.21) and established a first-order Liouville theorem as well as a large-scale $C^{1,\alpha}$-regularity estimate for $\mathbf{a}$-harmonic functions. In [117], Giunti, Höfer and Velázquez studied homogenization for the Poisson equation in a randomly perforated domain. In [19], Armstrong and the first author implemented the techniques of [25] to the percolation cluster to obtain quantitative homogenization results as well as a large-scale regularity theory.

4.1.4 Further outlook and conjecture

The results of this article present quantitative rates of convergence for the parabolic and elliptic Green's functions on the percolation cluster. We do not expect the result to be optimal: the quantitative rate of convergence $\frac{1}{2} - \delta$ and the stochastic integrability s in Theorem 4.1.1 can be improved and so is the case for Theorem 4.1.2. We expect the following conjecture to hold.

Conjecture 1. *Fix $s \in \left(0, \frac{2(d-1)}{d}\right)$, there exists a positive constant $C < \infty$ depending on the parameters $d, \mathfrak{p}, \lambda$ and s, such that, for each time $t > 0$ and each pair of points $x, y \in \mathbb{Z}^d$ such* *that* $|x-y| \leqslant t$, *conditionally on the event* $\{x, y \in \mathscr{C}_\infty\}$,

$$\left|p(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y)\right| \leqslant \begin{cases} \mathcal{O}_s\left(Ct^{-\frac{d}{2}-\frac{1}{2}}\exp\left(-\frac{|x-y|^2}{Ct}\right)\right) & \text{when } d \geqslant 3, \\ \mathcal{O}_s\left(C\log^{\frac{1}{2}}(1+t)\,t^{-\frac{3}{2}}\exp\left(-\frac{|x-y|^2}{Ct}\right)\right) & \text{when } d = 2, \end{cases}$$

where the notation $\mathcal{O}_s$ *is used to measure the stochastic integrability and is defined in Section 4.1.6. For the elliptic Green's function, a similar result holds:*

1. *In dimension* $d \geqslant 3$, *for each* $x, y \in \mathbb{Z}^d$, *conditionally on the event* $\{x, y \in \mathscr{C}_\infty\}$,

$$|g(x,y) - \bar{g}(x-y)| \leqslant \mathcal{O}_s\left(C|x-y|^{1-d}\right).$$

   where the function $\bar{g}$ *is defined in the equation (4.9).*

2. *In dimension* 2, *for each* $y \in \mathbb{Z}^d$, *conditionally on the event* $\{y \in \mathscr{C}_\infty\}$, *the limit*

$$K(y) := \lim_{x \to \infty} g(x,y) - \bar{g}(x-y),$$

   exist, is finite almost surely and satisfies the stochastic integrability estimate

$$|K(y)| \leqslant \mathcal{O}_s(C).$$

   Moreover, for every $x \in \mathbb{Z}^d$, *conditionally on the event* $\{x, y \in \mathscr{C}_\infty\}$, *one has*

$$|g(x,y) - \bar{g}(x-y) - K(y)| \leqslant \mathcal{O}_s\left(C\log^{\frac{1}{2}}(1+|x-y|)\,|x-y|^{-1}\right).$$

Remark. This statement cannot be stated with a minimal scale as in Theorems 4.1.1 and 4.1.2. This is due to the fact that the estimates scale optimally in time or space; the best possible statements involving a minimal scale are the ones of Theorems 4.1.1 and 4.1.2.

This result can be conjectured from the theory of stochastic homogenization in the uniformly elliptic setting (see [25, Theorem 9.11 and Corollary 9.12]). There is one main difference between the results in the uniformly elliptic setting and in the percolation setting, which is the stochastic integrability: we expect that the stochastic integrability will be reduced by a factor $(d-1)/d$. This is expected because of a surface order large deviation effect which can be heuristically explained as follows. In the uniformly elliptic setting and in a given ball B_R, to design a bad environment, i.e., an environment on which no good control on the heat kernel is valid, it is necessary to have a number of ill-behaved edges of order of the volume of the ball. In the percolation setting, one can design a bad environment with a number of ill-behaved edges of order of the surface of the ball: given a ball of size R, it is possible to disconnect it into two half-balls with cR^{d-1} closed edges. This should result in a deterioration of the stochastic integrability by a factor $(d-1)/d$.

The conjecture improves Theorems 4.1.1 and 4.1.2 in two distinct directions: the spatial scaling, where the coefficient $1/2-\delta$ is replaced by $1/2$ for the heat kernel and the coefficient $1-\delta$ is replaced by 1 for the elliptic Green's function, and the stochastic integrability, where the exponent s can take any value in the interval $\left(0, \frac{2(d-1)}{d}\right)$. We believe that the two improvements should follow from different techniques: for the spatial integrability, we think that it should follow by an adaptation of the techniques developed in [25, Chapter 9]. The improvement of the stochastic integrability seems to be a much harder problem which requires separate considerations and should rely on a precise understanding of the geometry of the percolation clusters.

We complete this section by mentioning that the results of this article pertain to the variable speed random walk, but similar results, with similar proofs, should hold for other related models of random walk on the infinite cluster such as the constant speed random walk and the simple random walk. This choice is motivated by the fact that the generator of the VSRW, written in (4.1), is more convenient to work with than the ones of the CSRW and the SRW, which simplifies the analysis.

4.1.5 Organization of the article

The rest of this article is organized as follows. The remaining section of this introduction is devoted to the presentation of some useful notations.

In Section 4.2, we record some preliminary results, including some results from the theory of quantitative stochastic homogenization on the infinite cluster from [19, 83, 134]: the quantitative sublinearity of the corrector and a quantitative estimate to control the H^{-1}-norm of the centered flux. In Section 4.3, we recall the Gaussian bounds on the heat kernel which were established by Barlow in [39] and establish a large-scale $C^{0,1}$-regularity theory for the heat kernel.

In Section 4.4, we establish Theorem 4.1.1. The proof is organized in three subsections: Section 4.4.1 is devoted to the proofs of three regularization steps, which can be seen as a preparation for the two-scale expansion in Section 4.4.2. The heart of the proof is Section 4.4.2, where we perform the two-scale expansion. In Section 4.4.3, we post-process the result from Section 4.4.2 and deduce the result of Theorem 4.1.1.

In Section 4.5, we use Theorem 4.1.1 to prove the homogenization of the elliptic Green function, i.e., Theorem 4.1.2.

Appendix 4.A and Appendix 4.B are devoted respectively to two technical estimates: a concentration inequality for the density of the infinite cluster in a cube and the proof of the quantitative estimate of the weak H^{-1}-norm of the centered flux which is stated in Section 4.2.3.

4.1.6 Notation and assumptions

General notations and assumptions

We let $\mathbb{Z}^d$ be the standard d-dimensional hypercubic lattice and $E_d := \left\{ \{x,y\} \, : \, x,y \in \mathbb{Z}^d, |x-y| = 1 \right\}$ denote the set of bonds. We also denote by $\overrightarrow{E_d}$ the set of *oriented* bonds, or *edges*, of $\mathbb{Z}^d$. We use the notation $\{x,y\}$ to refer to a bond and (x,y) to refer to an edge.

We denote the canonical basis of $\mathbb{R}^d$ by $\{e_1, \ldots, e_d\}$. For a vector $p \in \mathbb{R}^d$ and an integer $i \in \{1, \ldots, d\}$, we denote by $[p]_i$ its ith-component, i.e., $p = ([p]_1, \ldots, [p]_d)$. For $x, y \in \mathbb{Z}^d$, we write $x \sim y$ if x and y are nearest neighbors. We usually denote a generic edge by e. We fix an ellipticity parameter $\lambda \in (0,1]$ and denote by Ω the set of all functions $\mathbf{a} : E_d \to \{0\} \cup [\lambda, 1]$, i.e., $\Omega = (\{0\} \cup [\lambda, 1])^{E_d}$ and we denote by $\mathbf{a}$ a generic element of Ω. The Borel σ-algebra on Ω is denoted by $\mathcal{F}$. For each $U \subseteq \mathbb{Z}^d$, we let $\mathcal{F}(U) \subseteq \mathcal{F}$ denote σ-algebra generated by the projections $\mathbf{a} \mapsto \mathbf{a}(\{x,y\})$, for $x, y \in U$ with $x \sim y$.

We fix an i.i.d. probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$, that is, a measure of the form $\mathbb{P} = \mathbb{P}_0^{E_d}$ where $\mathbb{P}_0$ is a measure of probability supported in the set $\{0\} \cup [\lambda, 1]$ with the property that, for any fixed bond e,

$$\mathfrak{p} := \mathbb{P}_0 \left[\mathbf{a}(e) \neq 0 \right] > \mathfrak{p}_c(d),$$

where $\mathfrak{p}_c(d)$ is the bond percolation threshold for the lattice $\mathbb{Z}^d$. We say that a bond e is *open* if $\mathbf{a}(e) \neq 0$ and *closed* if $\mathbf{a}(e) = 0$. A connected component of open edges is called a *cluster*. Under the assumption $\mathfrak{p} > \mathfrak{p}_c(d)$, there exists almost surely a unique maximal infinite cluster, which is denoted by $\mathscr{C}_\infty$ and we also note $\theta(p) := \mathbb{P}[0 \in \mathscr{C}_\infty]$. From now on, we always consider environments $\mathbf{a} \in \Omega$ such that there exists a unique infinite cluster of open edges. We denote by $\mathbb{E}$ the expectation with respect to the measure $\mathbb{P}$.

Notation of $\mathcal{O}_s$

For a parameter $s > 0$, we use the notation $\mathcal{O}_s$ to measure the stochastic integrability of random variables. It is defined as follows, given a random variable X, we write

$$X \leqslant \mathcal{O}_s(\theta) \text{ if and only if } \mathbb{E}\left[\exp((\theta^{-1}X)_+^s)\right] \leqslant 2, \tag{4.22}$$

where $(\theta^{-1}X)_+$ means $\max\{\theta^{-1}X, 0\}$. From the inequality (4.22) and the Markov's inequality, one deduces the following estimate for the tail of the random variable X: for all $x > 0$, $\mathbb{P}[X \geqslant \theta x] \leqslant 2\exp(-x^s)$.

Given a random variable X satisfying the identity $X \leqslant \mathcal{O}_s(\theta)$, one can check that, for each $\lambda \in \mathbb{R}^+$, one has $\lambda X \leqslant \mathcal{O}_s(\lambda\theta)$. Additionally, one can reduce the stochastic integrability parameter s according to the following statement: for each $s' \in (0, s]$, there exists a constant $0 < C_{s'} < \infty$ such that $X \leqslant \mathcal{O}_{s'}(C_{s'}\theta)$.

To estimate the stochastic integrability of a sum of random variables, we use the following estimate, which can be found in [25, Lemma A.4 of Appendix A]: for each exponent $s > 0$ there exists a positive constant $C_s < \infty$ such that for any measure space $(E, \mathcal{S}, m)$ and any family of random variables $\{X(z)\}_{z\in E}$, one has

$$\forall z \in E, X(z) \leqslant \mathcal{O}_s(\theta(z)) \Longrightarrow \int_E X(z)m(dz) \leqslant \mathcal{O}_s\left(C_s \int_E \theta(z)m(dz)\right). \tag{4.23}$$

The previous statement allows to estimate the stochastic integrability of a sum of random variables: given $X_1, \ldots, X_n$ a collection of non-negative random variables and $C_1, \ldots, C_n$ a collection of non-negative constants such that, for any $i \in \{1, \ldots, n\}$, $X_i \leqslant \mathcal{O}_s(C_i)$, one has the estimate

$$\sum_{i=1}^n X_i \leqslant \mathcal{O}_s\left(C_s \sum_{i=1}^n C_i\right). \tag{4.24}$$

The following lemma is useful to construct minimal scales.

Lemma 4.1.1. *[19, Lemme 2.2] Fix $K \geqslant 1, s > 0$ and $\beta > 0$ and let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of non-negative random variables satisfying the inequality $X_n \leqslant \mathcal{O}_s(K3^{-n\beta})$ for every $n \in \mathbb{N}$. There exists a positive constant $C(s,\beta,K) < \infty$ such that the random scale $M := \sup\{3^n \in \mathbb{N} : X_n \geqslant 1\}$ satisfies the stochastic integrability estimate $M \leqslant \mathcal{O}_{\beta s}(C)$.*

Topology, functions and integration

For every subset $V \subseteq \mathbb{Z}^d$ and every environment $\mathbf{a} \in \Omega$, we consider two sets of bonds $E_d(V)$ and $E_d^{\mathbf{a}}(V)$. The first one is inherited from the set of bonds E_d of $\mathbb{Z}^d$, the second one is inherited from the bonds of non-zero conductance of the environment $\mathbf{a}$. They are defined by the formulas

$$E_d(V) := \{\{x,y\} : x, y \in V, x \sim y\}, \qquad E_d^{\mathbf{a}}(V) := \{\{x,y\} : x, y \in V, x \sim y, \mathbf{a}(\{x,y\}) \neq 0\}.$$

We similarly define the set of edges $\overrightarrow{E_d}(V)$ and $\overrightarrow{E_d^{\mathbf{a}}}(V)$.

The *interiors* of a set V with respect to $E_d(V)$ and $E_d^{\mathbf{a}}(V)$ are defined by the formulas

$$\mathrm{int}(V) := \{x \in V : y \sim x \implies y \in V\}, \qquad \mathrm{int}_{\mathbf{a}}(V) := \{x \in V : y \sim x, \mathbf{a}(\{x,y\}) \neq 0 \implies y \in V\},$$

and the *boundaries* of V are defined by $\partial V := V \backslash \mathrm{int}(V)$ and $\partial_{\mathbf{a}} V := V \backslash \mathrm{int}_{\mathbf{a}}(V)$. The cardinality of a subset $V \subseteq \mathbb{Z}^d$ is denoted by $|V|$ and called the *volume* of V. Given two sets $U, V \subseteq \mathbb{Z}^d$, we define the distance between U and V according to the formula $\mathrm{dist}(U,V) := \min_{x\in U, y\in V} |x-y|$ and the distance of a point $x \in \mathbb{Z}^d$ to a set $V \subseteq \mathbb{Z}^d$ by the notation $\mathrm{dist}(x,V) := \min_{y \in V} |x-y|$.

For a subset $V \subseteq \mathbb{Z}^d$, the spaces of functions with zero boundary condition are defined by

$$C_0(V) := \{v : V \to \mathbb{R} : v = 0 \text{ on } \partial V\}, \qquad C_0^{\mathbf{a}}(V) := \{v : V \to \mathbb{R} : v = 0 \text{ on } \partial_{\mathbf{a}} V\}. \tag{4.25}$$

Given a subset $U \subseteq \mathbb{Z}^d$ and a function $u : \mathscr{C}_\infty \cap U \to \mathbb{R}$ (resp. a function $F : E_d^{\mathbf{a}}(\mathscr{C}_\infty \cap U) \to \mathbb{R}$), the integration over the set $\mathscr{C}_\infty \cap U$ (resp. over $E_d^{\mathbf{a}}(\mathscr{C}_\infty \cap U)$) is denoted by

$$\int_{\mathscr{C}_\infty \cap U} u := \sum_{x \in \mathscr{C}_\infty \cap U} u(x), \qquad \text{resp.} \int_{\mathscr{C}_\infty \cap U} F := \sum_{e \in E_d^{\mathbf{a}}(\mathscr{C}_\infty \cap U)} F(e), \tag{4.26}$$

which means that we only integrate on the vertices (resp. open bonds) of the infinite cluster $\mathscr{C}_\infty$. We extend this notation to the setting of vector-valued functions $u : \mathscr{C}_\infty \cap U \to \mathbb{R}^n$. We also let $(u)_V := \frac{1}{|V|} \int_V u$ denote the mean of the function u over the finite subset $V \subseteq \mathbb{Z}^d$.

Given a subset $V \subseteq \mathbb{Z}^d$, a vector field is a function $\overrightarrow{F} : \overrightarrow{E_d}(V) \to \mathbb{R}$ satisfying the anti-symmetry property

$$\overrightarrow{F}(\{x,y\}) = -\overrightarrow{F}(\{y,x\}).$$

For $y \in \mathbb{Z}^d$ and $r > 0$, we denote by $B_r(y)$ the discrete Euclidean ball of radius $r > 0$ and center y; we often write B_r in place of $B_r(0)$. A *cube* Q is a subset of $\mathbb{Z}^d$ of the form

$$Q := \mathbb{Z}^d \cap \left(z + [-N,N]^d\right).$$

We define the *center* and the *size* of the cube given in the previous display above to be the point z and the integer N respectively. The size of the cube Q is denoted by $\mathrm{size}(Q)$. Given an integer $n \in \mathbb{N}$, we use the non-standard convention of denoting by nQ the cube

$$nQ := \mathbb{Z}^d \cap \left(z + [-nN, nN]^d\right). \tag{4.27}$$

A *triadic cube* is a cube of the form

$$\square_m(z) := z + \left(-\frac{3^m}{2}, \frac{3^m}{2}\right)^d, \; z \in 3^m \mathbb{Z}^d, \, m \in \mathbb{N}.$$

We usually write $\square_m := \square_m(0)$. Additionally, we note that $\mathrm{size}(\square_m) = 3^m$, denote by $\mathcal{T}_m := \left\{z + \square_m : z \in 3^m \mathbb{Z}^d\right\}$ the set of triadic cubes of size 3^m and by $\mathcal{T}$ the set of all triadic cubes, i.e., $\mathcal{T} := \cup_{m \in \mathbb{N}} \mathcal{T}_m$.

Discrete analysis and function spaces

In this article, we consider two types of objects: functions defined in the continuous space $\mathbb{R}^d$ and functions defined on the discrete space $\mathbb{Z}^d$.

Notations for discrete functions. Given a discrete subset $U \subseteq \mathbb{Z}^d$, an environment $\mathbf{a}$ such that there exists an infinite cluster $\mathscr{C}_\infty$ of open edges, and a function $u : \mathscr{C}_\infty \cap U \to \mathbb{R}$, we define its gradient ∇u to be the vector field defined on $\overrightarrow{E_d}$ by, for each edge $e = (x,y) \in \overrightarrow{E_d}$,

$$\nabla u(e) := \begin{cases} u(y) - u(x) & \text{if } x, y \in \mathscr{C}_\infty \text{ and } \mathbf{a}(\{x,y\}) \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.28}$$

For each $x \in \mathscr{C}_\infty$, we also define the norm of the gradient that $|\nabla u|(x) := \sum_{y \sim x} |u(y) - u(x)|$. We frequently abuse notation and write $|\nabla u(x)|$ instead of $|\nabla u|(x)$.

For a vector field $\overrightarrow{F} : \overrightarrow{E_d} \to \mathbb{R}$, we define the discrete divergence operator according to the formula, for each $x \in \mathbb{Z}^d$,

$$\nabla \cdot \overrightarrow{F}(x) := \sum_{y \sim x} \overrightarrow{F}(x,y).$$

By the discrete integration by parts, one has, for any discrete set $U \subseteq \mathbb{Z}^d$, any functions $v \in C_0^{\mathbf{a}}(\mathscr{C}_\infty \cap U)$ and $u : \mathscr{C}_\infty \cap U \to \mathbb{R}$,

$$\int_{\mathscr{C}_\infty \cap U} \nabla v \cdot \mathbf{a} \nabla u := \int_{\mathscr{C}_\infty \cap U} \mathbf{a}(e) \nabla v(e) \nabla u(e) = \int_{\mathscr{C}_\infty \cap U} v(-\nabla \cdot \mathbf{a} \nabla u), \tag{4.29}$$

where the finite difference elliptic operator $-\nabla \cdot \mathbf{a} \nabla$ is defined in (4.1).

For $p \in [1, \infty)$, we define the $L^p(\mathscr{C}_\infty \cap U)$-norm and the normalized $\underline{L}^p(\mathscr{C}_\infty \cap U)$-norm by the formulas

$$\|u\|_{L^p(\mathscr{C}_\infty \cap U)} := \left(\int_{\mathscr{C}_\infty \cap U} |u|^p \right)^{\frac{1}{p}}, \qquad \|u\|_{\underline{L}^p(\mathscr{C}_\infty \cap U)} := \left(\frac{1}{|\mathscr{C}_\infty \cap U|} \int_{\mathscr{C}_\infty \cap U} |u|^p \right)^{\frac{1}{p}}. \tag{4.30}$$

We also define the $L^p(\mathscr{C}_\infty \cap U)$-norm and the normalized $\underline{L}^p(\mathscr{C}_\infty \cap U)$-norm of the gradient of a function $u : \mathscr{C}_\infty \cap U \to \mathbb{R}$ by the formulas

$$\|\nabla u\|_{L^p(\mathscr{C}_\infty \cap U)} := \left(\int_{\mathscr{C}_\infty \cap U} |\nabla u|^p \right)^{\frac{1}{p}}, \qquad \|\nabla u\|_{\underline{L}^p(\mathscr{C}_\infty \cap U)} := \left(\frac{1}{|\mathscr{C}_\infty \cap U|} \int_{\mathscr{C}_\infty \cap U} |\nabla u|^p \right)^{\frac{1}{p}}. \tag{4.31}$$

We define the normalized discrete Sobolev norm $\underline{W}^{1,p}(\mathscr{C}_\infty \cap U)$ by

$$\|u\|_{\underline{W}^{1,p}(\mathscr{C}_\infty \cap U)} := |\mathscr{C}_\infty \cap U|^{-\frac{1}{d}} \|u\|_{\underline{L}^p(\mathscr{C}_\infty \cap U)} + \|\nabla u\|_{\underline{L}^p(\mathscr{C}_\infty \cap U)}, \tag{4.32}$$

and the dual norm $\underline{W}^{-1,p}(\mathscr{C}_\infty \cap U)$,

$$\|u\|_{\underline{W}^{-1,p}(\mathscr{C}_\infty \cap U)} := \sup_{v \in C_0^{\mathbf{a}}(\mathscr{C}_\infty \cap U), \|v\|_{\underline{W}^{1,p'}(\mathscr{C}_\infty \cap U)} \leq 1} \frac{1}{|\mathscr{C}_\infty \cap U|} \int_{\mathscr{C}_\infty \cap U} uv, \tag{4.33}$$

with $\frac{1}{p'} + \frac{1}{p} = 1$. We use the notation $\underline{H}^1(\mathscr{C}_\infty \cap U) := \underline{W}^{1,2}(\mathscr{C}_\infty \cap U)$ and $\underline{H}^{-1}(\mathscr{C}_\infty \cap U) := \underline{W}^{-1,2}(\mathscr{C}_\infty \cap U)$.

For a function $u : \mathbb{Z}^d \to \mathbb{R}$ and a vector $h \in \mathbb{Z}^d$, we denote by $T_h(u) := u(\cdot + h)$ the translation and by $\mathcal{D}_{e_k}$ the finite difference operator defined by, for any function $u : \mathbb{Z}^d \to \mathbb{R}$,

$$\mathcal{D}_{e_k} u := \begin{cases} \mathbb{Z}^d & \to & \mathbb{R}, \\ x & \mapsto & T_{e_k}(u)(x) - u(x). \end{cases}$$

We also define the vector-valued finite difference operator $\mathcal{D}u := (\mathcal{D}_{e_1}u, \mathcal{D}_{e_2}u, \cdots, \mathcal{D}_{e_d}u)$. This definition has two main differences with the gradient on graph defined in (4.28): it is defined on the vertices of $\mathbb{Z}^d$ (not on the edges) and it is vector-valued. This second definition of discrete derivative is introduced because it is convenient in the two-scale expansion (see (4.107)).

Given an environment $\mathbf{a} \in \Omega$, and a function $u : \mathscr{C}_\infty \to \mathbb{R}$, we define the functions $\mathbf{a}\mathcal{D}u$ and $\mathbf{1}_{\{\{\mathbf{a}\neq 0\}\}}\mathcal{D}u$ by, for each $x \in \mathscr{C}_\infty$,

$$\begin{aligned} \mathbf{a}\mathcal{D}u(x) &= (\mathbf{a}(\{x, x+e_1\})\mathcal{D}_{e_1}u(x), \ldots, \mathbf{a}(\{x, x+e_d\})\mathcal{D}_{e_d}u(x)), \\ \mathbf{1}_{\{\{\mathbf{a}\neq 0\}\}}\mathcal{D}u(x) &= \left(\mathbf{1}_{\{\{\mathbf{a}(\{x,x+e_1\})\neq 0\}\}}\mathcal{D}_{e_1}u(x), \ldots, \mathbf{1}_{\{\{\mathbf{a}(\{x,x+e_d\})\neq 0\}\}}\mathcal{D}_{e_d}u(x)\right). \end{aligned} \tag{4.34}$$

We extend these functions to the entire space $\mathbb{Z}^d$ by setting, for each point $x \in \mathbb{Z}^d \setminus \mathscr{C}_\infty$,

$$\mathbf{a}\mathcal{D}u(x) = \mathbf{1}_{\{\{\mathbf{a}\neq 0\}\}}\mathcal{D}u(x) = 0.$$

It is natural to introduce the dual operator $\mathcal{D}^*_{e_k}u := T_{-e_k}(u) - u$ and the divergence $\mathcal{D}^* \cdot$ defined by, for any vector-valued function $\widetilde{F} : \mathbb{Z}^d \to \mathbb{R}^d$, $\widetilde{F} = \left(\widetilde{F}_1, \widetilde{F}_2, \cdots \widetilde{F}_d\right)$,

$$\mathcal{D}^* \cdot \widetilde{F}(x) = \sum_{k=1}^d \mathcal{D}^*_{e_k}\widetilde{F}_k(x).$$

By the discrete integration by parts, one has the equality, for any $v \in C_0^{\mathbf{a}}(\mathscr{C}_\infty \cap U)$,

$$\int_{\mathscr{C}_\infty \cap U} v\,(\mathcal{D}^* \cdot \mathbf{a}\mathcal{D}u) = \int_{\mathscr{C}_\infty \cap U} \mathcal{D}v \cdot \mathbf{a}\mathcal{D}u. \tag{4.35}$$

In fact one can check that the identity $-\nabla \cdot \mathbf{a}\nabla = \mathcal{D}^* \cdot \mathbf{a}\mathcal{D}$ holds, which allows to interchange the two notations.

Moreover, given a vector $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, we denote by $|x| = \left(\sum_{j=1}^d x_j^2\right)^{\frac{1}{2}}$ its norm. This allows to extend the definition of the Sobolev norms (4.30), (4.32) and (4.33) to vector-valued functions, and we note that

$$c\,\|u\|_{\underline{W}^{1,p}(\mathscr{C}_\infty \cap U)} \leqslant |\mathscr{C}_\infty \cap U|^{-\frac{1}{d}}\,\|u\|_{\underline{L}^p(\mathscr{C}_\infty \cap U)} + \left\|\mathbf{1}_{\{\{\mathbf{a}\neq 0\}\}}\mathcal{D}u\right\|_{\underline{L}^p(\mathscr{C}_\infty \cap U)} \leqslant C\,\|u\|_{\underline{W}^{1,p}(\mathscr{C}_\infty \cap U)},$$

for some constants c, C which only depend on the dimension d.

Notations for continuous functions. We use the notations ∂_k, ∇, Δ for the standard derivative, gradient and Laplacian on $\mathbb{R}^d$, which are only applied to smooth functions. It will always be clear from context whether we refer to the continuous or discrete derivatives. We sometime slightly abuse the notation and denote by $|\nabla^k \eta|$ the norm of k-th derivatives of the function η.

Notations for parabolic functions. For $r > 0$, we define the time interval $I_r := \left(-r^2, 0\right]$ and $I_r(t) := t + \left(-r^2, 0\right]$. We frequently use the parabolic cylinders $I_r \times B_r$ and $I_r \times (\mathscr{C}_\infty \cap B_r)$ and define their volumes by

$$|(I_r \times B_r)| = r^2 \times |B_r| \qquad \text{and} \qquad |(I_r \times (\mathscr{C}_\infty \cap B_r))| = r^2 \times |\mathscr{C}_\infty \cap B_r|.$$

Given a function $u : I_r \times B_r \to \mathbb{R}$ (resp. $v : I_r \times (\mathscr{C}_\infty \cap B_r) \to \mathbb{R}$), we define the integrals

$$\int_{I_r \times B_r} u := \int_{-r^2}^0 \int_{B_r} u(t,x)\,dx\,dt \quad \text{and} \quad \int_{I_r \times (\mathscr{C}_\infty \cap B_r)} v := \int_{-r^2}^0 \int_{\mathscr{C}_\infty \cap B_r} v(t,x)\,dx\,dt,$$

and denote the mean of these functions by the notation

$$(u)_{I_r\times B_r} := \frac{1}{|(I_r\times B_r)|}\int_{-r^2}^{0}\int_{B_r} u(t,x)\,dx\,dt,$$
$$(v)_{I_r\times(\mathscr{C}_\infty\cap B_r)} := \frac{1}{|(I_r\times(\mathscr{C}_\infty\cap B_r))|}\int_{-r^2}^{0}\int_{B_r} v(t,x)\,dx\,dt.$$

Given a finite subset $V\subseteq\mathbb{Z}^d$ or $V\subseteq\mathbb{R}^d$, we denote by $\partial_\sqcup(I_r\times V)$ the parabolic boundary of the cylinder $I_r\times V$ defined by the formula

$$\partial_\sqcup(I_r\times V) := (I_r\times\partial V)\cup\left(\{-r^2\}\times V\right).$$

Given a real number $p\geqslant 1$ and a Lebesgue-measurable function $u: I_r\times V\to\mathbb{R}^d$, we define the norm $L^p\left(I_r\times V\right)$ and the normalized norm $\underline{L}^p\left(I_r\times V\right)$ according to the formulas

$$\|u\|_{L^p(I_r\times V)} := \left(\int_{-r^2}^{0}\|u(t,\cdot)\|_{L^p(V)}^p\,dt\right)^{\frac{1}{p}} \text{ and } \|u\|_{\underline{L}^p(I_r\times V)} := \left(r^{-2}\int_{-r^2}^{0}\|u(t,\cdot)\|_{\underline{L}^p(V)}^p\,dt\right)^{\frac{1}{p}}.$$

These notations are extended to the gradient of a function $u: I_r\times V\to\mathbb{R}^d$ by the formulas

$$\|\nabla u\|_{L^p(I_r\times V)} := \left(\int_{-r^2}^{0}\|\nabla u(t,\cdot)\|_{L^p(V)}^p\,dt\right)^{\frac{1}{p}} \text{ and } \|\nabla u\|_{\underline{L}^p(I_r\times V)} := \left(r^{-2}\int_{-r^2}^{0}\|\nabla u(t,\cdot)\|_{\underline{L}^p(V)}^p\,dt\right)^{\frac{1}{p}}.$$

Given a real number $q\geqslant 1$, we also define the space $L^q\left(I_r;W^{-1,p}(V)\right)$ by

$$L^q\left(I_r;W^{-1,p}(V)\right) := \left\{u: I_r\times V\to\mathbb{R}^n \,:\, \int_{-r^2}^{0}\|u\|_{\underline{W}^{-1,p}(V)}^q\,dt<\infty\right\},$$

and we equip this space with the normalized norm defined by

$$\|u\|_{\underline{L}^q(I_r;\underline{W}^{-1,p}(V))} := \left(r^{-2}\int_{-r^2}^{0}\|u(t,\cdot)\|_{\underline{W}^{-1,p}(V)}\,dt\right)^{\frac{1}{q}}.$$

We define the parabolic Sobolev space $W^{1,p}_{\mathrm{par}}(I_r\times V)$ to be the set of measurable functions $u: I_r\times V\to\mathbb{R}$ such that the time derivative $\partial_t u$, understood in the sense of distributions, belongs to the space $W^{-1,p'}(I_r\times V)$ with $\frac{1}{p}+\frac{1}{p'}=1$, i.e.,

$$W^{1,p}_{\mathrm{par}}(I_r\times V) := \left\{u\in L^p\left(I_r\times V\right) \,:\, \partial_t u\in L^{p'}\left(I_r;W^{-1,p'}(V)\right)\right\}.$$

We also make use of the notations $H^1_{\mathrm{par}}\left(I_r\times V\right) := W^{1,2}_{\mathrm{par}}(I_r\times V)$ for the H^1 parabolic space and $\underline{L}^2\left(I_r;\underline{H}^{-1}(V)\right) := \underline{L}^2\left(I_r;\underline{W}^{-1,2}(V)\right)$.

Convention for constants, exponents and minimal scales/times

Throughout this article, the symbols c and C denote positive constants which may vary from line to line. These constants may depend only on the dimension d, the ellipticity λ and the probability $\mathfrak{p}$. Similarly we use the symbols α, β, γ, δ to denote positive exponents which depend only on d, λ and $\mathfrak{p}$. Usually, we use the letter C for large constants (whose value is expected to belong to $[1,\infty)$) and c for small constants (whose value is expected to be in $(0,1]$). The values of the exponents α, β, γ, δ are always expected to be small. When the constants and exponents depend on other parameters, we write it explicitly and use the notation $C := C(d,\mathfrak{p},t)$ to mean that the constant C depends on the parameters $d,\mathfrak{p}$ and t. We also assume that all the minimal scales and times which appear in this article are larger than 1.

4.2 Preliminaries

In this section, we collect a few results from the theory of supercritical percolation which are important tools in the establishment of Theorems 4.1.1 and 4.1.2.

4.2.1 Supercritical percolation

A partition of good cubes

An important step to prove results on the behavior of the random walk on the infinite cluster $\mathscr{C}_\infty$ consists in understanding the geometry of this cluster. A general picture to keep in mind is that the geometry of $\mathscr{C}_\infty$ is similar, at least on large scales, to the one of the Euclidean lattice $\mathbb{Z}^d$. To give a precise mathematical meaning to this statement, the common strategy is to implement a renormalization structure for the infinite cluster. In this article, we use a strategy, which was first introduced by Armstrong and the first author in [19]. It relies on the following geometric definition and lemma which are due to Penrose and Pisztora [200].

Definition 4.2.1 (Pre-good cube)**.** We say that a discrete cube $\square \subseteq \mathbb{Z}^d$ of size N is *pre-good* if:

(i) There exists a cluster of open edges which intersects the $2d$ faces of the cube $\square$. This cluster is denoted by $\mathscr{C}_* (\square)$;

(ii) The diameter of all the other clusters is smaller than $\frac{N}{10}$.

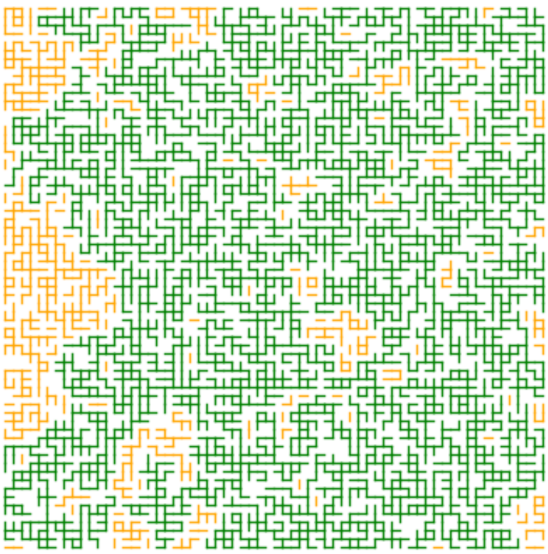

Figure 4.2: A pre-good cube, the cluster $\mathscr{C}_*(\square)$ is drawn in green and the clusters in yellow are the small clusters.

We then upgrade this definition into the following definition of good cubes.

Definition 4.2.2 (Good cube)**.** We say that a discrete cube $\square \subseteq \mathbb{Z}^d$ of size N is *good* if:

(i) The cube $\square$ is pre-good;

(ii) Every cube $\square'$ whose size is between $N/10$ and N and which has non-empty intersection with $\square$ is also pre-good.

We note that the event "the cube $\square$ is good" is $\mathcal{F}(2\square)$-measurable. The main reason to use good cubes instead of pre-good cubes is that they satisfy the following connectivity property, which can be obtained from straightforward geometric considerations and whose proof can be found in [19, Lemma 2.8].

Lemma 4.2.1 (Connectivity property)**.** *Let $\square_1, \square_2$ be two cubes of $\mathbb{Z}^d$ which are neighbors, i.e., which satisfy*

$$\operatorname{dist}(\square_1, \square_2) \leqslant 1,$$

which have comparable size in the sense that

$$\frac{1}{3} \leqslant \frac{\operatorname{size}(\square_1)}{\operatorname{size}(\square_2)} \leqslant 3,$$

and which are both good. Then there exists a cluster $\mathscr{C}$ such that

$$\mathscr{C}_*(\square_1) \cup \mathscr{C}_*(\square_2) \subseteq \mathscr{C} \subseteq \square_1 \cup \square_2.$$

The main interest in these definitions is that in the supercritical phase $\mathfrak{p} > \mathfrak{p}_c(d)$, the probability of a cube to be good is exponentially close to 1 in the size of the cube. Such a result is stated in the following proposition and is a direct consequence [201, Theorem 3.2] and [200, Theorem 5].

Proposition 4.2.1. *Consider a Bernoulli bond percolation of probability $\mathfrak{p} \in (\mathfrak{p}_c(d), 1]$. Then there exists a positive constant $C(d, \mathfrak{p}) < \infty$ such that, for every cube $\square \subseteq \mathbb{Z}^d$ of size N,*

$$\mathbb{P}[\square \textit{ is good}] \geqslant 1 - C \exp\left(-C^{-1} N\right). \tag{4.36}$$

The renormalization structure we want to implement relies on the observation that $\mathbb{Z}^d$ can be partitioned into good cubes of varying sizes. Thanks to the exponential stochastic integrability obtained by Penrose and Pisztora and stated in Proposition 4.2.1, we are able to build such a partition. The precise statement is given in the following proposition.

Proposition 4.2.2 (Propositions 2.1 and 2.4 of [19])**.** *Under the assumption $\mathfrak{p} > \mathfrak{p}_c$, $\mathbb{P}$–almost surely, there exists a partition $\mathcal{P}$ of $\mathbb{Z}^d$ into triadic cubes with the following properties:*

(i) All the predecessors of elements of $\mathcal{P}$ are good cubes, i.e., for every pair of triadic cubes $\square, \square' \in \mathcal{T}$, one has the property

$$\square' \in \mathcal{P} \text{ and } \square' \subseteq \square \implies \square \text{ is good.}$$

(ii) Neighboring elements of $\mathcal{P}$ have comparable sizes: for $\square, \square' \in \mathcal{P}$ such that $\operatorname{dist}(\square, \square') \leqslant 1$, we have

$$\frac{1}{3} \leqslant \frac{\operatorname{size}(\square')}{\operatorname{size}(\square)} \leqslant 3.$$

(iii) Estimate for the coarseness of $\mathcal{P}$: if we denote by $\square_{\mathcal{P}}(x)$ the unique element of $\mathcal{P}$ containing a given point $x \in \mathbb{Z}^d$, then there exists a constant $C(\mathfrak{p}, d) < \infty$ such that,

$$\operatorname{size}(\square_{\mathcal{P}}(x)) \leqslant \mathcal{O}_1(C). \tag{4.37}$$

(iv) *Minimal scale for $\mathcal{P}$. For each $q \in [1,\infty)$, there exists a constant $C := C(d,\mathfrak{p},q) < \infty$, a non-negative random variable $\mathcal{M}_q(\mathcal{P})$ and an exponent $r := r(d,\mathfrak{p},q) > 0$ such that*

$$\mathcal{M}_q(\mathcal{P}) \leqslant \mathcal{O}_r(C), \tag{4.38}$$

and for each radius R satisfying $R \geqslant \mathcal{M}_q(\mathcal{P})$,

$$R^{-d} \sum_{x \in \mathbb{Z}^d \cap B_R} \operatorname{size}(\square_{\mathcal{P}}(x))^q \leqslant C \quad \text{and} \quad \sup_{x \in \mathbb{Z}^d \cap B_R} \operatorname{size}(\square_{\mathcal{P}}(x)) \leqslant R^{\frac{1}{q}}. \tag{4.39}$$

Remark. To be precise, this proposition is a consequence of Propositions 2.1 and 2.4 of [19] and of Proposition 4.2.1 as is explained in [19, Section 2.2].

Additionally, the precise result stated in [19, Proposition 2.4] is that there exists a minimal scale $\mathcal{M}_q(\mathcal{P})$ above which $\sup_{x \in \mathbb{Z}^d \cap B_R} \operatorname{size}(\square_{\mathcal{P}}(x)) \leqslant R^{\frac{d}{d+q}}$, with an exponent $d/(d+q)$ instead of $1/q$. Nevertheless, it is straightforward to recover the statement of Proposition 4.2.2 from the one of [19, Proposition 2.4].

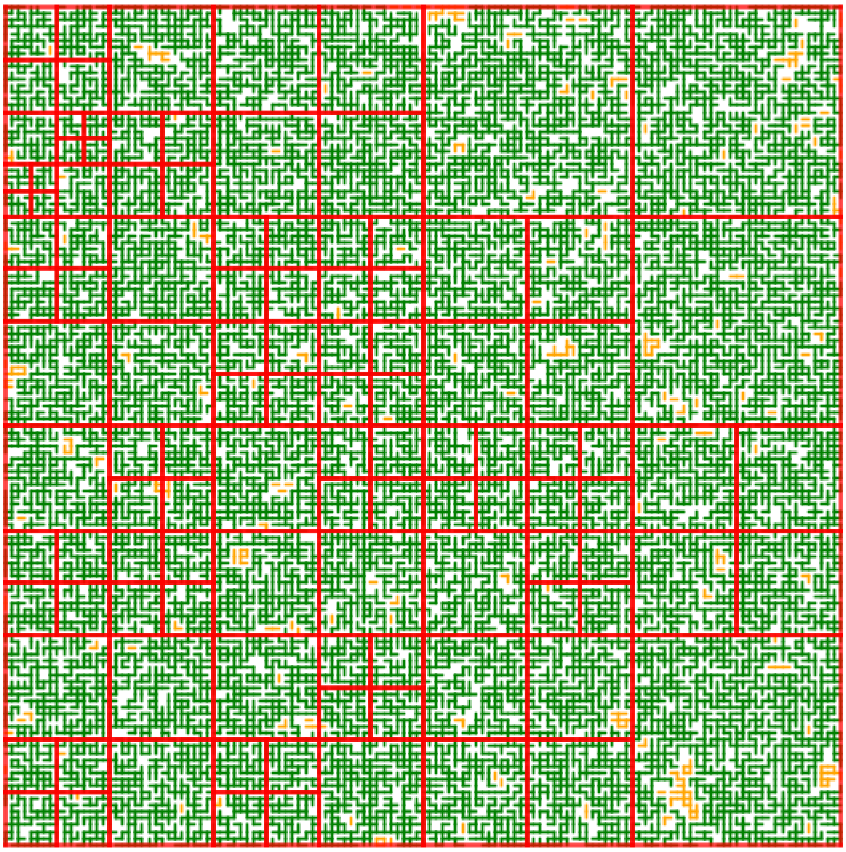

Figure 4.3: A realization of the partition $\mathcal{P}$, where the cluster in green is the maximal cluster and the cubes in red are elements of $\mathcal{P}$.

Figure 4.3 (drawn with dyadic cubes instead of triadic cubes to improve readability) illustrates what this partition looks like. It allows to extend functions defined on the infinite cluster to the whole space $\mathbb{R}^d$, as is explained below. We consider a function $u : \mathscr{C}_\infty \to \mathbb{R}$. For each point $x \in \mathbb{Z}^d$, we choose a point $z(x)$ in the cluster $\mathscr{C}_*(\square_{\mathcal{P}}(x))$ according to some deterministic procedure (for instance we choose the one which is the closest to the center of the cube and break ties by using the lexicographical order). We then define the coarsened function $[u]_{\mathcal{P}}$ on $\mathbb{Z}^d$ according to the formula, for each $x \in \mathbb{Z}^d$,

$$[u]_{\mathcal{P}}(x) := \begin{cases} u(x) & \text{if } x \in \mathscr{C}_\infty, \\ u(z(x)) & \text{otherwise,} \end{cases} \tag{4.40}$$

and extend it to the whole space $\mathbb{R}^d$ by setting it to be piecewise constant on the cubes $\left[x - \frac{1}{2}, x + \frac{1}{2}\right)^d$, for $x \in \mathbb{Z}^d$. When the function u is defined on the parabolic space $[0, \infty) \times \mathscr{C}_\infty$, we define its extension to the space $[0, \infty) \times \mathbb{Z}^d$, which we also denote by $[u]_{\mathcal{P}}$, according to the formula

$$[u]_{\mathcal{P}}(t, x) := \begin{cases} u(t, x) & \text{if } x \in \mathscr{C}_\infty, \\ u(t, z(x)) & \text{otherwise.} \end{cases} \tag{4.41}$$

For later purposes, we note that, given a function $u : \mathscr{C}_\infty \to \mathbb{R}$, the L^p-norm of the function $\nabla [u]_{\mathcal{P}}$ can be estimated in terms of L^p-norm of the function ∇u and the sizes of the cubes in the partition $\mathcal{P}$. Specifically, one has the formula, for any radius $r \geqslant \text{size}(\square_{\mathcal{P}}(0))$,

$$\|\nabla [u]_{\mathcal{P}}\|^p_{L^p(\mathbb{Z}^d \cap B_r)} \leqslant C \int_{B_r \cap \mathscr{C}_\infty} \text{size}(\square_{\mathcal{P}}(x))^{pd-1} |\nabla u(x)|^p \, dx. \tag{4.42}$$

The proof of this result can be found in [19, Lemma 3.3]. Additionally, one can estimate the L^p-norm of the function $[u]_{\mathcal{P}}$ in terms of the L^p-norm of the function u according to the formula, for any radius $r \geqslant \text{size}(\square_{\mathcal{P}}(0))$,

$$\|[u]_{\mathcal{P}}\|^p_{L^p(B_r)} \leqslant \int_{\hat{B}_r \cap \mathscr{C}_\infty} \text{size}(\square_{\mathcal{P}}(x))^d |u(x)|^p \, dx, \tag{4.43}$$

where $\hat{B}_r$ denotes the union of all the cubes in the partition $\mathcal{P}$ which intersect the ball B_r, i.e., $\hat{B}_r := \cup\{\square \in \mathcal{P} \, : \, \square \cap B_r \neq \varnothing\}$. This estimate is a consequence of the following argument: by definition of the coarsened function $[u]_{\mathcal{P}}$, one has the estimates, for any cube $\square$ of the partition $\mathcal{P}$,

$$\|[u]_{\mathcal{P}}\|^p_{L^p(\square)} \leqslant \text{size}(\square)^d \|[u]_{\mathcal{P}}\|^p_{L^\infty(\square)} \leqslant \text{size}(\square)^d \|u\|^p_{L^\infty(\mathscr{C}_\infty \cap \square)} \leqslant \text{size}(\square)^d \|u\|^p_{L^p(\mathscr{C}_\infty \cap \square)},$$

where we used the discrete $L^\infty - L^p$-estimate in the third inequality. Summing over all the cubes of the set $\hat{B}_r$ completes the proof of the estimate (4.43).

4.2.2 Functional inequalities on the infinite cluster

In this section, we state mostly without proofs, some functional inequalities which are valid on the infinite cluster $\mathscr{C}_\infty$. The partition of good cubes presented in Section 4.2.1 allows to prove these estimates and we refer to [19] for the details of the argument. Some of these inequalities were already proved by other renormalization technique: it is in particular the case of the Poincaré inequality which was established by Barlow in [39] (see also Mathieu, Remy [181] and Benjamini, Mossel [46]).

The fact that these bounds are stated on a random graph which has an irregular nature means that they are only valid on balls of size larger than some random minimal scales, denoted by $\mathcal{M}_{\text{Poinc}}$ and $\mathcal{M}_{\text{Meyers}}$ in the following statements, which depend on the environment $\mathbf{a}$ and are large when the environment is ill-behaved.

The first functional inequality we record is the Poincaré inequality, it can be found in [39, Theorem 2.18] for the L^2-version.

Proposition 4.2.3 (Poincaré inequality on $\mathscr{C}_\infty$)**.** *Fix a real number* $p \in \left[\frac{d}{d-1}, \infty\right)$*. There exist a constant* $C := C(d, \mathfrak{p}, p) < \infty$*, an exponent* $s := s(d, \mathfrak{p}, p) > 0$ *such that, for any* $y \in \mathbb{Z}^d$*, there exists a non-negative random variable* $\mathcal{M}_{L^p-\mathrm{Poinc}}(y)$ *which satisfies the stochastic integrability estimate*

$$\mathcal{M}_{L^p-\mathrm{Poinc}}(y) \leqslant \mathcal{O}_s(C), \tag{4.44}$$

such that for each radius $R \geqslant \mathcal{M}_{L^p-\mathrm{Poinc}}(y)$ *and each function* $u : \mathscr{C}_\infty \cap B_R(y) \to \mathbb{R}$*,*

$$\left\| u - (u)_{\mathscr{C}_\infty \cap B_R(y)} \right\|_{\underline{L}^p(\mathscr{C}_\infty \cap B_R(y))} \leqslant CR \left\| \nabla u \right\|_{\underline{L}^p(\mathscr{C}_\infty \cap B_R(y))}.$$

Moreover for each function $u : \mathscr{C}_\infty \cap B_R(y) \to \mathbb{R}$*, such that* $u = 0$ *on the boundary* $\partial\left(\mathbb{Z}^d \cap B_R(y)\right) \cap \mathscr{C}_\infty$*,*

$$\left\| u \right\|_{\underline{L}^p(\mathscr{C}_\infty \cap B_R(y))} \leqslant CR \left\| \nabla u \right\|_{\underline{L}^p(\mathscr{C}_\infty \cap B_R(y))}.$$

Remark. This inequality is frequently used in the case $p = 2$. To shorten the notation, we write $\mathcal{M}_{\mathrm{Poinc}}(y)$ to refer to the minimal scale $\mathcal{M}_{L^2-\mathrm{Poinc}}(y)$.

Proof. By translation invariance of the model, we can always assume $y = 0$. The proof relies on the Sobolev inequality as stated in [19, Proposition 3.4], together with the Hölder inequality by setting $\mathcal{M}_{L^p-\mathrm{Poinc}}(0) := \mathcal{M}_q(\mathcal{P})$, for a parameter q chosen large enough depending only on the dimension d and the exponent p. □

The second estimate we need to record is the parabolic Caccioppoli inequality. This estimate is valid on any subgraph of $\mathbb{Z}^d$ and is used in Section 4.3.2.

Proposition 4.2.4 (Parabolic Caccioppoli inequality on $\mathscr{C}_\infty$)**.** *There exists a finite positive constant* $C := C(d, \lambda)$ *such that, for each point* $y \in \mathbb{Z}^d$*, each radius* $R \geqslant 1$ *and each function* $u : I_R \times (\mathscr{C}_\infty \cap B_R(y)) \to \mathbb{R}$ *which is* **a***-caloric, i.e., which is a solution of the parabolic equation*

$$\partial_t u - \nabla \cdot \mathbf{a} \nabla u = 0 \ \text{in} \ I_R \times (\mathscr{C}_\infty \cap B_R(y)),$$

one has

$$\left\| \nabla u \right\|_{L^2(I_{R/2} \times (\mathscr{C}_\infty \cap B_{R/2}(y)))} \leqslant \frac{C}{R} \left\| u - (u)_{I_R \times (\mathscr{C}_\infty \cap B_R(y))} \right\|_{L^2(I_R \times (\mathscr{C}_\infty \cap B_R(y)))}.$$

Proof. The proof follows the standard arguments of the Caccioppoli inequality; the fact that the function u is defined on the infinite cluster does not affect the proof and we omit the details. □

The third estimate we record is an $L^\infty L^2$ gradient bound for **a**-caloric functions. The proof of this result can be found in [25, Lemma 8.2] in the uniformly elliptic setting, the extension to the percolation cluster makes no difference in the proof.

Lemma 4.2.2. *There exists a positive constant* $C := C(d, \lambda) < \infty$ *such that for any radius* $R \geqslant 1$*, any point* $y \in \mathbb{Z}^d$ *and any function* $u : I_R \times (\mathscr{C}_\infty \cap B_R(y)) \to \mathbb{R}$ *which satisfies*

$$\partial_t u - \nabla \cdot \mathbf{a} \nabla u = 0 \ \text{in} \ I_R \times (\mathscr{C}_\infty \cap B_R(y)),$$

one has the estimate

$$\sup_{t \in I_{R/2}} \left\| \nabla u(t, \cdot) \right\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_{R/2}(y))} \leqslant C \left\| \nabla u \right\|_{\underline{L}^2(I_R \times (\mathscr{C}_\infty \cap B_R(y)))}.$$

The last estimate we record in this section is the Meyers estimate for $\mathbf{a}$-caloric functions on the percolation cluster. This inequality is a non-concentration estimate and essentially states that the energy of solutions of a parabolic equation cannot concentrate in small volumes. It is used in Section 4.3.2.

Proposition 4.2.5 (Interior Meyers estimate on $\mathscr{C}_\infty$). *There exist a finite positive constant $C := C(d, \mathfrak{p}, \lambda)$, two exponents $s := s(d, \mathfrak{p}, \lambda) > 0$, $\delta_0 := \delta_0(d, \mathfrak{p}, \lambda) > 0$ such that, for each $y \in \mathbb{Z}^d$, there exists a non-negative random variable $\mathcal{M}_{\mathrm{Meyers}}(y)$ which satisfies the stochastic integrability estimate*

$$\mathcal{M}_{\mathrm{Meyers}}(y) \leqslant \mathcal{O}_s(C),$$

such that, for each radius $R \geqslant \mathcal{M}_{\mathrm{Meyers}}(y)$ and each function $u : I_R \times (\mathscr{C}_\infty \cap B_R(y)) \to \mathbb{R}$ solution of the equation

$$\partial_t u - \nabla \cdot \mathbf{a} \nabla u = 0 \text{ in } I_R \times (\mathscr{C}_\infty \cap B_R(y)),$$

one has

$$\|\nabla u\|_{\underline{L}^{2+\delta_0}(I_{R/2} \times (\mathscr{C}_\infty \cap B_{R/2}(y)))} \leqslant C \|\nabla u\|_{\underline{L}^2(I_R \times (\mathscr{C}_\infty \cap B_R(y)))}.$$

Proof. The classical proof of the interior Meyers estimate (cf. [112]) is based on an application of the Caccioppoli inequality, the Sobolev inequality and the Gehring's lemma (cf. [118]). The proof of this result on the percolation cluster for the elliptic problem is written in [19, Proposition 3.8]. For the parabolic problem considered here, the proof in the case of uniformly elliptic environments can be found in [18, Appendix B]. The argument can be adapted to the percolation cluster following the strategy developed in [19, Section 3]. Since the analysis does not contain any new idea regarding the method and the result can be obtained by essentially rewriting the proof, we skip the details.

□

4.2.3 Homogenization on percolation clusters

In this section, we collect some results of stochastic homogenization in supercritical percolation useful in the proof of Theorem 4.1.1. The proof of this theorem is based on a quantitative two-scale expansion, which relies on two important functions: the first-order corrector and its flux. They are introduced in Sections 4.2.3 and 4.2.3 respectively.

The first-order corrector

We let $\mathsf{A}_1(\mathscr{C}_\infty)$ be the random vector space of $\mathbf{a}$-harmonic functions on the infinite cluster $\mathscr{C}_\infty$ with at most linear growth. This latter condition is expressed in terms of average L^2-norm and we define

$$\mathsf{A}_1(\mathscr{C}_\infty) := \left\{ u : \mathscr{C}_\infty \to \mathbb{R} \ : \ -\nabla \cdot (\mathbf{a} \nabla u) = 0 \text{ in } \mathscr{C}_\infty \text{ and } \lim_{r \to \infty} r^{-2} \|u\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r)} = 0 \right\}.$$

It is known that this space is almost surely finite-dimensional and that its dimension is equal to $(d+1)$ (see [45]). Additionally, every function $u \in \mathsf{A}_1(\mathscr{C}_\infty)$ can be uniquely written as

$$u(x) = c + p \cdot x + \phi_p(x),$$

where $c \in \mathbb{R}$, $p \in \mathbb{R}^d$ and ϕ_p is a function called the corrector; it is defined up to a constant and satisfies the quantitative sublinearity property stated in the following proposition.

Proposition 4.2.6. *For any exponent $\alpha > 0$, there exist an exponent $s(d, \mathfrak{p}, \lambda, \alpha) > 0$ and a positive constant $C(d, \mathfrak{p}, \lambda, \alpha) < \infty$ such that, for any point $y \in \mathbb{Z}^d$, there exists a non-negative random variable $\mathcal{M}_{\mathrm{corr},\alpha}(y)$ satisfying the stochastic integrability estimate*

$$\mathcal{M}_{\mathrm{corr},\alpha}(y) \leqslant \mathcal{O}_s\left(C\right), \tag{4.45}$$

such that for every radius $r \geqslant \mathcal{M}_{\mathrm{corr},\alpha}(y)$, and every $p \in \mathbb{R}^d$,

$$\operatorname*{osc}_{x \in \mathscr{C}_\infty \cap B_r(y)} \phi_p := \left(\sup_{x \in \mathscr{C}_\infty \cap B_r(y)} \phi_p - \inf_{x \in \mathscr{C}_\infty \cap B_r(y)} \phi_p \right) \leqslant C|p| r^\alpha.$$

Proof. The proof of this result relies on the optimal scaling estimates for the corrector established in [83]. Indeed by [83, Theorem 1], one has the following result: there exists a constant $C := C(d, \lambda, \mathfrak{p}) < \infty$ and an exponent $s := s(d, \lambda, \mathfrak{p}) < \infty$ such that for each $x, y \in \mathbb{Z}^d$, and each $p \in \mathbb{R}^d$,

$$\left|\phi_p(x) - \phi_p(y)\right| \mathbf{1}_{\{\{x,y \in \mathscr{C}_\infty\}\}} \leqslant \begin{cases} \mathcal{O}_s(C|p|) & \text{if } d \geqslant 3, \\ \mathcal{O}_s(C|p| \log^{\frac{1}{2}} |x-y|) & \text{if } d = 2. \end{cases} \tag{4.46}$$

Proposition 4.2.6 is then a consequence of the previous estimate and an application of Lemma 4.1.1 with the sequence of random variables

$$X_n := 3^{-\alpha n} \sup_{x \in \mathbb{Z}^d \cap B_{3^n}(y)} \left|\phi_p(x) - \phi_p(y)\right| \mathbf{1}_{\{\{x,y \in \mathscr{C}_\infty\}\}}.$$

To be more precise, we use the estimate (4.24) to control the maximum of the random variables

$$\begin{aligned} X_n &= 3^{-\alpha n} \sup_{x \in \mathbb{Z}^d \cap B_{3^n}(y)} \left|\phi_p(x) - \phi_p(y)\right| \mathbf{1}_{\{\{x,y \in \mathscr{C}_\infty\}\}} \\ &\leqslant 3^{-\alpha n} \left(\sum_{x \in \mathbb{Z}^d \cap B_{3^n}(y)} \left(\left|\phi_p(x) - \phi_p(y)\right| \mathbf{1}_{\{\{x,y \in \mathscr{C}_\infty\}\}} \right)^{\frac{2d}{\alpha}} \right)^{\frac{\alpha}{2d}} \\ &\leqslant \begin{cases} \mathcal{O}_s\left(C|p| 3^{-\frac{\alpha n}{2}}\right) & \text{if } d \geqslant 3, \\ \mathcal{O}_s\left(C|p|\sqrt{n} 3^{-\frac{\alpha n}{2}}\right) & \text{if } d = 2. \end{cases} \end{aligned} \tag{4.47}$$

Then the sequence $\{X_n\}_{n \geqslant 1}$ satisfies the assumption of Lemma 4.1.1. □

The fact that the corrector is only defined up to a constant causes some technical difficulties in the proofs, in particular the two-scale expansion stated in (4.107) and used in the proof of Theorem 4.1.1 is ill-defined in this setting. To solve this issue, we choose the following (arbitrary) normalization for the corrector: given a point $y \in \mathbb{Z}^d$ and an environment $\mathbf{a}$ in the set of probability 1 on which the corrector is well-defined, we let $x \in \mathscr{C}_\infty$ which is the closest to the point y (and break ties by using the lexicographical order) and normalize the corrector by setting $\phi_p(x) = 0$. The choice of the point y will always be explicitly indicated to avoid confusions. We note that with this normalization, the corrector is not stationary.

The centered flux

A second important notion in the implementation of the two-scale expansion is the centered flux; it is defined in the following paragraph.

For a fixed vector $p = (p_1, \ldots, p_d) \in \mathbb{R}^d$, we consider the mapping $\mathbf{a}(p + \mathcal{D}\phi_p) : \mathscr{C}_\infty \to \mathbb{R}^d$ defined by the formula, for each $x \in \mathbb{Z}^d$,

$$\mathbf{a}(p + \mathcal{D}\phi_p)(x) := (\mathbf{a}(\{x, x+e_1\}) (p_1 + \mathcal{D}_{e_1}\phi_p(x)), \ldots, \mathbf{a}(\{x, x+e_d\}) (p_d + \mathcal{D}_{e_d}\phi_p(x))) .$$

This function oscillates quickly but it is close to the deterministic slope $\frac{1}{2}\bar{\sigma}^2 p$ in the $\underline{H}^{-1}$-norm on the infinite cluster, where $\bar{\sigma}$ is the diffusivity of the random walk introduced in (4.4). This motivates the following definition: for a fixed vector $p \in \mathbb{R}^d$, we define the centered flux $\widetilde{\mathbf{g}}_p : \mathscr{C}_\infty \to \mathbb{R}^d$ according to the formula

$$\widetilde{\mathbf{g}}_p := \mathbf{a}(\mathcal{D}\phi_p + p) - \frac{1}{2}\bar{\sigma}^2 p.$$

The following proposition estimates the $\underline{H}^{-1}$-norm of the centered flux. It it proved in Appendix 4.B, Proposition 4.B.1.

Proposition 4.2.7. *For any exponent $\alpha > 0$, there exist a positive constant $C(d, \mathfrak{p}, \lambda, \alpha) < \infty$ and two exponents $s(d, \mathfrak{p}, \lambda, \alpha) > 0$ and $\alpha(d, \mathfrak{p}, \lambda) > 0$ such that, for any $y \in \mathbb{Z}^d$, there exists a non-negative random variable $\mathcal{M}_{\mathrm{flux},\alpha}(y)$ satisfying the stochastic integrability estimate*

$$\mathcal{M}_{\mathrm{flux},\alpha}(y) \leqslant \mathcal{O}_s(C), \tag{4.48}$$

such that for each radius $r \geqslant \mathcal{M}_{\mathrm{flux},\alpha}(y)$,

$$\|\widetilde{\mathbf{g}}_p\|_{\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r(y))} \leqslant C|p|r^\alpha. \tag{4.49}$$

Remark. We emphasize that, in this article, the previous proposition is not a property of the diffusivity $\bar{\sigma}^2$ but its definition: building on former result from [19, 83, 134], we prove that there exists a coefficient such that the estimate (4.49) is satisfied and name this coefficient $\bar{\sigma}^2$. Thanks to the estimate (4.49), we are then able to prove Theorem 4.1.1 and the invariance principle (4.4) with the same coefficient $\bar{\sigma}^2$. We refer to (4.182) and Remark 4.B for a more detailed discussion.

4.2.4 Random walks on graphs

In this section, we record the Carne-Varopoulos bound pertaining to the transition kernel of the continuous-time random walk which holds on any infinite connected subgraph of $\mathbb{Z}^d$. This estimate is not as strong as the ones we are trying to establish (for instance the ones of Theorem 4.1.1, or of Theorem 4.3.1 proved in [39]) but can be applied in greater generality: it applies to any realization of the infinite cluster, i.e., to any environment $\mathbf{a}$ in the set of probability 1 where there exists a unique infinite cluster, without any consideration about its geometry. From a mathematical perspective, this means that there is no minimal scale in the statement of Proposition 4.2.8.

Proposition 4.2.8 (Carne-Varopoulos bound, Corollaries 11 and 12 of [87]). *Let $\mathcal{G}$ be an infinite, connected subgraph of $\mathbb{Z}^d$ and $\mathbf{a}$ be a function from the bonds of $\mathcal{G}$ into $[\lambda, 1]$. For $y \in \mathcal{G}$, we let $p(\cdot, \cdot, y)$ be the heat kernel associated to the parabolic equation*

$$\begin{cases} \partial_t p(\cdot, \cdot, y) - \nabla \cdot (\mathbf{a} \nabla p(\cdot, \cdot, y)) = 0 & \text{in } (0, \infty) \times \mathcal{G}, \\ p(0, \cdot, y) = \delta_y & \text{in } \mathcal{G}. \end{cases}$$

Then there exists a positive constant $C := C(d, \lambda) < \infty$ *such that for each point* $x \in \mathcal{G}$*,*

$$p(t,x,y) \leqslant \begin{cases} C \exp\left(-\frac{|x-y|^2}{Ct}\right) & \text{if } |x-y| \leqslant t, \\ C \exp\left(-\frac{|x-y|}{C}\left(1+\ln\frac{|x-y|}{t}\right)\right) & \text{if } |x-y| \geqslant t. \end{cases} \tag{4.50}$$

Remark. For later use, we note that, when $|x-y| \geqslant t$,

$$\exp\left(-\frac{|x-y|}{C}\left(1+\ln\frac{|x-y|}{t}\right)\right) \leqslant \exp\left(-\frac{t}{C}\right).$$

A consequence of this inequality is that, by increasing the value of the constant C, one can add a factor $t^{-d/2}$ in the second line of the right side of (4.50): for every constant $0 < C < \infty$ there exists a finite constant $C' > C$ such that, when $|x-y| \geqslant t$,

$$C \exp\left(-\frac{|x-y|}{C}\left(1+\ln\frac{|x-y|}{t}\right)\right) \leqslant C' t^{-d/2} \exp\left(-\frac{|x-y|}{C'}\left(1+\ln\frac{|x-y|}{t}\right)\right).$$

4.3 Decay and Lipschitz regularity of the heat kernel

In this section, we collect and establish some estimates about the decay of the parabolic Green's function. In Section 4.3.1, we record a result of Barlow in [39], which establishes Gaussian upper bounds on the parabolic Green's function on the infinite cluster. This result is a percolation version of the Nash-Aronson estimate [32], originally proved for uniformly elliptic divergence form diffusions. Building upon the result of Barlow, we then establish estimates on the gradient of the parabolic Green's function on the percolation cluster, stated in Theorem 4.1.3, thanks to a large-scale $C^{0,1}$-regularity estimate. The argument makes use of techniques from stochastic homogenization and follows a classical route which can be decomposed into three steps: we first establish a quantitative homogenization theorem for the parabolic Dirichlet problem (see Section 4.3.2), once this is achieved we prove a large-scale $C^{0,1}$-regularity estimate for **a**-caloric functions (see Section 4.3.2). In Section 4.3.2, we use this regularity estimate together with the heat kernel bound of Barlow to obtain the decay of the gradient of the heat kernel stated in Theorem 4.1.3.

4.3.1 Decay of the heat kernel

In this section, we record the result of Barlow [39], who established Gaussian bounds on the transition kernel. We first introduce the following function.

Definition 4.3.1. Given a point $x \in \mathbb{R}^d$, a time $t \in (0, \infty)$ and a constant $0 < C < \infty$, we define the function Φ_C according to the formula

$$\Phi_C(t,x) := \begin{cases} Ct^{-d/2} \exp\left(-\frac{|x|^2}{Ct}\right) & \text{if } |x| \leqslant t, \\ Ct^{-d/2} \exp\left(-\frac{|x|}{C}\left(1+\ln\frac{|x|}{t}\right)\right) & \text{if } |x| \geqslant t. \end{cases} \tag{4.51}$$

We note that this function is radial and increasing in the variable C. This function corresponds to a discrete heat kernel. For further use, we note that it satisfies the following semigroup property, for each $t_1, t_2 \in (0, \infty)$ and each $x, y \in \mathbb{Z}^d$

$$\int_{\mathbb{Z}^d} \Phi_C(t_1, x-z)\, \Phi_C(t_2, z-y)\, dz \leqslant \Phi_{C'}(t_1+t_2, x-y), \tag{4.52}$$

for some larger constant $C' > C$. This property is proved by an explicit computation or by using the semigroup property of the law of the random walk on $\mathbb{Z}^d$ (see Remark 4.3.1). We define the function Ψ_C according to the formula

$$\Psi_C(t,r) := \begin{cases} (0,\infty) \times [0,\infty) & \to \mathbb{R}, \\ (t,r) & \mapsto -\ln\left(\Phi_C(t,x)\right), \text{ where } x \in \mathbb{R}^d \text{ satisfies } |x| = r. \end{cases} \tag{4.53}$$

In particular, one has the identity,

$$\forall x \in \mathbb{R}^d, \ \Phi_C(t,x) = \exp\left(-\Psi_C(t,|x|)\right).$$

The function Ψ_C satisfies the following properties:

(i) It is decreasing in the variable C, increasing in the variable r, and continuous with respect to both variables;

(ii) It is convex with respect to the variable r.

We now record the result of Barlow.

Theorem 4.3.1 (Gaussian upper bound, Theorem 1 and Lemma 1.1 of [39])**.** *There exist an exponent $s := s(d,\mathfrak{p},\lambda) > 0$, a positive constant $C := C(d,\mathfrak{p},\lambda) < \infty$ such that for each $y \in \mathbb{R}^d$, there exists a random time $\mathcal{T}_{\mathrm{NA}}(y)$ satisfying the stochastic integrability estimate*

$$\mathcal{T}_{\mathrm{NA}}(y) \leqslant \mathcal{O}_s(C), \tag{4.54}$$

such that, on the event $\{y \in \mathscr{C}_\infty\}$, for every time $t \in (0,\infty)$ satisfying $t \geqslant \mathcal{T}_{\mathrm{NA}}(y)$, and every point $x \in \mathscr{C}_\infty$,

$$p(t,x,y) \leqslant \Phi_C(t,x-y). \tag{4.55}$$

Remark. The stochastic integrability estimate (4.54) is not stated in Theorem 1.1 of [39] but is mentioned in its remark equation (0.5) following the theorem.

Remark. The estimate in the regime $t \leqslant |x-y|$ does not require the assumption that t is larger than the minimal scale $\mathcal{T}_{\mathrm{NA}}(y)$ and is in fact a deterministic result: it is a consequence of Proposition 4.2.8 (proved in [87]) and Remark 4.2.4.

Remark. The function Φ_C can be used to obtain upper and lower bounds on the law of the random walk on the lattice $\mathbb{Z}^d$: there exist constants C_1, C_2 depending only on the dimension d such that

$$\Phi_{C_1}(t,x-y) \leqslant p^{\mathbb{Z}^d}(t,x,y) \leqslant \Phi_{C_2}(t,x-y), \tag{4.56}$$

where we used the notation $p^{\mathbb{Z}^d}(t,x,y) := \mathbb{P}_y[X_t = x]$, and where $(X_t)_{t\geqslant 0}$ denotes the VSRW on $\mathbb{Z}^d$ starting from the point y. We refer to the work [92] of Delmotte and the work [87] for this result. The estimates (4.56) can then be used to prove the property (4.52). Indeed, since the random walk $(X_t)_{t\geqslant 0}$ is a Markov process, its transition function $p^{\mathbb{Z}^d}$ has the semigroup property, and we can write

$$\begin{aligned} \int_{\mathbb{Z}^d} \Phi_{C_1}(t_1,x-z)\,\Phi_{C_1}(t_2,z-y)\,dz &\leqslant \int_{\mathbb{Z}^d} p^{\mathbb{Z}^d}(t_1,x,z)\,p^{\mathbb{Z}^d}(t_2,z,y)\,dz \\ &= p^{\mathbb{Z}^d}(t_1+t_2,x,y) \\ &\leqslant \Phi_{C_2}(t_1+t_2,x-y). \end{aligned}$$

This argument gives the estimate (4.52) in the case $C = C_1$, but can be easily extended to any constant $C > 0$.

We complete this section by mentioning that the result of Barlow is proved for the heat kernel associated to the constant speed random walk and on the percolation cluster only, i.e., when the conductances are only allowed to take the values 0 or 1. The adaptation to the variable speed random walk with uniformly elliptic conductances only requires a typographical change of the proof: all the computations performed in [39] to obtain the upper bound (4.55) can be adapted to our setting and so is the case of the existing results in the literature which are used in the proof.

4.3.2 Decay of the gradient of the Green's function

The main objective of this section is to prove Theorem 4.1.3. The proof of this result makes use of techniques from stochastic homogenization and can be split into three distinct steps, which correspond to the three following subsections. The first idea is to prove that the parabolic Green's function is close, on large scales, to a caloric function. This is carried out in Section 4.3.2 and the proof is based on a two-scale expansion. The analysis relies on the sublinearity of the corrector and the estimate on the $\underline{H}^{-1}$-norm of the centered flux stated in Section 4.2.3. This result is only necessary to establish a large-scale regularity theory for which sharp homogenization errors are not needed; we thus do not try to prove an optimal error estimate in the homogenization of the parabolic Dirichlet problem and only prove the result with an algebraic and suboptimal rate of convergence. Then, in Section 4.3.2, we use the homogenization estimate proved in Theorem 4.3.2 to establish a large-scale regularity theory in the spirit of [25, Chapter 3] or [19, Section 7]. Finally, in Section 4.3.2, we combine Proposition 4.3.1 and the heat kernel bound proved by Barlow and stated in Theorem 4.3.1 to deduce Theorem 4.1.3.

Homogenization of the parabolic Dirichlet problem

In this section, we prove a quantitative homogenization theorem for the parabolic Cauchy-Dirichlet problem on the infinite cluster. In the following statement, we let η be a smooth, non-negative function supported in the ball $B_{\frac{1}{2}}(0)$, and satisfying the identity $\int \eta = 1$. It is used as a smoothing operator in the convolution (4.60). We also define the set $\mathrm{Cv}(\mathbb{Z}^d \cap B_r(y))$ to be the convex hull of the set $\mathbb{Z}^d \cap B_r(y)$, i.e.,

$$\mathrm{Cv}(\mathbb{Z}^d \cap B_r(y)) := \left\{ z \in \mathbb{R}^d : z = \sum_i \alpha_i x_i, \quad x_i \in \mathbb{Z}^d \cap B_r(y), \quad 0 \leqslant \alpha_i \leqslant 1 \text{ and } \sum_i \alpha_i = 1 \right\}. \tag{4.57}$$

It is used to define the domain of the homogenized equation so that the boundary condition coincides.

Theorem 4.3.2. *Fix an exponent $\delta > 0$, then there exist a positive constant $C(d, \lambda, \mathfrak{p}, \delta) < \infty$, two exponents $s(d, \lambda, \mathfrak{p}, \delta) > 0$, $\alpha(d, \lambda, \mathfrak{p}, \delta) > 0$ such that for any point $y \in \mathbb{Z}^d$, there exists a non-negative random variable $\mathcal{M}_{\mathrm{hom},\delta}(y)$ satisfying*

$$\mathcal{M}_{\mathrm{hom},\delta}(y) \leqslant \mathcal{O}_s(C)$$

such that, for every $r > \mathcal{M}_{\mathrm{hom}}(y)$, and every boundary condition $f \in W^{1,2+\delta}_{\mathrm{par}}(I_r \times (\mathscr{C}_\infty \cap B_r(y)))$, the following statement is valid. Let u be the weak solution of the parabolic equation

$$\begin{cases} (\partial_t - \nabla \cdot \mathbf{a} \nabla) u = 0 & \text{in } I_r \times (\mathscr{C}_\infty \cap B_r(y)), \\ u = f & \text{on } \partial_{\sqcup} \left(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r(y)) \right) \cap (I_r \times (\mathscr{C}_\infty \cap B_r(y))), \end{cases} \tag{4.58}$$

and $\bar{u}$ be the weak solution of the homogenized, continuous in space, parabolic equation

$$\begin{cases} (\partial_t - \frac{1}{2}\bar{\sigma}^2\Delta)\bar{u} = 0 & \text{in } I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r(y)), \\ \bar{u} = \widetilde{f} & \text{on } \partial_{\sqcup}\left(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r(y))\right), \end{cases} \tag{4.59}$$

where the boundary condition $\widetilde{f}$ is the extension of f to the continuous parabolic cylinder defined by the formula

$$\widetilde{f} := [f]_{\mathcal{P}} \star \eta, \tag{4.60}$$

and the extension $[f]_{\mathcal{P}}$ is defined in the paragraph following Proposition 4.2.2. Then, the following estimate holds

$$\frac{1}{r}\left\|u - \bar{u}\right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r(y)))} \leqslant C r^{-\alpha}\left\|\nabla f\right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r(y)))}. \tag{4.61}$$

Remark. The equation (4.58) is discrete in space and continuous in time, while equation (4.59) is both continuous in space and time. The solution u and $\bar{u}$ coincide on the parabolic boundary $\partial_{\sqcup}\left(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r(y))\right) \cap (I_r \times (\mathscr{C}_\infty \cap B_r(y)))$ for the equation on the clusters. All the norms in the inequality (4.61) are discrete in space and continuous in time.

Remark. The reason we define the homogenized limit to be continuous is the following: we need to use a number of results (e.g., regularity theory for the homogenized equation, the Meyers estimate) which are usually stated in the continuous setting. Moreover, one has explicit formulas for the elliptic and parabolic Green's functions and the continuous object is better behaved regarding scaling properties. On a higher level, the correct limiting object should be the continuous function as, over large-scales, the discrete lattice approximates the continuum.

Proof of Theorem 4.3.2. By translation invariance of the model, we assume without loss of generality that $y = 0$ and do some preparation before the proof. We first define the minimal scale $\mathcal{M}_{\mathrm{hom},\delta}(0)$ to be equal to

$$\mathcal{M}_{\mathrm{hom},\delta}(0) := \max\left(\mathcal{M}_{\mathrm{Poinc}}(0), \mathcal{M}_q(\mathcal{P}), \mathcal{M}_{\mathrm{corr},\frac{1}{2}}(0), \mathcal{M}_{\mathrm{flux},\frac{1}{2}}(0)\right),$$

where the parameter q is assumed to be larger than $4d$ and will be fixed at the end of the proof. Using the stochastic integrability estimates (4.38), (4.44), (4.45), (4.48) on the four minimal scales together with the property (4.24) of the $\mathcal{O}_s$ notation, one has

$$\mathcal{M}_{\mathrm{hom},\delta}(0) \leqslant \mathcal{O}_s(C).$$

We record that under the assumption $r \geqslant \mathcal{M}_{\mathrm{hom},\delta}(0) \geqslant \mathcal{M}_{2d}(\mathcal{P})$, one has

$$c r^d \leqslant |\mathscr{C}_\infty \cap B_r| \leqslant C r^d, \tag{4.62}$$

which allows to compare the number of points of the infinite cluster in the ball B_r with the volume of the ball B_r. This estimate can be deduced by an application of the estimate (4.39) with the Cauchy-Schwarz inequality:

$$|B_r|^2 \leqslant \left(\sum_{\square\in\mathcal{P},\square\cap B_r\neq\varnothing}(\mathrm{size}(\square))^d\right)^2 \leqslant \left(\sum_{\square\in\mathcal{P},\square\cap B_r\neq\varnothing} 1\right)\left(\sum_{\square\in\mathcal{P},\square\cap B_r\neq\varnothing}(\mathrm{size}(\square))^{2d}\right) \leqslant C|\mathscr{C}_\infty \cap B_r| r^d.$$

We record the following interior regularity estimate for the homogenized function $\bar{u}$, which is standard for solutions of the heat equation (see [105, Theorem 9, Section 2.3]): for every pair

$(t,x) \in I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r)$, and every radii $r_1, r_2 > 0$ such that $I_{r_1}(t) \times B_{r_2}(x) \subseteq I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r)$, one has the inequality

$$\forall k,l \in \mathbb{N}, \qquad |\partial_t^l \nabla^{k+1} \bar{u}|(t,x) \leqslant C_{k+2l} (r_1)^{-2l} (r_2)^{-k} \left\| \nabla \bar{u} \right\|_{\underline{L}^2(I_{r_1}(t) \times B_{r_2}(x))}. \tag{4.63}$$

We remark that in [105, Theorem 9, Section 2.3] the inequality is stated in the case when $r_1 = r_2$; The estimate (4.63) can be recovered by a careful investigation of the proof.

We introduce a cut-off function Υ in the parabolic cylinder $I_r \times B_r$ constant equal to 1 in the interior of the cylinder and decreasing linearly to 0 in a mesoscopic boundary layer of size $r' \ll r$,

$$\begin{cases} \Upsilon(t,x) \equiv 1 & (t,x) \in I_r \times B_r, \ \mathrm{dist}(x, \partial B_r) \geqslant 2r' \text{ and } \mathrm{dist}(t, \partial I_r) \geqslant 2(r')^2, \\ 0 \leqslant \Upsilon(t,x) \leqslant 1 & (t,x) \in I_r \times B_r, \\ \Upsilon(t,x) \equiv 0 & (t,x) \in I_r \times B_r, \ \mathrm{dist}(x, \partial B_r) \leqslant r' \text{ or } \mathrm{dist}(t, \partial I_r) \leqslant (r')^2. \end{cases} \tag{4.64}$$

The precise value of the parameter r' is given by the formula $r' = r^{1-\beta}$ for some small exponent β whose value is decided at the end of the proof. We additionally assume that the function Υ is smooth and satisfies the estimate

$$\forall k,l \in \mathbb{N}, \qquad |\partial_t^l \nabla^k \Upsilon| \leqslant C_{k+2l} (r')^{-(k+2l)}. \tag{4.65}$$

With these quantities, we can prove the following lemma.

Lemma 4.3.1. *We have the estimate*

$$\forall k,l \in \mathbb{N}, \qquad \left\| \partial_t^l \nabla^k (\Upsilon \nabla \bar{u}) \right\|_{L^\infty(I_r \times B_r)} \leqslant C_{k+2l} (r')^{-(k+2l)} \left(\frac{r}{r'} \right)^{\frac{2+d}{2}} \left\| \nabla f \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))}. \tag{4.66}$$

Proof. First, by using the inequality (4.63) and the fact that the map Υ is supported outside a boundary layer of size r' in the parabolic cylinder $I_r \times B_r$, we obtain the estimate

$$\begin{aligned} \forall (t,x) \in \mathrm{supp}(\Upsilon), \forall k,l \in \mathbb{N}, \quad |\partial_t^l \nabla^k (\Upsilon \nabla \bar{u})|(t,x) &\leqslant C_{k+2l} (r')^{-(k+2l)} \left\| \nabla \bar{u} \right\|_{\underline{L}^2(I_{r'}(t) \times B_{r'}(x))} \\ &\leqslant C_{k+2l} (r')^{-(k+2l)} \left(\frac{r}{r'} \right)^{\frac{2+d}{2}} \left\| \nabla \bar{u} \right\|_{\underline{L}^2(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))}. \end{aligned} \tag{4.67}$$

The inequality (4.67) implies the L^∞-estimate

$$\forall k,l \in \mathbb{N}, \quad \left\| \partial_t^l \nabla^k (\Upsilon \nabla \bar{u}) \right\|_{L^\infty(I_r \times B_r)} \leqslant C_{k+2l} (r')^{-(k+2l)} \left(\frac{r}{r'} \right)^{\frac{2+d}{2}} \left\| \nabla \bar{u} \right\|_{\underline{L}^2(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))}. \tag{4.68}$$

We then state the global Meyers estimate for the map $\bar{u}$: there exists an exponent $\delta_0 (d, \lambda, \mathfrak{p}) > 0$ such that for every $\delta' \in [0, \delta_0]$,

$$\left\| \nabla \bar{u} \right\|_{L^{2+\delta'}(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))} \leqslant C \left\| \nabla \widetilde{f} \right\|_{L^{2+\delta'}(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))}. \tag{4.69}$$

A proof of this result can be found in [113, Proposition 5.1], where the statement is given for cubes instead of parabolic cylinders (the adaptation to the setting considered here does not affect the proof). Moreover, one can estimate the L^p-norm of the (continuous) gradient of the function $\widetilde{f}$ in terms of the L^p-norm of the (discrete) gradient of the maps f and the sizes of the cubes of the partition. The formula is a consequence of [19, Lemma 3.3] and recalled in (4.42): for any $p \geqslant 1$, and any radius $r \geqslant \operatorname{size}(\square_{\mathcal{P}}(0))$,

$$\left\| \nabla \widetilde{f} \right\|_{L^p(\mathrm{Cv}(\mathbb{Z}^d \cap B_r))} \leqslant \int_{B_r \cap \mathscr{C}_\infty} \operatorname{size}(\square_{\mathcal{P}}(x))^{pd-1} \left| \nabla f(x) \right|^p \, dx.$$

Applying the Hölder inequality to this estimate with $r \geqslant \mathcal{M}_{\mathrm{hom},\delta}(0)$, using the assumption the minimal scale $\mathcal{M}_{\mathrm{hom},\delta}(0)$ is larger than the minimal scale $\mathcal{M}_q(\mathcal{P})$ so that Proposition 4.2.2 is valid, and choosing the parameter q to be large enough (larger than the value $\frac{(4+2\delta)\left((2+\frac{1}{2}\delta)d-1\right)}{\delta}$), one obtains the following inequality: for any $p \in \left[2, 2+\frac{1}{2}\delta\right]$,

$$\left\| \nabla \widetilde{f} \right\|_{\underline{L}^p(\mathrm{Cv}(\mathbb{Z}^d \cap B_r))} \leqslant C \left\| \nabla f \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))}. \tag{4.70}$$

Together with (4.69), this shows the inequality, for any exponent $\delta' \in \left[0, \min\left(\delta_0, \frac{1}{2}\delta\right)\right]$,

$$\left\| \nabla \bar{u} \right\|_{\underline{L}^{2+\delta'}(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))} \leqslant C \left\| \nabla f \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))}. \tag{4.71}$$

Putting the inequality (4.71) back into the estimate (4.68) concludes the proof of (4.66). □

The key ingredient in the proof of Theorem 4.3.2 is to use a modified two-scale expansion on the percolation cluster, defined for each $(t,x) \in I_r \times (\mathscr{C}_\infty \cap B_r)$ by the formula

$$w(t,x) := \bar{u}(t,x) + \Upsilon(t,x) \sum_{k=1}^d \mathcal{D}_{e_k} \bar{u}(t,x) \phi_{e_k}(x), \tag{4.72}$$

as an intermediate quantity: we prove that the function w is close to both functions u and $\bar{u}$. Here and in the rest of this section, the map ϕ_{e_k} is the first order corrector normalized according the procedure described in Section 4.2.3 around the point $y = 0$. The proof of Theorem 4.3.2 can be decomposed into five steps.

Step 1: Control over $\frac{1}{r} \left\| w - \bar{u} \right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r))}$. We use the estimate (4.66) to compute

$$\begin{aligned} \frac{1}{r} \left\| w - \bar{u} \right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r)))} &= \frac{1}{r} \left\| \sum_{k=1}^d \Upsilon \left(\mathcal{D}_{e_k} \bar{u} \right) \phi_{e_k} \right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r)))} \\ &\leqslant \frac{1}{r} \left\| \Upsilon \nabla \bar{u} \right\|_{L^\infty(I_r \times B_r)} \sum_{k=1}^d \left\| \phi_{e_k} \right\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r)} \\ &\leqslant \frac{C}{r} \left(\frac{r}{r'} \right)^{\frac{2+d}{2}} \left\| \nabla f \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))} \sum_{k=1}^d \left\| \phi_{e_k} \right\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r)}. \end{aligned}$$

Using the assumption $r \geqslant \mathcal{M}_{\mathrm{hom},\delta}(0) \geqslant \mathcal{M}_{\mathrm{corr},\frac{1}{2}}(0)$, we deduce

$$\frac{1}{r} \left\| w - \bar{u} \right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r)))} \leqslant C r^{-\frac{1}{2}} \left(\frac{r}{r'} \right)^{\frac{2+d}{2}} \left\| \nabla f \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))}.$$

The proof of Step 1 is complete.

Step 2: Control of $\frac{1}{r} \left\| w - u \right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r))}$ *by the norm* $\left\| (\partial_t - \nabla \cdot \mathbf{a} \nabla) w \right\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathscr{C}_\infty \cap Br))}$. We first note that the functions w and u are equal on the boundary of the parabolic cylinder $I_r \times (\mathscr{C}_\infty \cap B_r)$, and use the assumption $r \geqslant \mathcal{M}_{\mathrm{hom},\delta}(0) \geqslant \mathcal{M}_{\mathrm{Poinc}}(0)$ to apply the Poincaré inequality for each fixed time t and then integrate over time. This proves

$$\frac{1}{r} \left\| w - u \right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r))} \leqslant \left\| \nabla (w - u) \right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r))}.$$

Then, we use an integration by part and the uniform ellipticity of the environment on the infinite cluster

$$\begin{aligned}\|\nabla(w-u)\|^2_{\underline{L}^2(I_r\times(\mathscr{C}_\infty\cap B_r))} &\leqslant \frac{1}{\lambda|I_r\times(\mathscr{C}_\infty\cap B_r)|}\int_{I_r}\int_{\mathscr{C}_\infty\cap B_r}\nabla(w-u)\cdot\mathbf{a}\nabla(w-u)\\ &= \frac{1}{\lambda|I_r\times(\mathscr{C}_\infty\cap B_r)|}\int_{I_r}\int_{\mathscr{C}_\infty\cap B_r}\left(-\nabla\cdot\mathbf{a}\nabla(w-u)\right)(w-u).\end{aligned}$$

The fact that the functions w,u have the same initial condition over $\mathscr{C}_\infty\cap B_r$ implies that the following integral is non-negative

$$\begin{aligned}\int_{I_r}\int_{\mathscr{C}_\infty\cap B_r}\left(\partial_t(w-u)\right)(w-u) &= \frac{1}{2}\left(\|(w-u)(0,\cdot)\|^2_{L^2(\mathscr{C}_\infty\cap B_r)}-\left\|(w-u)(-r^2,\cdot)\right\|^2_{L^2(\mathscr{C}_\infty\cap B_r)}\right)\\ &= \frac{1}{2}\|(w-u)(0,\cdot)\|^2_{L^2(\mathscr{C}_\infty\cap B_r)}\geqslant 0.\end{aligned}$$

We combine this formula and equation (4.58) to obtain

$$\begin{aligned}\|\nabla(w-u)\|^2_{\underline{L}^2(I_r\times(\mathscr{C}_\infty\cap B_r))} &\leqslant \frac{1}{\lambda|I_r\times(\mathscr{C}_\infty\cap B_r)|}\int_{I_r}\int_{\mathscr{C}_\infty\cap B_r}\left((\partial_t-\nabla\cdot\mathbf{a}\nabla)(w-u)\right)(w-u)\\ &\leqslant \frac{1}{\lambda}\|w-u\|_{\underline{L}^2(I_r;\underline{H}^1(\mathscr{C}_\infty\cap B_r))}\|(\partial_t-\nabla\cdot\mathbf{a}\nabla)w\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r))}.\end{aligned}$$

This shows that

$$\frac{1}{r}\|w-u\|_{\underline{L}^2(\mathscr{C}_\infty\cap B_r)}\leqslant C\|(\partial_t-\nabla\cdot\mathbf{a}\nabla)w\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r))}.$$

Step 3: Control over $\|(\partial_t-\nabla\cdot\mathbf{a}\nabla)w\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r))}$. In this step, we adopt the finite difference notation and recall the identity $(\partial_t-\nabla\cdot\mathbf{a}\nabla)w=(\partial_t+\mathcal{D}^*\cdot\mathbf{a}\mathcal{D})w$. To estimate the $\underline{H}^{-1}$-norm of $(\partial_t-\nabla\cdot\mathbf{a}\nabla)w$, the idea is to derive an explicit formula for this quantity by using the definition of w given in (4.72) and to make a centered flux $\tilde{\mathbf{g}}_{e_k}=\mathbf{a}(\mathcal{D}\phi_{e_k}+e_k)-\frac{1}{2}\bar{\sigma}^2e_k$ appear. We first calculate $\partial_t w$ and $\mathcal{D}w$ and obtain the formulas

$$\begin{cases}\partial_t w &= \partial_t\bar{u}+\sum_{k=1}^d\partial_t(\Upsilon\mathcal{D}_{e_k}\bar{u})\phi_{e_k},\\ \mathcal{D}w &= (1-\Upsilon)\mathcal{D}\bar{u}+\sum_{k=1}^d(\Upsilon\mathcal{D}_{e_k}\bar{u})(e_k+\mathcal{D}\phi_{e_k})+\sum_{k=1}^d\mathcal{D}(\Upsilon\mathcal{D}_{e_k}\bar{u})\phi_{e_k}.\end{cases}$$

We combine the two equations to calculate $(\partial_t+\mathcal{D}^*\cdot\mathbf{a}\mathcal{D})w$,

$$\begin{aligned}(\partial_t+\mathcal{D}^*\cdot\mathbf{a}\mathcal{D})w &= \partial_t\bar{u}+\sum_{k=1}^d\partial_t(\Upsilon\mathcal{D}_{e_k}\bar{u})\phi_{e_k}+\mathcal{D}^*\cdot((1-\Upsilon)\mathbf{a}\mathcal{D}\bar{u})\\ &\quad+\sum_{k=1}^d\mathcal{D}^*\cdot((\Upsilon\mathcal{D}_{e_k}\bar{u})\mathbf{a}(e_k+\mathcal{D}\phi_{e_k}))+\sum_{k=1}^d\mathcal{D}^*\cdot(\mathbf{a}\mathcal{D}(\Upsilon\mathcal{D}_{e_k}\bar{u})\phi_{e_k}).\end{aligned}\tag{4.73}$$

Then, we use equation (4.59) which reads $\partial_t\bar{u}=\frac{1}{2}\bar{\sigma}^2\Delta\bar{u}$ to replace the term $\partial_t\bar{u}$ in the equation above. Notice that here $\frac{1}{2}\bar{\sigma}^2\Delta\bar{u}$ refers to the continuous Laplacian, but using the regularity properties on the function $\bar{u}$ stated in (4.66), we can replace this term by the discrete Laplacian $-\frac{1}{2}\bar{\sigma}^2\mathcal{D}^*\cdot\mathcal{D}\bar{u}$ by paying only a small error. The advantage of this operation is that we can use the two terms $-\frac{1}{2}\bar{\sigma}^2\mathcal{D}^*\cdot(\Upsilon\mathcal{D}\bar{u})$ and $\sum_{k=1}^d\mathcal{D}^*\cdot((\Upsilon\mathcal{D}_{e_k}\bar{u})\mathbf{a}(e_k+\mathcal{D}\phi_{e_k}))$ to make the flux appear: we have

$$\begin{aligned}\sum_{k=1}^d\mathcal{D}^*\cdot((\Upsilon\mathcal{D}_{e_k}\bar{u})\mathbf{a}(e_k+\mathcal{D}\phi_{e_k}))-\frac{1}{2}\bar{\sigma}^2\mathcal{D}^*\cdot(\Upsilon\mathcal{D}\bar{u}) &= \sum_{k=1}^d\mathcal{D}^*\cdot\left((\Upsilon\mathcal{D}_{e_k}\bar{u})\left(\mathbf{a}(e_k+\mathcal{D}\phi_{e_k})-\frac{1}{2}\bar{\sigma}^2e_k\right)\right)\\ &= \sum_{k=1}^d\mathcal{D}^*(\Upsilon\mathcal{D}_{e_k}\bar{u})\cdot\tilde{\mathbf{g}}^*_{e_k},\end{aligned}\tag{4.74}$$

where $\widetilde{\mathbf{g}}^*_{e_k}$ is a translated version of the flux $\widetilde{\mathbf{g}}_{e_k}$ defined by the formula, for each $x \in \mathscr{C}_\infty$,

$$\widetilde{\mathbf{g}}^*_{e_k}(x) := \begin{pmatrix} T_{-e_1}\left[\mathbf{a}\left(\mathcal{D}\phi_{e_k} + e_k\right) - \frac{1}{2}\bar{\sigma}^2 e_k\right]_1 \\ \vdots \\ T_{-e_d}\left[\mathbf{a}\left(\mathcal{D}\phi_{e_k} + e_k\right) - \frac{1}{2}\bar{\sigma}^2 e_k\right]_d \end{pmatrix},$$

where we recall the notation $\left[\mathbf{a}\left(\mathcal{D}\phi_{e_k} + e_k\right) - \frac{1}{2}\bar{\sigma}^2 e_k\right]_i$ introduced in Section 4.1.6 for the ith-component of the vector $\mathbf{a}\left(\mathcal{D}\phi_{e_k} + e_k\right) - \frac{1}{2}\bar{\sigma}^2 e_k$. In Appendix 4.B, it is proved that the translated flux $\widetilde{\mathbf{g}}^*_{e_k}$ has similar properties as the centered flux $\widetilde{\mathbf{g}}_{e_k}$. In particular, it is proved in Remark 4.B that for every radius $r \geqslant \mathcal{M}_{\mathrm{corr},\frac{1}{2}}(0)$,

$$\left\|\widetilde{\mathbf{g}}^*_{e_k}\right\|_{\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r)} \leqslant C r^{\frac{1}{2}}.$$

Combining the identities (4.73) and (4.74), one obtains

$$\begin{aligned}(\partial_t - \mathcal{D}^* \cdot \mathbf{a}\mathcal{D})w = \underbrace{\frac{1}{2}\left(\nabla \cdot \bar{\sigma}^2(\Upsilon \nabla \bar{u}) - (-\mathcal{D}^* \cdot \bar{\sigma}^2(\Upsilon \mathcal{D}\bar{u}))\right)}_{(4.75)\text{-a}} + \underbrace{\sum_{k=1}^d \partial_t(\Upsilon \mathcal{D}_{e_k}\bar{u})\phi_{e_k}}_{(4.75)\text{-b}} \\ + \underbrace{\mathcal{D}^* \cdot ((1-\Upsilon)\mathbf{a}\mathcal{D}\bar{u})}_{(4.75)\text{-c1}} + \underbrace{\frac{1}{2}(\nabla \cdot (\bar{\sigma}^2(1-\Upsilon)\nabla \bar{u})}_{(4.75)\text{-c2}} + \underbrace{\sum_{k=1}^d \mathcal{D}^*(\Upsilon \mathcal{D}_{e_k}\bar{u}) \cdot \widetilde{\mathbf{g}}^*_{e_k}}_{(4.75)\text{-d}} \\ + \underbrace{\sum_{k=1}^d \mathcal{D}^* \cdot (\mathbf{a}\mathcal{D}(\Upsilon \mathcal{D}_{e_k}\bar{u})\phi_{e_k})}_{(4.75)\text{-e}}.\end{aligned} \tag{4.75}$$

There remains to use triangle inequality and estimate the $\underline{L}^2(I_r; \underline{H}^{-1}(\mathscr{C}_\infty \cap B_r))$-norm of each term. The following estimates will be used several times: for $A : I_r \times (\mathscr{C}_\infty \cap B_r) \to \mathbb{R}$ and $B : I_r \times (\mathscr{C}_\infty \cap B_r) \to \mathbb{R}$, one has

$$\begin{aligned}\|AB\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r))} &= \sup_{\|v\|_{\underline{L}^2(I_r;\underline{H}^1(\mathscr{C}_\infty \cap B_r))} \leqslant 1} \frac{1}{|I_r \times (\mathscr{C}_\infty \cap B_r)|} \int_{I_r \times (\mathscr{C}_\infty \cap B_r)} ABv \\ &\leqslant \|A\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r))} \sup_{\|v\|_{\underline{L}^2(I_r;\underline{H}^1(\mathscr{C}_\infty \cap B_r))} \leqslant 1} \|Bv\|_{\underline{L}^2(I_r;\underline{H}^1(\mathscr{C}_\infty \cap B_r))} \\ &\leqslant \|A\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r))} \left(\|B\|_{L^\infty(I_r \times (\mathscr{C}_\infty \cap B_r))} + r\|\nabla B\|_{L^\infty(I_r \times (\mathscr{C}_\infty \cap B_r))}\right).\end{aligned} \tag{4.76}$$

From the definition of the $\underline{L}^2(I_r; \underline{H}^{-1}(\mathscr{C}_\infty \cap B_r))$-norm, one also has the estimate

$$\|A\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r))} \leqslant r\|A\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r))} \leqslant r\|A\|_{L^\infty(I_r \times (\mathscr{C}_\infty \cap B_r))}. \tag{4.77}$$

The term (4.75)-a is a difference between a discrete derivative and a continuous derivative; it can be estimated in terms of the third derivative of the function $\bar{u}$. Using the estimates (4.66) and (4.77) shows

$$\begin{aligned}\|(4.75)\text{-a}\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r))} &\leqslant r\left\|\nabla^2(\Upsilon \nabla \bar{u})\right\|_{L^\infty(I_r \times B_r)} \\ &\leqslant C r^{-1}\left(\frac{r}{r'}\right)^{3+\frac{d}{2}} \|\nabla f\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))}.\end{aligned} \tag{4.78}$$

A similar strategy can be used to estimate the term (4.75)-b

$$\begin{aligned}\|(4.75)\text{-b}\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r))} &\leqslant \left(r\,\|\partial_t\nabla(\Upsilon\nabla\bar{u})\|_{L^\infty(I_r\times B_r)} + \|\partial_t(\Upsilon\nabla\bar{u})\|_{L^\infty(I_r\times B_r)}\right)\sum_{k=1}^d \|\phi_{e_k}\|_{\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r)} \\ &\leqslant C\frac{r}{(r')^3}\left(\frac{r}{r'}\right)^{\frac{2+d}{2}}\|\nabla f\|_{\underline{L}^{2+\delta}(I_r\times(\mathscr{C}_\infty\cap B_r))}\sum_{k=1}^d\|\phi_{e_k}\|_{\underline{L}^2(\mathscr{C}_\infty\cap B_r)} \\ &\leqslant Cr^{-\frac{3}{2}}\left(\frac{r}{r'}\right)^{4+\frac{d}{2}}\|\nabla f\|_{\underline{L}^{2+\delta}(I_r\times(\mathscr{C}_\infty\cap B_r))},\end{aligned} \tag{4.79}$$

where we use the assumption $r \geqslant \mathcal{M}_{\mathrm{hom},\delta}(0) \geqslant \mathcal{M}_{\mathrm{corr},\frac{1}{2}}(0)$ to obtain the sublinearity of the corrector and the regularity estimate (4.66) to go from the second line to the third line.

To estimate the term (4.75)-c1, we note that the function $(1-\Upsilon)$ is equal to 0 outside a mesoscopic boundary layer of size r' of the ball B_r. We thus apply the Meyers estimate (4.69), with the exponent $\delta' = \min\left(\delta_0, \frac{1}{2}\delta\right)$, and the Hölder inequality. This shows

$$\begin{aligned}\|(4.75)\text{-c1}\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r))} &\leqslant \|(1-\Upsilon)\mathbf{a}\mathcal{D}\bar{u}\|_{\underline{L}^2(I_r\times(\mathscr{C}_\infty\cap B_r))} \\ &\leqslant \|1-\Upsilon\|_{\underline{L}^{\frac{4+2\delta'}{\delta'}}(I_r\times B_r)}\|\nabla\bar{u}\|_{\underline{L}^{2+\delta'}(I_r\times B_r)} \\ &\leqslant C\left(\frac{r'}{r}\right)^{\frac{\delta'}{4+2\delta'}}\|\nabla f\|_{\underline{L}^{2+\delta}(I_r\times(\mathscr{C}_\infty\cap B_r))}.\end{aligned} \tag{4.80}$$

where we used the Hölder inequality to go from the first line to the second line and the Meyers estimate to go from the second line to the third line.

We want to apply a similar technique to treat the term (4.75)-c2 since it is also a boundary layer term. However, we should notice that here the derivative ∇ is the continuous gradient defined on $\mathbb{R}^d$ and there is no conductance $\mathbf{a}$, thus we cannot apply a discrete integration by part on the cluster. We will focus on this term later in Step 4.

To estimate the term (4.75)-d, we apply the inequality (4.76), the regularity estimate (4.66), and we use the assumption $r \geqslant \mathcal{M}_{\mathrm{hom},\delta}(0) \geqslant \mathcal{M}_{\mathrm{flux},\frac{1}{2}}(0)$. We obtain

$$\begin{aligned}\|(4.75)\text{-d}\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r))} &\leqslant \left(\|\nabla(\Upsilon\nabla\bar{u})\|_{L^\infty(I_r\times B_r)} + r\left\|\nabla^2(\Upsilon\nabla\bar{u})\right\|_{L^\infty(I_r\times B_r)}\right)\sum_{k=1}^d\left\|\tilde{\mathbf{g}}^*_{e_k}\right\|_{\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r)} \\ &\leqslant Cr^{-\frac{1}{2}}\left(\frac{r}{r'}\right)^{3+\frac{d}{2}}\|\nabla f\|_{\underline{L}^{2+\delta}(I_r\times(\mathscr{C}_\infty\cap B_r))}.\end{aligned} \tag{4.81}$$

The term (4.75)-e can be estimated thanks to an integration by part and the regularity estimate (4.66). This yields

$$\begin{aligned}\|(4.75)\text{-e}\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r))} &\leqslant \|\nabla(\Upsilon\nabla\bar{u})\|_{L^\infty(I_r\times B_r)}\sum_{k=1}^d\|\phi_{e_k}\|_{\underline{L}^2(I_r\times(\mathscr{C}_\infty\cap B_r))} \\ &\leqslant Cr^{-\frac{1}{2}}\left(\frac{r}{r'}\right)^{2+\frac{d}{2}}\|\nabla f\|_{\underline{L}^{2+\delta}(I_r\times(\mathscr{C}_\infty\cap B_r))}.\end{aligned} \tag{4.82}$$

Step 4: Control over the term $\left\|\nabla\cdot(\bar{\sigma}^2(1-\Upsilon)\nabla\bar{u})\right\|_{\underline{L}^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r))}$. As was already mentioned, we cannot use a discrete integration by parts to estimate the $L^2(I_r;\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r))$-norm of this term. The strategy relies on the interior regularity estimate (4.63) which requires careful treatments since it is close to the boundary. We apply the Whitney decomposition on the ball B_r stated below with a minor adaptation to triadic cubes.

Lemma 4.3.2 (Whitney decomposition). *There exits a family of closed triadic cubes* $\{Q_j\}_{j\geqslant 0}$ *such that*

(i) $B_r = \bigcup_j Q_j$ *and the cubes* Q_j *have disjoint interiors;*

(ii) $\sqrt{d}\,\mathrm{size}(Q_j) \leqslant \mathrm{dist}(Q_j, \partial B_r) \leqslant 4\sqrt{d}\,\mathrm{size}(Q_j)$*;*

(iii) Two neighboring cubes Q_j *and* Q_k *have comparable sizes in the sense that*

$$\frac{1}{3} \leqslant \frac{\mathrm{size}(Q_k)}{\mathrm{size}(Q_j)} \leqslant 3;$$

(iv) Each cube Q_j *has at most* $C(d)$ *neighbors.*

We skip the construction of this partition, refer to [216, Theorem 3] or [127, Appendix J] for the proof and to Figure 4.4 for an illustration. With the help of this decomposition, we can estimate the norm $\left\| \nabla \cdot (\bar\sigma^2 (1-\Upsilon) \nabla \bar u) \right\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathscr{C}_\infty \cap B_r))}$. We first relabel the cubes of the decomposition according to their size; we write

$$\{Q_j\}_{i \geqslant 1} := \bigcup_{n=0}^{\infty} \bigcup_{k=1}^{M_n} \{Q_{n,k}\}, \qquad 3^{-(n+1)} r \leqslant \mathrm{size}(Q_{n,k}) < 3^{-n} r.$$

where M_n is the number of the cubes whose size belongs to the interval $[3^{-(n+1)}r, 3^{-n}r)$. Then, we decompose the set $\mathrm{supp}(1-\Upsilon)$ into two parts (see Figure 4.4)

$$\begin{aligned} \mathrm{supp}(1-\Upsilon) &= \Pi_1 \sqcup \Pi_2, \\ &\Pi_1 := \left\{ (t,x) \in I_r \times B_r : \mathrm{dist}(x, \partial B_r) \leqslant 2r', -r^2 + 2(r')^2 \leqslant t \leqslant 0 \right\}, \\ &\Pi_2 := \left\{ (t,x) \in I_r \times B_r : -r^2 \leqslant t \leqslant -r^2 + 2(r')^2 \right\}. \end{aligned}$$

We estimate the weak norm thanks to its definition: we let φ be a function from $I_r \times (\mathscr{C}_\infty \cap B_r)$ to $\mathbb{R}$ which satisfies $\|\varphi\|_{\underline{L}^2(I_r; \underline{H}^1(\mathscr{C}_\infty \cap B_r))} \leqslant 1$ and is equal to 0 on the boundary $I_r \times \partial_{\mathbf{a}} (\mathscr{C}_\infty \cap B_r)$. We split the integral

$$\int_{I_r} \int_{\mathscr{C}_\infty \cap B_r} \nabla\cdot(\bar\sigma^2(1-\Upsilon)\nabla \bar u)\varphi = \int_{(I_r \times \mathbb{Z}^d) \cap \Pi_1} \nabla\cdot(\bar\sigma^2(1-\Upsilon)\nabla \bar u)\varphi + \int_{(I_r \times \mathbb{Z}^d) \cap \Pi_2} \nabla\cdot(\bar\sigma^2(1-\Upsilon)\nabla \bar u)\varphi,$$

and treat the two terms separately.

Step 4.1: Control of the weak norm over Π_1. For the term involving the set Π_1, we use the Whitney decomposition to integrate on every cube of the partition. We first introduce the time intervals, for $m, n \in \mathbb{N}$, $I_{1,m} := -m + (-1, 0]$ and $nI_{1,m} := -m + (-n, 0]$, and partition the boundary layer Π_1 according to the formula

$$\Pi_1 = \bigcup_{m=0}^{\lfloor r^2 - (r')^2 \rfloor} \bigcup_{n=0}^{\infty} \bigcup_{k=1}^{M_n} \left(I_{1,m} \times Q_{n,k} \right) \cap \Pi_1.$$

We remark that we can restrict our attention to the cubes whose sizes is between 1 and r' thanks to the properties of the Whitney decomposition. Indeed, the cubes of size larger than r' remain outside the boundary layer Π_1, since the distance of a cube to the boundary is comparable to its size. On the other hand, the cubes of size smaller than 1 will not contain a point in the lattice $\mathrm{int}(\mathrm{Cv}(\mathbb{Z}^d \cap B_r))$, as these cubes are too close to the boundary, and the definition (4.57) implies that all the points of the lattice in the interior $\mathrm{int}(\mathrm{Cv}(\mathbb{Z}^d \cap B_r))$

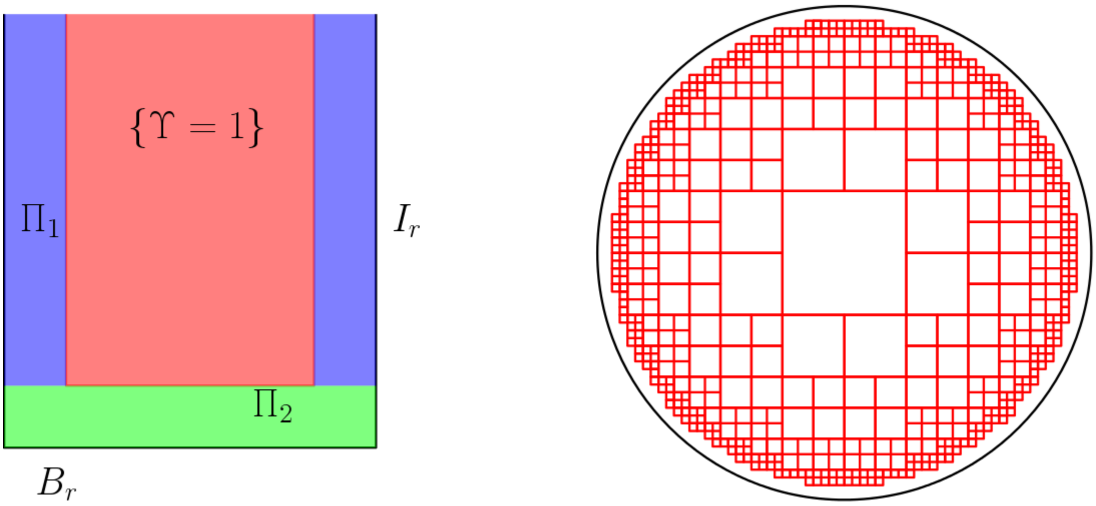

Figure 4.4: The figure on the left illustrates the partition of the cylinder $I_r \times B_r$, where the domain in blue stands for the set Π_1, the one in green for the set Π_2 and the one in red for $\{(t,x) : \Upsilon(t,x) = 1\}$. The figure on the right illustrates a Whitney decomposition in the ball B_r, where we use dyadic cubes to improve the readability.

are at distance at least $\frac{1}{\sqrt{2}}$ from the boundary. We thus have the following partition, if we denote by n_0 and n_1 the integers such that $3^{-(n_0+1)} r \leqslant r' < 3^{-n_0} r$, and $3^{-(n_1+1)} r \leqslant 1 < 3^{-n_1} r$,

$$\Pi_1 = \bigcup_{m=0}^{\lfloor r^2-(r')^2 \rfloor} \bigcup_{n=n_0}^{n_1} \bigcup_{k=1}^{M_n} \left(I_{1,m} \times Q_{n,k} \right) \cap \Pi_1.$$

Using this partition, we can split the integral

$$\int_{(I_r \times \mathscr{C}_\infty) \cap \Pi_1} \nabla \cdot (\bar\sigma^2 (1-\Upsilon) \nabla \bar u) \varphi = \sum_{m=0}^{\lfloor r^2-(r')^2 \rfloor} \sum_{n=n_0}^{n_1} \sum_{k=1}^{M_n} \int_{(I_r \times \mathscr{C}_\infty) \cap (I_{1,m} \times Q_{n,k})} \nabla \cdot (\bar\sigma^2 (1-\Upsilon) \nabla \bar u) \varphi. \tag{4.83}$$

We fix a cylinder $I_{1,m} \times Q_{n,k}$, apply the Cauchy-Schwarz inequality and use the interior regularity estimate (4.63) of the function $\bar u$ in the cylinder $I_{1,m} \times Q_{n,k}$ with the property that the distance $\mathrm{dist}(Q_{n,k}, \partial B_r)$ is larger than $\sqrt{d}\,\mathrm{size}(Q_{n,k})$ and the inclusion $2Q_{n,k} \subseteq B_r$. We obtain

$$\begin{aligned}\left| \int_{(I_r \times \mathscr{C}_\infty) \cap (I_{1,m} \times Q_{n,k})} \nabla \cdot (\bar\sigma^2 (1-\Upsilon) \nabla \bar u) \varphi \right| &\leqslant \left\| \nabla \cdot (\bar\sigma^2 (1-\Upsilon) \nabla \bar u) \right\|_{L^2(I_{1,m} \times Q_{n,k})} \|\varphi\|_{L^2(I_{1,m} \times (\mathscr{C}_\infty \cap Q_{n,k})} \\ &\leqslant C (3^{-n} r)^{-1} \|\nabla \bar u\|_{L^2(2I_{1,m} \times 2Q_{n,k})} \|\varphi\|_{L^2(I_{1,m} \times (\mathscr{C}_\infty \cap Q_{n,k})}.\end{aligned}$$

We sum over all the cubes $\{Q_{n,k}\}_{1 \leqslant k \leqslant M_n}$ and apply the Cauchy-Schwarz inequality

$$\begin{aligned}&\sum_{k=1}^{M_n} \left| \int_{(I_r \times \mathscr{C}_\infty) \cap (I_{1,m} \times Q_{n,k})} \nabla \cdot (\bar\sigma^2 (1-\Upsilon) \nabla \bar u) \varphi \right| \\ &\quad \leqslant \sum_{k=1}^{M_n} C (3^{-n} r)^{-1} \|\nabla \bar u\|_{L^2(2I_{1,m} \times 2Q_{n,k})} \|\varphi\|_{L^2(I_{1,m} \times (\mathscr{C}_\infty \cap Q_{n,k})} \\ &\quad \leqslant C (3^{-n} r)^{-1} \left(\sum_{k=1}^{M_n} \|\nabla \bar u\|^2_{L^2(2I_{1,m} \times 2Q_{n,k})} \right)^{\frac12} \|\varphi\|_{L^2\left(I_{1,m} \times \left(\mathscr{C}_\infty \cap \left(\sqcup_{k=1}^{M_n} Q_{n,k}\right)\right)\right)}.\end{aligned}$$

We then use the following three ingredients:

- Given a discrete set $A \subseteq \mathbb{Z}^d$, the L^2-norm of coarsened function $[\varphi]_{\mathcal{P}}$ over A is larger than the one of the function φ over the set $\mathscr{C}_\infty \cap A$;
- We have the inclusion $\sqcup_{k=1}^{M_n} Q_{n,k} \subseteq \{x \in B_r : \mathrm{dist}(x, \partial B_r) \leqslant 5 \times 3^{-n}\sqrt{d} r\}$;
- We choose the vertex $z(\cdot)$ (defined in (4.40)) to be a point on the boundary ∂B_r for the cubes of the partition intersecting ∂B_r. With this convention, the coarsened function $[\varphi]_{\mathcal{P}}$ is equal to zero on ∂B_r, so we can apply the Poincaré inequality for $[\varphi]_{\mathcal{P}}$ in the boundary layer $\{x \in B_r : \mathrm{dist}(x, \partial B_r) \leqslant 5 \times 3^{-n}\sqrt{d} r\}$.

We obtain the estimate

$$\begin{aligned}(3^{-n} r)^{-1} \left\| \varphi \right\|_{L^2\left(I_{1,m} \times \left(\mathscr{C}_\infty \cap \left(\sqcup_{k=1}^{M_n} Q_{n,k}\right)\right)\right)} &\leqslant C (3^{-n} r)^{-1} \left\| [\varphi]_{\mathcal{P}} \mathbf{1}_{\left\{\{x \in B_r : \mathrm{dist}(x, \partial B_r) \leqslant 5 \times 3^{-n}\sqrt{d} r\}\right\}} \right\|_{L^2(I_{1,m} \times (\mathbb{Z}^d \cap B_r))} \\ &\leqslant C \left\| \nabla [\varphi]_{\mathcal{P}} \mathbf{1}_{\left\{\{x \in B_r : \mathrm{dist}(x, \partial B_r) \leqslant 5 \times 3^{-n}\sqrt{d} r\}\right\}} \right\|_{L^2(I_{1,m} \times (\mathbb{Z}^d \cap B_r))} \\ &\leqslant C \left\| \nabla [\varphi]_{\mathcal{P}} \right\|_{L^2(I_{1,m} \times (\mathbb{Z}^d \cap B_r))}.\end{aligned}$$

We put these estimates back into (4.83) and apply once again the Cauchy-Schwarz inequality. We notice that summing over the integers between n_0 and n_1 gives an additional error term of order $\log^{\frac{1}{2}}(1+r)$,

$$\begin{aligned}&\left| \int_{(I_r \times \mathscr{C}_\infty) \cap \Pi_1} \nabla \cdot (\bar{\sigma}^2 (1 - \Upsilon) \nabla \bar{u}) \varphi \right| \\ &\leqslant C \sum_{m=0}^{\lfloor r^2 - (r')^2 \rfloor} \sum_{n=n_0}^{n_1} \left(\sum_{k=1}^{M_n} \left\| \nabla \bar{u} \right\|^2_{L^2(2 I_{1,m} \times 2 Q_{n,k})} \right)^{\frac{1}{2}} \left\| \nabla [\varphi]_{\mathcal{P}} \right\|_{L^2(I_{1,m} \times (\mathbb{Z}^d \cap B_r))} \\ &\leqslant C \log^{\frac{1}{2}}(1+r) \left(\sum_{m=0}^{\lfloor r^2 - (r')^2 \rfloor} \sum_{n=n_0}^{n_1} \sum_{k=1}^{M_n} \left\| \nabla \bar{u} \right\|^2_{L^2(2 I_{1,m} \times 2 Q_{n,k})} \right)^{\frac{1}{2}} \left\| \nabla [\varphi]_{\mathcal{P}} \right\|_{L^2(I_r \times (\mathbb{Z}^d \cap B_r))}.\end{aligned} \tag{4.84}$$

We then estimate the norm $\left\| \nabla [\varphi]_{\mathcal{P}} \right\|_{L^2(I_r \times (\mathbb{Z}^d \cap B_r))}$ thanks to the inequalities (4.42), (4.39), and the assumption $r > \mathcal{M}_q(\mathcal{P})$. We obtain

$$\frac{1}{|I_r \times (\mathscr{C}_\infty \cap B_r)|^{\frac{1}{2}}} \left\| \nabla [\varphi]_{\mathcal{P}} \right\|_{L^2(I_r \times (\mathbb{Z}^d \cap B_r))} \leqslant C r^{\frac{2d-1}{2q}} \left\| \nabla \varphi \right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r))} \leqslant C r^{\frac{2d-1}{2q}}.$$

Moreover, by the properties of the Whitney covering, the sum $\sum_{i=0}^{\lfloor r^2 - (r')^2 \rfloor} \sum_{n=n_0}^{n_1} \sum_{k=1}^{M_n} \left\| \nabla \bar{u} \right\|^2_{L^2(2 I_{1,m} \times 2 Q_{n,k})}$ can be estimated by the L^2-norm of the function $\nabla \bar{u}$ in a boundary layer of size $6r'$ of the parabolic cylinder $I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r)$ (since every point in the ball B_r belongs to at most $C(d)$ cubes of the form $2Q_j$). More specifically, we have the estimate

$$\left(\frac{1}{|I_r \times (\mathscr{C}_\infty \cap B_r)|} \sum_{m=0}^{\lfloor r^2 - (r')^2 \rfloor} \sum_{n=n_0}^{n_1} \sum_{k=1}^{M_n} \left\| \nabla \bar{u} \right\|^2_{L^2(2 I_{1,m} \times 2 Q_{n,k})} \right)^{\frac{1}{2}} \leqslant C \left\| \nabla \bar{u} \right\|_{\underline{L}^2(I_r \times \{x \in B_r : \mathrm{dist}(x, \partial B_r) \leqslant 6r'\})}.$$

We then apply the Hölder's inequality and the global Meyers estimate (4.71) with the exponent $\delta' = \min\left(\delta_0, \frac{1}{2}\delta\right)$. We obtain

$$\begin{aligned}\left(\frac{1}{|I_r \times (\mathscr{C}_\infty \cap B_r)|} \sum_{m=0}^{\lfloor r^2 - (r')^2 \rfloor} \sum_{n=n_0}^{n_1} \sum_{k=1}^{M_n} \left\| \nabla \bar{u} \right\|^2_{L^2(2 I_{1,m} \times 2 Q_{n,k})} \right)^{\frac{1}{2}} &\leqslant C \left\| \nabla \bar{u} \right\|_{\underline{L}^2(I_r \times \{x \in B_r : \mathrm{dist}(x, \partial B_r) \leqslant 6r'\})} \\ &\leqslant C \left(\frac{r'}{r} \right)^{\frac{\delta'}{4 + 2\delta'}} \left\| \nabla f \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))}.\end{aligned}$$

We conclude that

$$\left| \frac{1}{|I_r \times (\mathscr{C}_\infty \cap B_r)|} \int_{(I_r \times \mathscr{C}_\infty) \cap \Pi_1} \nabla \cdot (\bar{\sigma}^2 (1 - \Upsilon) \nabla \bar{u}) \varphi \right| \\ \leqslant C \log^{\frac{1}{2}}(1+r) r^{\frac{2d-1}{2q}} \left(\frac{r'}{r} \right)^{\frac{\delta'}{4+2\delta'}} \left\| \nabla f \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))}. \tag{4.85}$$

Step 4.2: Control the weak norm over Π_2. One can repeat all the arguments above to estimate the weak norm over the set Π_2, but we should pay attention to the decomposition over the time interval I_r since now the support of Π_2 is close to the time boundary (see Figure 4.4). We define the time intervals

$$\forall m \in \mathbb{N}, \qquad I_{2,m} := -r^2 + \left(\frac{2}{3}\right)^m (r')^2 + \left(-\frac{1}{3} \times \left(\frac{2}{3}\right)^m (r')^2, 0\right],$$

so that they satisfy $2I_{2,m} \subseteq I_r$. We can then apply the same arguments as in the estimates (4.84) and (4.85) to obtain the inequality

$$\left| \frac{1}{|I_r \times (\mathscr{C}_\infty \cap B_r)|} \int_{(I_r \times \mathscr{C}_\infty) \cap \Pi_2} \nabla \cdot (\bar{\sigma}^2 (1 - \Upsilon) \nabla \bar{u}) \varphi \right| \\ \leqslant C \log^{\frac{1}{2}}(1+r) r^{\frac{2d-1}{2q}} \left(\frac{1}{r^2} \int_{-r^2}^{-r^2 + 2(r')^2} \left\| \nabla \bar{u}(t, \cdot) \right\|^2_{\underline{L}^2(\mathrm{Cv}(\mathbb{Z}^d \cap B_r))} \, dt \right)^{\frac{1}{2}}.$$

Then, we apply Hölder's inequality in the time variable and the estimate (4.71) to obtain

$$\left(\frac{1}{r^2} \int_{-r^2}^{-r^2 + 2(r')^2} \left\| \nabla \bar{u}(t, \cdot) \right\|^2_{\underline{L}^2(\mathrm{Cv}(\mathbb{Z}^d \cap B_r))} \, dt \right)^{\frac{1}{2}} \leqslant \left(\frac{r'}{r} \right)^{\frac{2\delta'}{4+2\delta'}} \left\| \nabla f \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))}. \tag{4.86}$$

This gives an estimate for the weak norm of the map $\nabla \cdot (\bar{\sigma}^2 (1 - \Upsilon) \nabla \bar{u})$ over the set Π_2. Finally, we combine the estimates (4.85) and (4.86) to conclude that

$$\left\| \nabla \cdot (\bar{\sigma}^2 (1 - \Upsilon) \nabla \bar{u}) \right\|_{\underline{L}^2(I_r; \underline{H}^{-1}(\mathscr{C}_\infty \cap B_r))} \leqslant C \log^{\frac{1}{2}}(1+r) r^{\frac{2d-1}{2q}} \left(\frac{r'}{r} \right)^{\frac{\delta'}{4+2\delta'}} \left\| \nabla f \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))}. \tag{4.87}$$

Step 5: Choice of the parameters q, β *and conclusion.* We conclude the proof by combing the estimates (4.78), (4.79), (4.80), (4.81), (4.82), (4.87) and by choosing $r' = r^{1-\beta}$ for some small exponent $\beta \in (0, 1/2]$ to obtain

$$\frac{1}{r} \left\| u - \bar{u} \right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r))} \leqslant C \mathcal{E}(r, \beta, q) \left\| \nabla f \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))}, \tag{4.88}$$

where the quantity $\mathcal{E}(r, \beta, q)$ is defined by the formula

$$\mathcal{E}(r, \beta, q) := r^{-\frac{1}{2} + \beta\left(3 + \frac{d}{2}\right)} + \log^{\frac{1}{2}}(1+r) r^{\frac{2d-1}{2q} - \frac{\beta\delta'}{4+2\delta'}}, \tag{4.89}$$

where we recall that $\delta' = \min\left(\delta_0, \frac{1}{2}\delta\right)$ and that δ_0 is the exponent given by the Meyers estimate stated in (4.69).

It remains to select a value for the exponents β and q. We first choose the value of the exponent β and set $\beta := \frac{1}{12+2d}$ so that the first term in the right side of (4.89) is equal to $r^{-\frac{1}{4}}$.

Then we set $q := \frac{(12+2d)(2d-1)(4+2\delta')}{\delta'}$ so that $\frac{2d-1}{2q} = \frac{\beta\delta'}{2(4+2\delta')}$. With this choice, the second term in the right side of (4.89) is equal to $\log^{\frac{1}{2}}(1+r) r^{-\frac{\beta\delta'}{8+4\delta'}}$.

We obtain that Theorem 4.3.2 holds with the exponent $\alpha := \frac{\delta'}{(12+2d)(16+8\delta')} > 0$. The proof is complete.

□

Large-scale $C^{0,1}$-regularity estimate

The objective of this section is to prove the following $C^{0,1}$-large-scale regularity estimate for $\mathbf{a}$-caloric functions on the infinite cluster.

Proposition 4.3.1. *There exist a constant $C(d,\lambda,\mathfrak{p}) < \infty$, an exponent $s(d,\lambda,\mathfrak{p}) > 0$ such that for each point $y \in \mathbb{Z}^d$, there exists a non-negative random variable $\mathcal{M}_{C^{0,1}-\mathrm{reg}}(y)$ satisfying*

$$\mathcal{M}_{C^{0,1}-\mathrm{reg}}(y) \leqslant \mathcal{O}_s(C) \tag{4.90}$$

such that, for every $r \geqslant \mathcal{M}_{C^{0,1}-\mathrm{reg}}(y)$, and every weak solution $u \in H^1_{\mathrm{par}}(I_R \times (\mathscr{C}_\infty \cap B_R(y)))$ of the equation

$$\partial_t u - \nabla \cdot (\mathbf{a}\nabla u) = 0 \text{ in } I_R \times (\mathscr{C}_\infty \cap B_R(y)),$$

one has the estimate, for every radius $r \in [\mathcal{M}_{C^{0,1}-\mathrm{reg}}(y), R]$,

$$\sup_{t\in I_r} \|\nabla u(t,\cdot)\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r(y))} \leqslant \frac{C}{R} \left\| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_R \times B_R(y)} \right\|_{\underline{L}^2(I_R \times B_R(y))}. \tag{4.91}$$

Remark. The right side of the estimate involves the coarsened function $[u]_{\mathcal{P}}$ and we do not try to remove the coarsening to obtain a result of the form

$$\sup_{t\in I_r} \|\nabla u(t,\cdot)\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r(y))} \leqslant \frac{C}{R} \left\| u - (u)_{I_R \times (\mathscr{C}_\infty \cap B_R(y))} \right\|_{\underline{L}^2(I_R \times (\mathscr{C}_\infty \cap B_R(y)))}, \tag{4.92}$$

even though such a result would be more natural and should be provable. There are two reasons motivating this choice. First, the estimate (4.91) involving the coarsening is simpler to prove than the inequality (4.92) and this choice reduces the amount of technicalities in the proof. Second, the objective of this section is to prove the Lipschitz regularity on the heat-kernel stated in Theorem 4.1.3 and the estimate (4.91) is sufficient in this regard.

This proposition proves that there exists a large random scale above which one has a good control on the gradient of $\mathbf{a}$-caloric functions. Such result belongs to the theory of large-scale regularity which is an important aspect of stochastic homogenization. The result presented above is a percolation version of a known result in the uniformly elliptic setting (see [25, Theorem 8.7]) and can be considered a first step toward the establishment of a general large-scale regularity theory for the parabolic problem on the infinite percolation cluster.

We do not establish such a general theory here but we believe that it should follow from similar arguments: in the elliptic setting a general large-scale regularity theory was established in [19] and the generalization to the parabolic setting should be achievable. The reason justifying this choice is that our objective is to prove an estimate on the gradient of the Green's function (Theorem 4.1.3) and we do not need the full strength of the large-scale regularity theory to prove this result.

The main idea of the proof is that, thanks to Theorem 4.3.2, an $\mathbf{a}$-caloric function is well-approximated by a $\bar{\sigma}^2$-caloric function. It is then possible to transfer the regularity known for $\bar{\sigma}^2$-caloric functions to $\mathbf{a}$-caloric functions following the classical ideas of the regularity theory. Such result can only hold when the $\mathbf{a}$-caloric function is well-approximated by a $\bar{\sigma}^2$-caloric function which, according to Theorem 4.3.2, only holds on large scales.

This strategy has been carried out in [25] and is summarized in the following lemma, for which we refer to [25, Lemma 8.9].

Lemma 4.3.3 (Lemma 8.9 of [25]). *Fix an exponent $\beta \in (0,1]$, $k \geqslant 1$ and $X \geqslant 1$. Let $R \geqslant 4X$ and $v \in L^2(I_R \times B_R)$ have the property that, for every $r \in \left[X, \frac{1}{4}R\right]$, there exists a function $w \in H^1_{\mathrm{par}}(I_r \times B_r)$ which is a weak solution of*

$$\partial_t w - \frac{\bar{\sigma}^2}{2} \Delta w = 0 \text{ in } I_r \times B_r, \tag{4.93}$$

satisfying

$$\|v - w\|_{\underline{L}^2(I_{r/2} \times B_{r/2})} \leqslant K r^{-\beta} \left\| v - (v)_{I_{4r} \times B_{4r}} \right\|_{\underline{L}^2(I_{4r} \times B_{4r})}.$$

Then there exists a constant $C := C(\beta, K, d, \lambda) < \infty$ such that for every radius $r \in [X, R]$,

$$\frac{1}{r} \left\| v - (v)_{I_r \times B_r} \right\|_{\underline{L}^2(I_r \times B_r)} \leqslant \frac{C}{R} \left\| v - (v)_{I_R \times B_R} \right\|_{\underline{L}^2(I_R \times B_R)}.$$

To prove Proposition 4.3.1, we apply the previous lemma and combine it with Theorem 4.3.2. One has to face the following difficulty: we want to apply the previous result in the setting of percolation where the functions are only defined on the infinite cluster and not on $\mathbb{R}^d$ as in the statement of Lemma 4.3.3.

To overcome this issue, the idea is to use the partition $\mathcal{P}$ to extend the function u, using the definition of the coarsened map $[u]_{\mathcal{P}}$ stated in (4.40). The strategy of the proof is then the following:

(i) Proving that the function $[u]_{\mathcal{P}}$ is a good approximation of the function u. In particular we wish to prove that if the map u is well-approximated by a $\bar{\sigma}^2$-caloric function, then the map $[u]_{\mathcal{P}}$ is also well-approximated by a $\bar{\sigma}^2$-caloric function.

(ii) Apply Lemma 4.3.3 to the function $[u]_{\mathcal{P}}$ to obtain a large-scale $C^{0,1}$-regularity estimate for this map.

(iii) Transfer the result from the function $[u]_{\mathcal{P}}$ to the function u.

The details are carried out in the following proof.

Proof. Using the translation invariance of the model, we can assume without loss of generality that $y = 0$. We also let δ_0 be the exponent which appears in the Meyers estimate stated in Proposition 4.2.5 and consider the minimal scale $\mathcal{M}_{\mathrm{Meyers}}(0)$ given by Proposition 4.2.5. We consider the minimal scale $\mathcal{M}_{\mathrm{hom},\delta_0}(0)$ given by Theorem 4.3.2 and we let α be the exponent which appears in the estimate (4.61). We set $q := \max\left(\frac{d}{\alpha}, 4d\right)$ and let $\mathcal{M}_q(\mathcal{P})$ be a minimal scale provided by Proposition 4.2.2. In particular, one has, for each radius $r \geqslant \mathcal{M}_q(\mathcal{P})$,

$$r^{-d} \sum_{x \in \mathbb{Z}^d \cap B_r} \mathrm{size}\left(\square_{\mathcal{P}}(x)\right)^q \leqslant C \text{ and } \sup_{x \in \mathbb{Z}^d \cap B_r} \mathrm{size}\left(\square_{\mathcal{P}}(x)\right) \leqslant r^{\frac{1}{q}}. \tag{4.94}$$

The reasons justifying the choice of the exponents q will be become clear later in the proofs. By Proposition 4.2.2, one knows that the minimal scale $\mathcal{M}_q(\mathcal{P})$ satisfies the stochastic integrability estimate

$$\mathcal{M}_q(\mathcal{P}) \leqslant \mathcal{O}_s(C).$$

We then let $\mathcal{M}_{C^{0,1}-\mathrm{reg}}(0)$ be the minimal scale defined by the formula

$$\mathcal{M}_{C^{0,1}-\mathrm{reg}}(0) := \max\left(\mathcal{M}_q(\mathcal{P}), \mathcal{M}_{\mathrm{Meyers}}(0), \mathcal{M}_{\mathrm{hom},\delta_0}(0)\right).$$

In the rest of the proof, we assume that the radii r and R are always larger than this minimal scale. We also note that, under the assumption $r \geqslant \mathcal{M}_{C^{0,1}-\mathrm{reg}}(0)$, we can compare the volume of the ball B_r and the cardinality of $\mathscr{C}_\infty \cap B_r$, and we have the estimate

$$cr^d \leqslant |\mathscr{C}_\infty \cap B_r| \leqslant Cr^d.$$

This is a consequence of the assumption $\mathcal{M}_{C^{0,1}-\mathrm{reg}}(0) \geqslant \mathcal{M}_{\mathrm{hom},\delta_0}(0)$ and the estimate (4.62).

We apply Theorem 4.3.2 to the function u on the parabolic cylinder $I_r \times (\mathscr{C}_\infty \cap B_r)$, with the boundary condition $f = u$ and with the exponent δ_0; this proves that there exists a function $\bar{u} \in H^1_{\mathrm{par}}(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))$, which is a solution of the equation (4.93), such that

$$\frac{1}{r}\left\|u - \bar{u}\right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r))} \leqslant Cr^{-\alpha}\left\|\nabla u\right\|_{\underline{L}^{2+\delta_0}(I_r \times (\mathscr{C}_\infty \cap B_r))}. \tag{4.95}$$

We split the proof into 4 steps. In the first two steps, we prove that we can apply Lemma 4.3.3 with the coarsened function $[u]_{\mathcal{P}}$, we then post-process the result in Steps 3 and 4 and deduce Proposition 4.3.1.

Step 1. In this step, we post-process the result of Theorem 4.3.2: in the statement of the estimate (4.95), the right-hand side is expressed with an $L^{2+\delta_0}$-norm, for some small strictly positive exponent δ_0. The goal of this step is to remove this additional assumption. To this end, we use the assumption $r \geqslant \mathcal{M}_{C^{0,1}-\mathrm{reg}}(0) \geqslant \mathcal{M}_{\mathrm{Meyers}}(0)$, which implies

$$\left\|\nabla u\right\|_{\underline{L}^{2+\delta_0}(I_r \times (\mathscr{C}_\infty \cap B_r))} \leqslant C\left\|\nabla u\right\|_{\underline{L}^2(I_{2r} \times (\mathscr{C}_\infty \cap B_{2r}))}.$$

We then apply the parabolic Caccioppoli inequality, which is stated in Proposition 4.2.4, and reads

$$\left\|\nabla u\right\|_{\underline{L}^2(I_{2r} \times (\mathscr{C}_\infty \cap B_{2r}))} \leqslant \frac{C}{r}\left\|u - (u)_{I_{4r} \times (\mathscr{C}_\infty \cap B_{4r})}\right\|_{\underline{L}^2(I_{4r} \times (\mathscr{C}_\infty \cap B_{4r}))}.$$

Combining the two previous displays with the inequality (4.95) shows

$$\left\|u - \bar{u}\right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r))} \leqslant Cr^{-\alpha}\left\|u - (u)_{I_{4r} \times (\mathscr{C}_\infty \cap B_{4r})}\right\|_{\underline{L}^2(I_{4r} \times (\mathscr{C}_\infty \cap B_{4r}))}, \tag{4.96}$$

and Step 1 is complete.

Step 2. The goal of this step is to prove that the L^2-norm of the difference $[u]_{\mathcal{P}} - \bar{u}$ on the continuous parabolic cylinder $I_{r/2} \times B_{r/2}$ is small: we prove that there exists an exponent $\beta > 0$ such that

$$\left\|[u]_{\mathcal{P}} - \bar{u}\right\|_{\underline{L}^2(I_{r/2} \times B_{r/2})} \leqslant Cr^{-\beta}\left\|[u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_{4r} \times B_{4r}}\right\|_{\underline{L}^2(I_{4r} \times B_{4r})}.$$

The proof of this inequality relies on the estimate (4.96), which establishes that $\bar{u}$ is a good approximation of u on the infinite cluster, together with the following parabolic regularity result: since $\bar{u}$ is $\bar{\sigma}^2$-caloric on the parabolic cylinder $I_r \times B_r$, one has the estimate

$$\left\|\nabla \bar{u}\right\|_{L^\infty(I_{r/2} \times B_{r/2})} \leqslant Cr^{-1}\left\|\bar{u} - (\bar{u})_{I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r)}\right\|_{\underline{L}^2(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))}. \tag{4.97}$$

We then consider a (continuous) triadic cube $\square$ of the partition $\mathcal{P}$ such that $\square \cap B_{r/2} \neq \emptyset$. By definition of the coarsening stated in (4.41), one sees that, for each time $t \in I_{r/2}$,

$$\begin{aligned}\left\| [u]_{\mathcal{P}}(t,\cdot) - \bar{u}(t,\cdot) \right\|_{L^\infty(\square)} &\leqslant \left\| [u]_{\mathcal{P}}(t,\cdot) - [\bar{u}]_{\mathcal{P}}(t,\cdot) \right\|_{L^\infty(\square)} + \left\| [\bar{u}]_{\mathcal{P}}(t,\cdot) - \bar{u}(t,\cdot) \right\|_{L^\infty(\mathscr{C}_\infty \cap \square)} \\ &= \left\| [u - \bar{u}]_{\mathcal{P}}(t,\cdot) \right\|_{L^\infty(\square)} + \left\| [\bar{u}]_{\mathcal{P}}(t,\cdot) - \bar{u}(t,\cdot) \right\|_{L^\infty(\mathscr{C}_\infty \cap \square)}.\end{aligned}$$

We then note that, by definition of the coarsening stated in (4.41), the L^∞-norm of the function $[u - \bar{u}]_{\mathcal{P}}$ is smaller than the L^∞-norm of the function $u - \bar{u}$. Combining this observation with the estimate (4.43), we obtain

$$\left\| [u]_{\mathcal{P}}(t,\cdot) - \bar{u}(t,\cdot) \right\|_{L^\infty(\square)} \leqslant \left\| u(t,\cdot) - \bar{u}(t,\cdot) \right\|_{L^\infty(\mathscr{C}_\infty \cap \square)} + \operatorname{size}(\square) \left\| \nabla \bar{u}(t,\cdot) \right\|_{L^\infty(\hat{\square})},$$

where the set $\hat{\square}$ stands for the union of the cube $\square$ and all its neighbors in the partition $\mathcal{P}$. We use the $L^\infty - L^2$ estimate, valid in the discrete setting,

$$\left\| u(t,\cdot) - \bar{u}(t,\cdot) \right\|_{L^\infty(\mathscr{C}_\infty \cap \square)} \leqslant \left\| u(t,\cdot) - \bar{u}(t,\cdot) \right\|_{L^2(\mathscr{C}_\infty \cap \square)},$$

together with the regularity estimate (4.97) and the estimate (4.94) on the sizes of the cubes of the partition $\mathcal{P}$ to deduce

$$\begin{aligned}\left\| [u]_{\mathcal{P}} - \bar{u} \right\|_{L^2(I_{r/2} \times \square)} &\leqslant \operatorname{size}(\square)^{d/2} \left\| u - \bar{u} \right\|_{L^2(I_{r/2} \times (\mathscr{C}_\infty \cap \square))} \\ &\quad + C \operatorname{size}(\square)^{1+d/2} r^{-1} \left\| \bar{u} - (\bar{u})_{I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r)} \right\|_{\underline{L}^2(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))}.\end{aligned}$$

We then use the bounds (4.94) to estimate the size of the cube $\square$ and sum over all the cubes of the partition $\mathcal{P}$ which intersect the ball $B_{r/2}$, and use the estimate $1 + \frac{d}{2} \leqslant d$,

$$\left\| [u]_{\mathcal{P}} - \bar{u} \right\|_{\underline{L}^2(I_{r/2} \times B_{r/2})} \leqslant C r^{\frac{\alpha}{2}} \left\| u - \bar{u} \right\|_{\underline{L}^2(I_r \times (\mathscr{C}_\infty \cap B_r))} + C r^{-1+\alpha} \left\| \bar{u} - (\bar{u})_{I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r)} \right\|_{\underline{L}^2(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))}.$$

An estimate on the first term on the right side is provided by the estimate (4.96). For the second term on the right-hand side, we apply the Poincaré inequality ([18, Corollary 3.4]), use the estimate (4.71) for $f = u$, and apply Proposition 4.2.4 (the parabolic Caccioppoli inequality) and Proposition 4.2.5 (the interior Meyers estimate)

$$\begin{aligned}\left\| \bar{u} - (\bar{u})_{I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r)} \right\|_{\underline{L}^2(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))} &\leqslant r \left\| \nabla \bar{u} \right\|_{\underline{L}^2(I_r \times \mathrm{Cv}(\mathbb{Z}^d \cap B_r))} \leqslant r \left\| \nabla u \right\|_{\underline{L}^{2+\delta}(I_r \times (\mathscr{C}_\infty \cap B_r))} \\ &\leqslant C r \left\| \nabla u \right\|_{\underline{L}^2(I_{2r} \times (\mathscr{C}_\infty \cap B_{2r}))} \leqslant C \left\| u - (u)_{I_{4r} \times (\mathscr{C}_\infty \cap B_{4r})} \right\|_{\underline{L}^2(I_{4r} \times (\mathscr{C}_\infty \cap B_{4r}))}.\end{aligned}$$

Thus, we combine the two estimates and obtain

$$\begin{aligned}\left\| [u]_{\mathcal{P}} - \bar{u} \right\|_{\underline{L}^2(I_{r/2} \times B_{r/2})} &\leqslant r^{\frac{\alpha}{2}} r^{-\alpha} \left\| u - (u)_{I_{4r} \times (\mathscr{C}_\infty \cap B_{4r})} \right\|_{\underline{L}^2(I_{4r} \times (\mathscr{C}_\infty \cap B_{4r}))} \\ &\quad + C r^{-1+\alpha} \left\| u - (u)_{I_{4r} \times (\mathscr{C}_\infty \cap B_{4r})} \right\|_{\underline{L}^2(I_{4r} \times (\mathscr{C}_\infty \cap B_{4r}))}. \qquad (4.98)\end{aligned}$$

We then set the value $\beta := \frac{\alpha}{2}$. Since, by the identity (4.89), the value of the exponent α is smaller than $\frac{1}{2}$, we have the inequality $1 - \alpha \geqslant \frac{\alpha}{2}$. This implies

$$\left\| [u]_{\mathcal{P}} - \bar{u} \right\|_{\underline{L}^2(I_{r/2} \times B_{r/2})} \leqslant C r^{-\beta} \left\| u - (u)_{I_{4r} \times (\mathscr{C}_\infty \cap B_{4r})} \right\|_{\underline{L}^2(I_{4r} \times (\mathscr{C}_\infty \cap B_{4r}))}.$$

By definition of the coarsened function $[u]_{\mathcal{P}}$, we also have

$$\begin{aligned}\left\| u - (u)_{I_{4r}\times(\mathscr{C}_\infty\cap B_{4r})} \right\|_{\underline{L}^2(I_{4r}\times(\mathscr{C}_\infty\cap B_{4r}))} &\leqslant \left\| u - ([u]_{\mathcal{P}})_{I_{4r}\times B_{4r}} \right\|_{\underline{L}^2(I_{4r}\times(\mathscr{C}_\infty\cap B_{4r}))} \\ &\leqslant C \left\| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_{4r}\times B_{4r}} \right\|_{\underline{L}^2(I_{4r}\times B_{4r})}. \end{aligned} \tag{4.99}$$

The proof of Step 2 is complete.

Step 3. In the two previous steps, we proved that the coarsened function $[u]_{\mathcal{P}}$ satisfies the assumption of Lemma 4.3.3, with the choice $X = \mathcal{M}_{C^{0,1}-\mathrm{reg}}(0)$. We consequently apply the lemma and obtain that there exists a constant $C := C(d,\mathfrak{p},\lambda) < \infty$ such that, for every pair of radii r, R satisfying $R \geqslant r \geqslant \mathcal{M}_{C^{0,1}-\mathrm{reg}}(0)$,

$$\frac{1}{r}\left\| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_r\times B_r} \right\|_{\underline{L}^2(I_r\times B_r)} \leqslant \frac{C}{R}\left\| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_R\times B_R} \right\|_{\underline{L}^2(I_R\times B_R)}. \tag{4.100}$$

Then, we apply once again the estimate (4.99) for the left-hand side of (4.100) in $I_r\times(\mathscr{C}_\infty\cap B_r)$ and we obtain

$$\frac{1}{r}\left\| u - (u)_{I_r\times(\mathscr{C}_\infty\cap B_r)} \right\|_{\underline{L}^2(I_r\times(\mathscr{C}_\infty\cap B_r))} \leqslant \frac{C}{R}\left\| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_R\times B_R} \right\|_{\underline{L}^2(I_R\times B_R)}. \tag{4.101}$$

Step 4. In this final step, we upgrade the large-scale $C^{0,1}$-regularity estimate into the estimate (4.91). The strategy is to use an $L^\infty_t L^2_x$ regularity estimate which is valid for the $\mathbf{a}$-caloric functions since the environment $\mathbf{a}$ is assumed to be time independent. This result is stated in Lemma 4.2.2 and we apply it to the function u to obtain, for each $r \geqslant 1$,

$$\sup_{t\in I_{r/2}} \left\| \nabla u(t,\cdot) \right\|_{\underline{L}^2(\mathscr{C}_\infty\cap B_{r/2})} \leqslant \frac{C}{r}\left\| u - (u)_{I_r\times(\mathscr{C}_\infty\cap B_r)} \right\|_{\underline{L}^2(I_r\times(\mathscr{C}_\infty\cap B_r))}.$$

Combining this result with (4.101) completes the proof of Proposition 4.3.1 with the radius $\frac{r}{2}$ instead of r; this is a minor difference which can be fixed by standard arguments. □

Decay of the gradient of the heat kernel

The objective of this section is to post-process the regularity theory established in Proposition 4.3.1 and to apply it to the heat kernel. Together with Theorem 4.3.1, we deduce Theorem 4.1.3.

Proof of Theorem 4.1.3. We fix a time $t \in (0,\infty)$, two points $x, y \in \mathbb{Z}^d$ and work on the event $\{y \in \mathscr{C}_\infty\}$. We let $\mathcal{M}_{C^{0,1}-\mathrm{reg}}(x)$ be the minimal scale provided by Proposition 4.3.1 and, for $z \in \mathbb{Z}^d$, we let $\mathcal{T}_{\mathrm{NA}}(z)$ be the minimal time provided by Theorem 4.3.1. We first define the minimal time $\mathcal{T}'_{\mathrm{NA}}(x)$ by the formula

$$\mathcal{T}'_{\mathrm{NA}}(x) := \sup\left\{t \in [1,\infty) \ : \ \exists z \in B_t(x) \text{ such that } \mathcal{T}_{\mathrm{NA}}(z) \geqslant t\right\}, \tag{4.102}$$

so that for every time t larger than this minimal time, every point $z \in \mathscr{C}_\infty \cap B_t(x)$, and every point $z' \in \mathscr{C}_\infty$, one has the estimate

$$p(t,z,z') \leqslant \Phi_C\left(t, z - z'\right).$$

By using the symmetry of the heat kernel, we also have, for each time $t \geqslant \mathcal{T}'_{\mathrm{NA}}(x)$, for each point $z \in \mathscr{C}_\infty \cap B_t(x)$, and each point $z' \in \mathscr{C}_\infty$,

$$p(t,z',z) \leqslant \Phi_C\left(t, z - z'\right). \tag{4.103}$$

Additionally, we claim that this minimal time satisfies the stochastic integrability estimate

$$\mathcal{T}'_{\mathrm{NA}}(x) \leqslant \mathcal{O}_s(C). \tag{4.104}$$

The proof of the estimate (4.104) relies on an application of Lemma 4.1.1 by choosing

$$X_n := \sup_{z \in \mathbb{Z}^d \cap B_{3^n}(x)} 3^{-n} \mathcal{T}_{\mathrm{NA}}(z),$$

and we refer to the computation (4.47) for the details of the argument. We then define the minimal scale

$$\mathcal{M}_{\mathrm{reg}}(x) := \max\left(\mathcal{M}_{C^{0,1}-\mathrm{reg}}(x), \sqrt{\mathcal{T}'_{\mathrm{NA}}(x)}\right). \tag{4.105}$$

Using the definition of the $\mathcal{O}_s$ notation, the stochastic integrability estimates (4.90) and (4.104), one has, by reducing the size of the exponent s if necessary,

$$\mathcal{M}_{\mathrm{reg}}(x) \leqslant \mathcal{O}_s(C).$$

In particular the tail of the random variable $\mathcal{M}_{\mathrm{reg}}(x)$ satisfies the inequality (4.20). We define, for $\tau \in [-t, \infty)$ and $z \in \mathscr{C}_\infty$,

$$u(\tau, z) := p(t + 1 + \tau, z, y).$$

We let $R := \frac{\sqrt{t}}{2}$ and note that the function u is solution of the parabolic equation on the cylinder $I_R \times (\mathscr{C}_\infty \cap B_R(x))$. Applying Proposition 4.3.1 with the values r, R and using the assumption $R \geqslant r \geqslant \mathcal{M}_{\mathrm{reg}}(x)$, we obtain

$$\sup_{\tau \in I_r} \|\nabla u(\tau, \cdot)\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r(x))} \leqslant \frac{C}{t^{1/2}} \left\| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_R \times B_R(x)} \right\|_{\underline{L}^2(I_R \times B_R(x))}. \tag{4.106}$$

We then note that the assumption $R \geqslant \mathcal{M}_{\mathrm{reg}}(x)$ implies $\frac{t}{4} \geqslant \mathcal{T}'_{\mathrm{NA}}(x)$ and allows to apply the estimate (4.103) to bound the right side of the previous display. This shows

$$\left\| [u]_{\mathcal{P}} - ([u]_{\mathcal{P}})_{I_R \times B_R(x)} \right\|_{\underline{L}^2(I_R \times B_R(x))} \leqslant \Phi_C(t, y - x).$$

Combining the previous display with (4.106), considering the specific value $\tau = -1$, and increasing the value of the constant C shows

$$\|\nabla p(t, \cdot, y)\|_{\underline{L}^2(\mathscr{C}_\infty \cap B_r(x))} \leqslant t^{-1/2} \Phi_C(t, y - x).$$

The proof of Theorem 4.1.3 is complete.

□

4.4 Quantitative homogenization of the heat kernel

In this section, we establish the theorem of quantitative homogenization of the parabolic Green's function, Theorem 4.1.1. From now on, we fix a point $y \in \mathbb{Z}^d$ and only work on the event $\{y \in \mathscr{C}_\infty\}$. The proof of this result relies on a two-scale expansion which takes the following form:

$$h(t, x, y) := \theta(\mathfrak{p})^{-1} \left(\bar{p}(t, x - y) + \sum_{k=1}^{d} \mathcal{D}_{e_k} \bar{p}(t, x - y) \phi_{e_k}(x) \right), \tag{4.107}$$

where the correctors ϕ_{e_k} are normalized around the point y, following the procedure described after Proposition 4.2.6.

As is common for two-scale expansions, the proof relies on two important ingredients:

- The sublinearity of the corrector, stated in Proposition 4.2.6;
- The sublinearity of the flux stated in Proposition 4.2.7.

The analysis also requires the estimates on the parabolic Green's function and its gradient provided by Theorem 4.3.1 and Theorem 4.1.3. We now present a sketch of the proof of Theorem 4.1.1. The result can be rewritten by using the Φ_C notation: for each exponent $\delta > 0$ and each $x \in \mathscr{C}_\infty$,

$$|p(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y)| \leqslant t^{-\frac{1}{2}+\delta}\Phi_C(t,x-y), \qquad \sqrt{t} \geqslant \mathcal{T}_{\mathrm{par},\delta}(y). \tag{4.108}$$

To prove this result, we first prove a weighted L^2-estimate (see (4.53) for the definition of Ψ_C)

$$\left\| \big(p(t,\cdot,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,\cdot - y)\big) \exp\left(\Psi_C(t,|\cdot - y|)\right) \right\|_{L^2(\mathscr{C}_\infty)} \leqslant Ct^{-\frac{d}{4}-\frac{1}{2}+\delta}, \tag{4.109}$$

and deduce the estimate (4.108) thanks to the semigroup property, this is proved in Section 4.4.3. Proving the estimate (4.109) is the core of the proof. To this end, we need to introduce a mesoscopic time $1 \ll \tau \ll t$ and a number of intermediate functions which are listed below:

- The two-scale expansion h defined in (4.107);
- The function $q := q(\cdot,\cdot,\tau,y)$ introduced in (4.110);
- The function $v := v(\cdot,\cdot,\tau,y)$ introduced in (4.124);
- The function w defined by the formula $w := h - v - q$.

The idea is to use these functions to split the difference

$$\begin{aligned} p(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y) = {} & (p(t,x,y) - q(t,x,\tau,y)) - v(t,x,\tau,y) \\ & - w(t,x,\tau,y) + \big(h(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y))\big), \end{aligned}$$

and then to prove that the L^2-norm of each of the terms is smaller than $t^{-\frac{d}{4}-\frac{1}{2}+\delta}$. More specifically, we organize the proof as follows:

(i) in Lemma 4.4.1, we prove that the term corresponding to the difference $(p-q)$ is small;

(ii) in Lemma 4.4.2, we prove that the term corresponding to the function v is small;

(iii) in Proposition 4.4.1, we prove that the term corresponding to the function w is small;

(iv) the term $\big(h(t,x,y) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-y))\big)$ is proved to be small by using the sublinearity of the corrector, the proof is straightforward and not stated in a specific lemma.

The rest of this section is organized as follows. Section 4.4.1 is devoted to the proof of Lemmas 4.4.1 and 4.4.2. Section 4.4.2 is devoted to the proof of Proposition 4.4.1 and is the core of the analysis: we make use of the regularization Lemmas proved in Section 4.4.1 as well as the various results recorded in the previous sections to perform the two-scale expansion. In Section 4.4.3, we post-process the results and prove the quantitative convergence of the heat kernel, Theorem 4.1.1.

4.4.1 Two regularization steps

We now introduce the function q. For a fixed initial time $\tau > 0$ and a vertex $y \in \mathscr{C}_\infty$, we let $(t,x) \mapsto q(t,x,\tau,y)$ be the solution of the parabolic problem

$$\begin{cases} \partial_t q - \nabla \cdot (\mathbf{a} \nabla q) = 0 & \text{in } (\tau, \infty) \times \mathscr{C}_\infty, \\ q(\tau, \cdot, \tau, y) = \theta(\mathfrak{p})^{-1} \bar{p}(\tau, \cdot - y) & \text{on } \mathscr{C}_\infty. \end{cases} \tag{4.110}$$

A reason justifying this construction is that the initial condition of the heat kernel p, which is a Dirac at y, is too singular and one cannot perform the two-scale expansion due to this lack of regularity. The idea is thus to replace the Dirac by a smoother function, the function $\bar{p}(\tau,\cdot)$, and to exploit its more favorable regularity properties to perform the two-scale expansion (see Section 4.4.2). For this strategy to work, one needs to choose the value of the time τ to be both:

- Large enough so that so that the function $\bar{p}(\tau,\cdot)$ is regular enough: in particular we want $\tau \gg 1$;
- Small enough so that the function q is a good approximation of the heat kernel p: we want $\tau \ll t$.

The choice of τ will be $\tau := t^{1-\kappa}$, for some small exponent κ whose value is decided at the end of the proof.

The following lemma states that the function q is a good approximation of the heat kernel p, when the coefficient τ is chosen such that $1 \ll \tau \ll t$. In the following lemma, given an exponent $\alpha > 0$, we use the notation $\mathcal{T}_{\mathrm{dense},\alpha}(y)$ to denote the minimal time introduced in Proposition 4.A.3, above which the mass of the homogenized heat kernel $\bar{p}$ is almost equal to the density of the infinite cluster $\mathscr{C}_\infty$ up to an error of order $t^{-\frac{1}{2}+\alpha}$.

Lemma 4.4.1. *For each exponent $\alpha > 0$ and each vertex $y \in \mathbb{Z}^d$, we let $\mathcal{T}_{\mathrm{approx},\alpha}(y)$ be the minimal time defined by the formula*

$$\mathcal{T}_{\mathrm{approx},\alpha}(y) := \max\left(\mathcal{M}_{\mathrm{reg}}(y)^2, \mathcal{T}_{\mathrm{dense},\alpha}(y)\right).$$

This random variable satisfies the stochastic integrability estimate

$$\mathcal{T}_{\mathrm{approx},\alpha}(y) \leqslant \mathcal{O}_s(C),$$

and the following property: there exists a positive constant $C := C(d,\mathfrak{p},\lambda,\alpha) < \infty$ such that, on the event $\{y \in \mathscr{C}_\infty\}$, for every pair of times $t,\tau \in (0,\infty)$ such that $t \geqslant 3\tau$ and $\tau \geqslant \mathcal{T}_{\mathrm{approx},\alpha}(y)$, and for every $x \in \mathscr{C}_\infty$, one has

$$|q(t,x,\tau,y) - p(t,x,y)| \leqslant \left(\left(\frac{\tau}{t}\right)^{\frac{1}{2}} + \tau^{-\frac{1}{2}+\alpha}\right) \Phi_C(t, x-y). \tag{4.111}$$

Proof. Before starting the proof, we note that the assumptions of the lemma imply the following results:

- The two inequalities $t \geqslant 3\tau$ and $\tau \geqslant \mathcal{T}_{\mathrm{approx},\alpha}(y)$ imply the estimates $t \geqslant \mathcal{T}_{\mathrm{approx},\alpha}(y)$ and $t - \tau \geqslant \mathcal{T}_{\mathrm{approx},\alpha}(y)$;
- By definition, the minimal time $\mathcal{T}_{\mathrm{approx},\alpha}(y)$ is larger than the minimal times $\mathcal{T}_{\mathrm{dense},\alpha}(y)$, $\mathcal{T}'_{NA}(y)$ and the square of the minimal scales $\mathcal{M}_{\mathrm{Poinc}}(y)$ and $\mathcal{M}_{\mathrm{reg}}(y)$. We can thus apply the corresponding results in the proof.

Step 1: Set up. We fix a vertex $y \in \mathbb{Z}^d$ and work on the event $\{y \in \mathscr{C}_\infty\}$. We first record the following estimate: under the assumption $\mathcal{T}_{\mathrm{dense},\alpha}(y)$, for each radius $r \geqslant \sqrt{\tau}$, we can compare the volume of the ball $B_r(y)$ and the cardinality of the set $\mathscr{C}_\infty \cap B_r(y)$, and we have the estimate

$$cr^d \leqslant |\mathscr{C}_\infty \cap B_r(y)| \leqslant Cr^d.$$

We consider a point x in the infinite cluster $\mathscr{C}_\infty$ and two times $t, \tau > 0$ such that $t \geqslant \tau \geqslant \mathcal{T}_{\mathrm{approx},\alpha}(y)$. By Duhamel's principle, one has

$$q(t,x,\tau,y) - p(t,x,y) = \int_{\mathscr{C}_\infty} \left(\theta(\mathfrak{p})^{-1}\bar{p}(\tau, z-y) - p(\tau,z,y)\right) p(t-\tau,x,z)\,dz. \tag{4.112}$$

Using the inequality $\tau \geqslant \mathcal{T}_{\mathrm{dense},\alpha}(y)$ and Proposition 4.A.3, we have the inequality

$$\left|\theta(\mathfrak{p})^{-1}\int_{\mathscr{C}_\infty} \bar{p}(\tau, z-y)\;dz - 1\right| \leqslant C\tau^{-\frac{1}{2}+\alpha}.$$

Since the mass of the transition kernel $p(\tau,\cdot,y)$ on the infinite cluster is equal to 1, we deduce that

$$\left|\int_{\mathscr{C}_\infty} \left(\theta(\mathfrak{p})^{-1}\bar{p}(\tau, z-y) - p(\tau,z,y)\right)\,dz\right| \leqslant C\tau^{-\frac{1}{2}+\alpha}.$$

We can thus subtract a constant term equal to $(p(\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)}$ in the right side of (4.112) up to a small cost of order $\tau^{-\frac{1}{2}+\alpha}$,

$$\begin{aligned}
&|q(t,x,\tau,y) - p(t,x,y)| \\
&\quad \leqslant \left|\int_{\mathscr{C}_\infty} \left(\theta(\mathfrak{p})^{-1}\bar{p}(\tau, z-y) - p(\tau,z,y)\right)\left(p(t-\tau,x,z) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)}\right)\,dz\right| \\
&\qquad + C\tau^{-\frac{1}{2}+\alpha}\left|(p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)}\right|.
\end{aligned}$$

Using the inequality $t-\tau \geqslant \mathcal{T}'_{\mathrm{NA}}(y)$ and (4.102), we apply Theorem 4.3.1 to obtain

$$\begin{aligned}
&|q(t,x,\tau,y) - p(t,x,y)| \\
&\leqslant \int_{\mathscr{C}_\infty} \Phi_C(\tau, z-y)\left|p(t-\tau,x,z) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)}\right|\,dz + \tau^{-\frac{1}{2}+\alpha}\Phi_C(t, x-y).
\end{aligned} \tag{4.113}$$

We then treat the first term on the right side of (4.113). The strategy is to split the integral into scales: for each integer $n \geqslant 1$, we let A_n be the dyadic annulus

$$A_n := \left\{z \in \mathbb{Z}^d \;:\; 2^n\sqrt{\tau} \leqslant |z-y| < 2^{n+1}\sqrt{\tau}\right\}$$

and then compute

$$\begin{aligned}
&\int_{\mathscr{C}_\infty} \Phi_C(\tau, z-y)\left|p(t-\tau,x,z) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)}\right|\,dz \\
&= \int_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \Phi_C(\tau, z-y)\left|p(t-\tau,x,z) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)}\right|\,dz \\
&\quad + \sum_{n=0}^{\infty}\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z-y)\left|p(t-\tau,x,z) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)}\right|\,dz.
\end{aligned} \tag{4.114}$$

Step 2: Multiscale analysis in the ball $B_{\sqrt{\tau}}(y)$. The term pertaining to small scales $B_{\sqrt{\tau}}(y)$ can be estimated thanks to the estimate $\Phi_C(\tau, z-y) \leqslant C\tau^{-d/2}$ and the Poincaré inequality. This latter inequality can be applied since we assumed $\tau \geqslant \mathcal{M}_{\mathrm{Poinc}}(y)^2$. This gives

$$\begin{aligned}\int_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} & \Phi_C(\tau, z-y) \left| p(t-\tau, x, z) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \right| dz \\ & \leqslant C \left\| p(t-\tau,x,\cdot) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \right\|_{\underline{L}^2\left(\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)\right)} \\ & \leqslant C\sqrt{\tau} \left\| \nabla_y p(t-\tau,x,\cdot) \right\|_{\underline{L}^2\left(\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)\right)},\end{aligned} \tag{4.115}$$

where the notation ∇_y means that the gradient is on the second spatial variable. We now estimate the term on the right side thanks to Theorem 4.1.3, or more precisely Remark 4.1.2, which can be applied since we assumed $t - \tau \geqslant \mathcal{M}_{\mathrm{reg}}(y)^2$. We have

$$\begin{aligned}\left\| \nabla p(t-\tau,x,\cdot) \right\|_{\underline{L}^2\left(\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)\right)} & \leqslant C(t-\tau)^{-1/2} \Phi_C\left(t-\tau, x-y\right) \\ & \leqslant C t^{-1/2} \Phi_C\left(t, x-y\right),\end{aligned}$$

where to go from the first line to the second one we used that $t - \tau \geqslant \frac{2}{3} t$ and increased the value of the constant C.

A combination of the two previous displays shows

$$\int_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \Phi_C(\tau, z-y) \left| p(t-\tau, x, z) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \right| dz \leqslant C \left(\frac{\tau}{t}\right)^{1/2} \Phi_C\left(t, x-y\right). \tag{4.116}$$

This completes the estimate of the term corresponding to the small scales in (4.114).

Step 3: Multiscale analysis in the annuli A_n. To estimate the terms corresponding to the dyadic annuli, we fix some integer $n \in \mathbb{N}$ and study the integral in the region A_n; thanks to the triangle inequality, we insert a constant term equal to $(p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)}$ in the integral,

$$\begin{aligned}\int_{\mathscr{C}_\infty \cap A_n} & \Phi_C(\tau, z-y) \left| p(t-\tau, x, z) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \right| dz \\ & \leqslant \underbrace{\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z-y) \left| p(t-\tau, x, z) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \right| dz}_{(4.117)-a} \\ & \quad + \underbrace{\frac{\left(\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z-y)\, dz\right)}{\left|\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)\right|} \int_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \left| p(t-\tau, x, z) - (p(t-\tau,x,\cdot))_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \right| dz}_{(4.117)-b}.\end{aligned} \tag{4.117}$$

Step 3.1: Estimate for the term (4.117)-a in the annuli A_n. We estimate (4.117)-a and distinguish three types of scales:

(i) The small scales which are defined as the annuli A_n such that $2^{n+2}\sqrt{\tau} \leqslant \sqrt{t}$;

(ii) The intermediate scales which are defined as the annuli A_n such that $t \geqslant 2^{n+2}\sqrt{\tau} > \sqrt{t}$;

(iii) The large scales which are defined as the annuli A_n such that $2^{n+2}\sqrt{\tau} > t$.

The following estimate for the function Φ_C is easy to check by its definition (4.51) and is used several times in the proof: for any constant $C' > C$, there exists a constant $c > 0$ depending only on the values of C and C' such that

$$\left(\sup_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) \leqslant e^{-c2^n\sqrt{\tau}} \left(\inf_{\mathscr{C}_\infty \cap A_n} \Phi_{C'}(\tau, \cdot - y) \right).$$

This implies that for any positive integer $k \in \mathbb{N}$, there exists a constant C_k depending only on the integer k and the constant C such that

$$\left(\sup_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) \leqslant \left(2^n\sqrt{\tau}\right)^{-k} \left(\inf_{\mathscr{C}_\infty \cap A_n} \Phi_{C_k}(\tau, \cdot - y) \right). \tag{4.118}$$

Step 3.1.1: Estimate for (4.117)*-a in the small scales* $2^{n+2}\sqrt{\tau} \leqslant \sqrt{t}$. We first focus on the small scales and apply the Poincaré inequality. With a computation similar to (4.115), one can estimate the first term in the right side of (4.117)

$$\begin{aligned}
&\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z - y) \left| p(t-\tau, x, z) - (p(t-\tau, x, \cdot))_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}} \right| dz \\
&\quad \leqslant \left(\sup_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) \int_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \left| p(t-\tau, x, z) - (p(t-\tau, x, \cdot))_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \right| dz \\
&\quad \leqslant \left(\sup_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) C (2^{n+1}\sqrt{\tau})^{\frac{d}{2}+1} \left\| \nabla_y p(t-\tau, x, \cdot) \right\|_{L^2\left(\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)\right)}.
\end{aligned}$$

Using the assumption $t - \tau \geqslant \mathcal{M}^2_{\mathrm{reg}}(y)$, we apply Theorem 4.1.3. This shows

$$\left\| \nabla_y p(t-\tau, x, \cdot) \right\|_{L^2\left(\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)\right)} \leqslant C t^{-1/2} \left(2^{n+1}\sqrt{\tau}\right)^{\frac{d}{2}} \Phi_C(t, x - y).$$

Using the explicit formula for the function Φ_C stated in (4.51), one has the estimate

$$\left(\sup_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) (2^{n+1}\sqrt{\tau})^d \leqslant C 2^{-2n}.$$

Combining the three previous displays shows, for each integer $n \in \mathbb{N}$ such that $2^{n+2}\sqrt{\tau} \leqslant \sqrt{t}$,

$$\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z - y) \left| p(t-\tau, x, z) - (p(t-\tau, x, \cdot))_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}} \right| dz \leqslant C 2^{-n} \left(\frac{\tau}{t} \right)^{\frac{1}{2}} \Phi_C(t, x - y). \tag{4.119}$$

Step 3.1.2: Estimate for (4.117)*-a in the intermediate scales* $t \geqslant 2^{n+2}\sqrt{\tau} > \sqrt{t}$. We now treat the case of the intermediate scales. In this case, we have $2^{n+2} \geqslant \sqrt{\frac{t}{\tau}}$, thus by the estimate (4.118) for $k = 2$, one has

$$\left(\sup_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \right) \leqslant 2^{-n} \left(\frac{\tau}{t} \right)^{\frac{1}{2}} \left(\inf_{\mathscr{C}_\infty \cap A_n} \Phi_{C_2}(\tau, \cdot - y) \right). \tag{4.120}$$

Using the assumptions $t - \tau \geqslant \mathcal{T}'_{\mathrm{NA}}(y)$, we can apply Theorem 4.3.1, the previous estimate (4.120) and the convolution property (4.52) for the map Φ_C. We obtain

$$\begin{aligned}
&\left(\sup_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y)\right) \int_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \left| p(t-\tau, x, z) - (p(t-\tau, x, \cdot))_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \right| \, dz \qquad (4.121)\\
\leqslant & 2^{-n} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} \left(\inf_{\mathscr{C}_\infty \cap A_n} \Phi_{C_2}(\tau, \cdot - y)\right) \int_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \Phi_C\left(t-\tau, x-z\right) \, dz \\
\leqslant & 2^{-n} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} \int_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \Phi_{C_2}(\tau, y-z) \Phi_C\left(t-\tau, x-z\right) \, dz \\
\leqslant & 2^{-n} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} \Phi_C\left(t, x-y\right),
\end{aligned}$$

by increasing the value of the constant C and using (4.52) in the last line.

Step 3.1.3: Estimate for the term (4.117)-a in the large scales $2^{n+2}\sqrt{\tau} > t$. The computation is similar to the one performed in (4.121) up to two differences listed below:

(i) For these scales, we cannot apply the Gaussian bounds on the heat kernel given by Theorem 4.3.1. Instead, we thus apply the Carne-Varopoulos bound which is stated in Proposition 4.2.8 and can be rewritten with the notation Φ_C: for each $x, z \in \mathscr{C}_\infty$,

$$p(t, x, z) \leqslant t^{\frac{d}{2}} \Phi_C\left(t, x-z\right);$$

(ii) We use the inequality $2^{n+2}\sqrt{\tau} > t$ and the estimate (4.118) for $k = \frac{d}{2} + 2$ to obtain the bound in the annulus A_n

$$\left(\sup_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y)\right) \leqslant 2^{-n} t^{-\frac{d}{2}} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} \left(\inf_{\mathscr{C}_\infty \cap A_n} \Phi_{C'}(\tau, \cdot - y)\right),$$

for some constant $C' > C$.

We can then perform the computation (4.121) and obtain the estimate

$$\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z-y) \left| p(t-\tau, x, z) - (p(t-\tau, x, \cdot))_{B_{2^{n+1}\sqrt{\tau}}(y)} \right| \, dz \leqslant C 2^{-n} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} \Phi_C(t, x-y). \tag{4.122}$$

Combining the estimates (4.119), (4.121) and (4.122), we have obtained, for each integer $n \in \mathbb{N}$,

$$\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z-y) \left| p(t-\tau, x, z) - (p(t-\tau, x, \cdot))_{B_{2^{n+1}\sqrt{\tau}}(y)} \right| \, dz \leqslant C 2^{-n} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} \Phi_C(t, x-y). \tag{4.123}$$

Step 3.2: Estimate for (4.117)-b in the annuli A_n. The second term (4.117)-b can be estimated thanks to a similar strategy: we first apply the inequality (4.118) with $k = d$

$$\begin{aligned}
\frac{\left(\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z-y) \, dz\right)}{\left|\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)\right|} &\leqslant \frac{\left|\mathscr{C}_\infty \cap A_n\right|}{\left|\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)\right|} \sup_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \\
&\leqslant 2^{dn} \sup_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, \cdot - y) \\
&\leqslant \inf_{\mathscr{C}_\infty \cap A_n} \Phi_{C'}(\tau, \cdot - y),
\end{aligned}$$

for some constant $C' > C$. Using this estimate, we deduce

$$\frac{\left(\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z-y)\, dz\right)}{\left|\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)\right|} \int_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \left| p(t-\tau, x, z) - (p(t-\tau, x, \cdot))_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \right| dz$$
$$\leqslant \inf_{\mathscr{C}_\infty \cap A_n} \Phi_{C'}(\tau, \cdot - y) \int_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \left| p(t-\tau, x, z) - (p(t-\tau, x, \cdot))_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \right| dz.$$

We can then apply the same proof as for the first term in the right side of (4.117). This proves the inequality

$$\frac{\left(\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z-y)\, dz\right)}{\left|\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)\right|} \int_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \left| p(t-\tau, x, z) - (p(t-\tau, x, \cdot))_{\mathscr{C}_\infty \cap B_{2^{n+1}\sqrt{\tau}}(y)} \right| dz$$
$$\leqslant C 2^{-n} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} \Phi_C(t, x-y).$$

Combining the previous inequality with the estimate (4.123), we obtain, for each integer $n \in \mathbb{N}$,

$$\int_{\mathscr{C}_\infty \cap A_n} \Phi_C(\tau, z-y) \left| p(t-\tau, x, z) - (p(t-\tau, x, \cdot))_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \right| dz \leqslant C 2^{-n} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} \Phi_C(t, x-y).$$

Combining this inequality with (4.113), (4.114), (4.116), and summing over the integer $n \in \mathbb{N}$ completes the proof of Lemma 4.4.1. □

We now introduce a second intermediate function useful in the proof of Theorem 4.1.1, the function v which is defined as follows. For some fixed $(\tau, y) \in (0, \infty) \times \mathscr{C}_\infty$, we let $(t, x) \mapsto v(t, x, \tau, y)$ be the solution of the parabolic equation

$$\begin{cases} \partial_t v(\cdot, \cdot, \tau, y) - \nabla \cdot \mathbf{a} \nabla v(\cdot, \cdot, \tau, y) = 0 & \text{in } (\tau, \infty) \times \mathscr{C}_\infty, \\ v(\tau, \cdot, \tau, y) = h(\tau, \cdot, y) - \theta(\mathfrak{p})^{-1} \bar{p}(\tau, \cdot - y) & \text{in } \mathscr{C}_\infty. \end{cases} \tag{4.124}$$

To define this function we run the parabolic equation starting from time τ and until time t with the initial condition given by the difference between the two-scale expansion h defined in (4.107) and the homogenized heat kernel $\theta(\mathfrak{p})^{-1} \bar{p}(\tau, \cdot - y)$. By the sublinearity of the corrector stated in Proposition 4.2.6, we expect the function $h(\tau, \cdot, y) - \theta(\mathfrak{p})^{-1} \bar{p}(\tau, \cdot - y)$ to be small. The following proposition states that the solution of the parabolic equation with this initial condition remains small (in the sense of the inequality (4.125)).

Lemma 4.4.2. *For any exponent $\alpha > 0$, there exists a positive constant $C(d, \lambda, \mathfrak{p}, \alpha) < \infty$ such that for each pair of times $t, \tau \in (0, \infty)$ satisfying $t \geqslant 3\tau$, $(t-\tau) \geqslant \mathcal{T}'_{\mathrm{NA}}(y)$, and $\sqrt{\tau} \geqslant \mathcal{M}_{\mathrm{corr}, \alpha}(y)$, the following estimate holds*

$$|v(t, x, \tau, y)| \leqslant C \tau^{-\frac{1}{2} + \frac{\alpha}{2}} \Phi_C(t, x-y). \tag{4.125}$$

Proof. The proof relies on two main ingredients: the quantitative sublinearity of the corrector and an explicit formula for the function v in terms of the heat kernel p.

First, by the definition (4.124), we have the formula

$$\begin{aligned} v(t, x, \tau, y) &= \int_{\mathscr{C}_\infty} \left(h(\tau, z, y) - \theta(\mathfrak{p})^{-1} \bar{p}(\tau, z-y) \right) p(t-\tau, x, z)\, dz \\ &= \int_{\mathscr{C}_\infty} \theta(\mathfrak{p})^{-1} \left(\sum_{k=1}^d \mathcal{D}_{e_k} \bar{p}(\tau, z-y) \phi_{e_k}(z) \right) p(t-\tau, x, z)\, dz. \end{aligned}$$

We then apply the four following estimates:

(i) The sublinearity of the corrector: under the assumption $\sqrt{\tau} \geqslant \mathcal{M}_{\mathrm{corr},\alpha}(y)$ and the normalization convention chosen for the corrector in the definition of the two-scale expansion h stated in (4.107), one has, for each $k \in \{1, \ldots, d\}$,

$$|\phi_{e_k}(z)| \leqslant \begin{cases} \tau^{\frac{\alpha}{2}} & \text{if } |z-y| \leqslant \tau^{\frac{1}{2}}, \\ |z-y|^{\alpha} & \text{if } |z-y| \geqslant \tau^{\frac{1}{2}}; \end{cases}$$

(ii) The Gaussian bounds on the transition kernel $p(t-\tau, x, \cdot)$, valid under the assumptions $\tau \geqslant \mathcal{T}'_{\mathrm{NA}}(y)$ and $t \geqslant 3\tau$: for each $z \in \mathscr{C}_\infty \cap B_{t-\tau}(y)$,

$$p(t-\tau, x, z) \leqslant \Phi_C\left(t-\tau, x-z\right) \leqslant \Phi_{C'}\left(t, x-z\right),$$

where the second inequality follows from the inequality $t - \tau \geqslant \frac{2}{3}t$ (by increasing the value of the constant C);

(iii) The bound on the transition kernel $p\left(t-\tau, x, z\right)$: for any point $z \in \mathscr{C}_\infty \setminus B_{t-\tau}(y)$,

$$p(t-\tau, x, z) \leqslant (t-\tau)^{d/2} \Phi_C\left(t-\tau, z-x\right),$$

which is a consequence of the definition of the map Φ_C stated in (4.51), Proposition 4.2.8 and the assumption $t \geqslant 3\tau$ (by increasing the value of the constant C);

(iv) The estimate on the homogenized heat kernel $\bar{p}$, which follows from standard results from the regularity theory,

$$|\mathcal{D}_{e_k}\bar{p}(\tau, z-y)| \leqslant C\tau^{-\frac{1}{2}} \Phi_C(\tau, z-y).$$

We obtain the inequality

$$\begin{aligned} |v(t,x,\tau,y)| \leqslant C\tau^{-\frac{1}{2}+\frac{\alpha}{2}} & \int_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \Phi_C\left(\tau, z-y\right) \Phi_C\left(t-\tau, x-z\right) dz \\ & + C\tau^{-\frac{1}{2}} \int_{\mathscr{C}_\infty \cap \left(B_{t-\tau}(y) \setminus B_{\sqrt{\tau}}(y)\right)} |z-y|^{\alpha} \Phi_C\left(\tau, z-y\right) \Phi_C\left(t-\tau, x-z\right) dz \\ & + C\tau^{-\frac{1}{2}}(t-\tau)^{d/2} \int_{\mathscr{C}_\infty \setminus B_{t-\tau}(y)} |z-y|^{\alpha} \Phi_C\left(\tau, z-y\right) \Phi_C\left(t-\tau, x-z\right) dz. \end{aligned} \tag{4.126}$$

We then estimate the three terms in the right side of (4.126) separately. For the first term, we use the inequality (4.52) and obtain, for some constant $C' > C$,

$$\begin{aligned} \tau^{-\frac{1}{2}+\frac{\alpha}{2}} \int_{\mathscr{C}_\infty \cap B_{\sqrt{\tau}}(y)} \Phi_C\left(\tau, z-y\right) \Phi_C\left(t-\tau, x-z\right) dz & \leqslant \tau^{-\frac{1}{2}+\frac{\alpha}{2}} \int_{\mathbb{Z}^d} \Phi_C\left(\tau, z-y\right) \Phi_C\left(t-\tau, x-z\right) dz \\ & \leqslant \tau^{-\frac{1}{2}+\frac{\alpha}{2}} \Phi_{C'}\left(t, x-y\right). \end{aligned} \tag{4.127}$$

To estimate the second term, we use the estimate (4.118) and deduce that

$$|z-y|^{\alpha} \Phi_C\left(\tau, z-y\right) \leqslant \tau^{\frac{\alpha}{2}} \Phi_{C'}\left(\tau, z-y\right), \tag{4.128}$$

for some constant $C' > C$. To estimate the third term in the right side of (4.126), we use the estimate (4.118) again with the value $k = d/2 + \alpha$ (the estimate applies even though k is not an integer). We obtain that for some constant $C' > C$ and for any point $z \in \mathscr{C}_\infty \setminus B_{t-\tau}(y)$,

$$(t-\tau)^{d/2} |z-y|^{\alpha} \Phi_C\left(\tau, z-y\right) \leqslant |z-y|^{d/2+\alpha} \Phi_C\left(\tau, z-y\right) \leqslant \Phi_{C'}\left(\tau, z-y\right). \tag{4.129}$$

Finally combining the identity (4.126) and the estimates (4.127), (4.128), (4.129), we obtain

$$|v(t,x,\tau,y)| \leqslant C\tau^{-\frac{1}{2}+\frac{\alpha}{2}} \Phi_{C'}\left(t, x-y\right).$$

The proof of Lemma 4.4.2 is complete. □

4.4.2 The two-scale expansion

The main objective of this section is to prove that the weighted L^2 norm of the function w is small in the sense of (4.130). Before starting the proof, we recall the notation for the function Ψ_C introduced in (4.53). We also recall the notation convention for discrete and continuous derivatives:

- In the proof of Proposition 4.4.1, the functions $\bar{p}$, ψ and η are defined on $\mathbb{R}^d$ and valued in $\mathbb{R}$, for these functions, we use the symbols ∇ and Δ to denote respectively the continuous gradient and the continuous Laplacian. To refer to the discrete derivatives, we use the notations $\mathcal{D}, \mathcal{D}^*, \mathcal{D}^2, \mathcal{D}^3$ etc.
- All the other functions are defined on the discrete lattice $\mathbb{Z}^d$ or on the infinite cluster $\mathscr{C}_\infty$, for these functions, we use the notation ∇ to denote the discrete gradient defined on the edges and the notation $\mathcal{D}$ for the discrete derivative defined on the vertices, following the conventions of Section 4.1.6.

Proposition 4.4.1. *For every exponent $\alpha > 0$, there exists a positive constant $C(d, \mathfrak{p}, \lambda, \alpha) < \infty$ such that for every point $y \in \mathbb{Z}^d$ and every time $t \in (0,\infty)$ such that $\sqrt{t} \geqslant \max\left(\mathcal{M}_{\mathrm{corr},\alpha}(y), \mathcal{M}_{\mathrm{flux},\alpha}(y)\right)$, one has, on the event $\{y \in \mathscr{C}_\infty\}$,*

$$\left\| w(t,\cdot,\tau,y) \exp(\Psi_C(t, |\cdot - y|)) \right\|_{L^2(\mathscr{C}_\infty)} \leqslant C \left(\frac{t}{\tau}\right)^{\frac{1}{2}} \tau^{-\frac{d}{4} - \frac{1}{2} + \frac{\alpha}{2}}. \tag{4.130}$$

Proof. The key is to develop a differential inequality for the function w. The proof is decomposed into five steps and is organized as follows. In Step 1, we use the explicit formula for w and apply the parabolic operator $\partial_t - \nabla \cdot \mathbf{a} \nabla$ to the map w to obtain the formulas (4.131) and (4.132). In Step 2, we test the equation obtained in (4.131) with the function ψw, where ψ is a map which is either equal to the constant 1 or equal to the function $x \mapsto \exp(\Psi_C(t, |x-y|)$. In the three remaining steps, we treat the different terms obtained and complete the proof of the estimate (4.130).

Step 1 : Establishing the equation for w. We claim that the function w satisfies the equation

$$\begin{cases} \partial_t w(\cdot,\cdot,\tau,y) - \nabla \cdot \mathbf{a} \nabla w(\cdot,\cdot,\tau,y) = f(\cdot,\cdot,y) + \mathcal{D}^* \cdot F(\cdot,\cdot,y) + \xi(\cdot,\cdot,y) & \text{in } (\tau,\infty) \times \mathscr{C}_\infty, \\ w(\tau,\cdot,y) = 0 & \text{in } \mathscr{C}_\infty, \end{cases} \tag{4.131}$$

where the three functions $f : (0,\infty) \times \mathscr{C}_\infty \times \mathscr{C}_\infty \to \mathbb{R}$, $F : (0,\infty) \times \mathscr{C}_\infty \times \mathscr{C}_\infty \to \mathbb{R}^d$ and $\xi : (0,\infty) \times \mathscr{C}_\infty \times \mathscr{C}_\infty \to \mathbb{R}$ are defined by the formulas, for each $(t,y) \in (0,\infty) \times \mathscr{C}_\infty$,

$$\begin{cases} f(t,\cdot,y) = \dfrac{1}{2}\bar{\sigma}^2 \left(\Delta \bar{p}(t,\cdot - y) - (-\mathcal{D}^* \cdot \mathcal{D}\bar{p}(t,\cdot - y)) \right) + \displaystyle\sum_{k=1}^d \left(\partial_t \mathcal{D}_{e_k} \bar{p}(t,\cdot - y) \right) \phi_{e_k}(\cdot), \\ [F]_i(t,\cdot,y) = \displaystyle\sum_{k=1}^d [\mathbf{a}\mathcal{D}\mathcal{D}_{e_k}\bar{p}(t,\cdot - y)]_i T_{e_i}(\phi_{e_k})(\cdot), \qquad \forall i \in \{1,\ldots,d\}, \\ \xi(t,\cdot,y) = \displaystyle\sum_{k=1}^d \mathcal{D}^* \mathcal{D}_{e_k} \bar{p}(t,\cdot - y) \cdot \widetilde{\mathbf{g}}^*_{e_k}(\cdot), \end{cases} \tag{4.132}$$

where $\widetilde{\mathbf{g}}^*_{e_k}$ is a translated version of the flux $\widetilde{\mathbf{g}}_{e_k}$ defined by the formula, for each $x \in \mathscr{C}_\infty$,

$$\widetilde{\mathbf{g}}^*_{e_k}(x) := \begin{pmatrix} T_{-e_1}\left[\mathbf{a}\left(\mathcal{D}\phi_{e_k} + e_k\right) - \frac{1}{2}\bar{\sigma}^2 e_k \right]_1 \\ \vdots \\ T_{-e_d}\left[\mathbf{a}\left(\mathcal{D}\phi_{e_k} + e_k\right) - \frac{1}{2}\bar{\sigma}^2 e_k \right]_d \end{pmatrix},$$

and we recall the notation $\left[\mathbf{a}\left(\mathcal{D}\phi_{e_k}+e_k\right)-\frac{1}{2}\bar{\sigma}^2 e_k\right]_i$ introduced in Section 4.1.6 to denote the ith-component of the vector $\mathbf{a}\left(\mathcal{D}\phi_{e_k}+e_k\right)-\frac{1}{2}\bar{\sigma}^2 e_k$. In Appendix 4.B, it is proved that the translated flux $\widetilde{\mathbf{g}}^*_{e_k}$ has similar properties as the centered flux $\widetilde{\mathbf{g}}_{e_k}$. In particular, it is proved in Remark 4.B that, for every radius $r \geqslant \mathcal{M}_{\mathrm{flux},\alpha}(y)$,

$$\left\|\widetilde{\mathbf{g}}^*_{e_k}\right\|_{\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r(y))} \leqslant Cr^\alpha.$$

The proof follows from a direct calculation. Since one has the equality $(\partial_t - \nabla\cdot\mathbf{a}\nabla)(v+q)(\cdot,\cdot,s,y)=0$, it suffices to focus on the term h in the definition of w, the details are left to the reader.

Step 2 : A differential inequality. In this step and the rest of the proof, we let ψ be the function from $(0,\infty)\times\mathbb{R}^d$ to $\mathbb{R}$ which is either:

(i) The constant function equal to 1;

(ii) The function $\exp(\Psi_C(t,|\cdot-y|))$, for some large constant $C>0$.

We note that in both cases, the function ψ is smooth and satisfies the following property

$$\exists\widetilde{C}>1,\quad \forall|h|\leqslant 1,\quad |\nabla\psi(\cdot+h)|\leqslant\widetilde{C}|\nabla\psi|. \tag{4.133}$$

This estimate allows to replace the discrete derivative $\mathcal{D}\psi$, by the continuous derivative $\nabla\psi$ by only paying a constant. The strategy of the proof is then to test the equation (4.131) with the function $w\psi^2$ to derive a differential inequality. We first write

$$\underbrace{\int_{\mathscr{C}_\infty}\left(\left(\partial_t w-\nabla\cdot\mathbf{a}\nabla w\right)\psi^2 w\right)}_{\text{LHS}} = \underbrace{\int_{\mathscr{C}_\infty} fw\psi^2 + \int_{\mathscr{C}_\infty} F\cdot\mathcal{D}(w\psi^2) + \int_{\mathscr{C}_\infty}\xi w\psi^2}_{\text{RHS}}. \tag{4.134}$$

We then estimate the terms on the left and right-hand sides separately. Before starting the computation, we mention that in the following paragraphs all the derivatives and integrations are acting on the first spatial variable. For the left-hand side of (4.134), we have

$$\begin{aligned}\text{LHS} &= \int_{\mathscr{C}_\infty}\left(\left(\partial_t w-\nabla\cdot\mathbf{a}\nabla w\right)\psi^2 w\right)\\ &= \int_{\mathscr{C}_\infty}\frac{1}{2}\partial_t(\psi^2 w^2) - \int_{\mathscr{C}_\infty}(\partial_t\psi)\psi w^2 + \int_{\mathscr{C}_\infty}\nabla(\psi^2 w)\cdot\mathbf{a}\nabla w.\end{aligned}$$

Using (4.133) and Young's inequality, we have, for any $a\in(0,\infty)$,

$$\int_{\mathscr{C}_\infty}\nabla(\psi^2 w)\cdot\mathbf{a}\nabla w \geqslant \lambda\int_{\mathscr{C}_\infty}|\nabla w|^2\psi^2 - a\int_{\mathscr{C}_\infty}|\nabla w|^2\psi^2 - \frac{\widetilde{C}^2}{a}\int_{\mathscr{C}_\infty}|\nabla\psi|^2 w^2,$$

where, following the notation convention recalled at the beginning of this section, the notation $\nabla\psi$ denotes the continous gradient for functions ψ, while the notation $\nabla(\psi w)$ refers to the discrete gradient on the infinite cluster since the map w is only defined on $\mathscr{C}_\infty$. By choosing the value $a=\frac{\lambda}{2}$, the previous display can be rewritten

$$\text{LHS} \geqslant \int_{\mathscr{C}_\infty}\frac{1}{2}\partial_t(\psi^2 w^2) + \frac{\lambda}{2}\int_{\mathscr{C}_\infty}|\nabla w|^2\psi^2 - \int_{\mathscr{C}_\infty}(\partial_t\psi)\psi w^2 - \frac{2\widetilde{C}^2}{\lambda}\int_{\mathscr{C}_\infty}|\nabla\psi|^2 w^2. \tag{4.135}$$

We now focus on terms on the right-hand side of the equality (4.134). For the first two terms, we use Young's inequality and obtain

$$\begin{aligned}\int_{\mathscr{C}_\infty} fw\psi^2 + \int_{\mathscr{C}_\infty} F\cdot\mathcal{D}(w\psi^2) \leqslant \int_{\mathscr{C}_\infty} tf^2\psi^2 + \frac{1}{4t}\int_{\mathscr{C}_\infty} w^2\psi^2 + \frac{\widetilde{C}^2}{\lambda}\int_{\mathscr{C}_\infty}|F|^2\psi^2 + \frac{\lambda}{4}\int_{\mathscr{C}_\infty}|\nabla w|^2\psi^2\\ + \frac{\lambda}{2}\int_{\mathscr{C}_\infty}|F|^2\psi^2 + \frac{2\widetilde{C}^2}{\lambda}\int_{\mathscr{C}_\infty}|\nabla\psi|^2 w^2.\end{aligned}$$

A combination of the two previous displays and the identity (4.134) shows

$$\int_{\mathscr{C}_\infty} \left(\frac{1}{2}\partial_t(\psi^2 w^2) + \frac{\lambda}{4}|\nabla w|^2 \psi^2 \right) \leqslant \int_{\mathscr{C}_\infty} w^2 \left((\partial_t \psi)\psi + \frac{1}{4t}\psi^2 + \frac{4\widetilde{C}^2}{\lambda}|\nabla \psi|^2 \right) \\ +2\int_{\mathscr{C}_\infty} tf^2\psi^2 + \left(\frac{\widetilde{C}^2}{\lambda} + \frac{\lambda}{2} \right) \int_{\mathscr{C}_\infty} |F|^2\psi^2 + \int_{\mathscr{C}_\infty} \xi w \psi^2. \tag{4.136}$$

The value of the constants in the second line of the estimate (4.136) does not need to be tracked in the proof, we thus rewrite it in the following form

$$\int_{\mathscr{C}_\infty} \left(\frac{1}{2}\partial_t(\psi^2 w^2) + \frac{\lambda}{4}|\nabla w|^2 \psi^2 \right) \leqslant \int_{\mathscr{C}_\infty} w^2 \left((\partial_t \psi)\psi + \frac{1}{4t}\psi^2 + \frac{4\widetilde{C}^2}{\lambda}|\nabla \psi|^2 \right) \\ + C\left(\int_{\mathscr{C}_\infty} tf^2\psi^2 + |F|^2\psi^2 + \xi w \psi^2 \right). \tag{4.137}$$

To complete the proof, we need to prove that the quantities on the second line of the previous display are small:

- One needs to prove that the term $\int_{\mathscr{C}_\infty} \xi(t,\cdot,y)(\psi^2(t,\cdot,y)w(t,\cdot,s,y))$ is small, this is proved in Step 3;
- One needs to prove that the term $\int_{\mathscr{C}_\infty} \left(tf^2(t,\cdot,y) + |F|^2(t,\cdot,y)\right)\psi^2(t,\cdot,y)$ is small, this is proved in Step 4.

Step 3 : Estimate of the term $\int_{\mathscr{C}_\infty} \xi(t,\cdot,y)(\psi^2(t,\cdot,y)w(t,\cdot,s,y))$. The term ξ involves the centered flux $\widetilde{\mathbf{g}}^*_{e_k}$, to prove that this integral is small, the strategy is to use the weak norm estimate on this function stated in Proposition 4.2.7 and a multiscale argument. Specifically, the goal of this step is to prove the inequality

$$\int_{\mathscr{C}_\infty} \xi\psi^2 w \leqslant Ct^{-\frac{d}{4}-1+\frac{\alpha}{2}} \left\| \nabla(w\psi) \right\|_{L^2(\mathscr{C}_\infty)} + Ct^{-\frac{d}{4}-\frac{3}{2}+\frac{\alpha}{2}} \left\| w\psi \right\|_{L^2(\mathscr{C}_\infty)}. \tag{4.138}$$

As in Lemma 4.4.1, we need to split the space into scales and we define the dyadic annuli: for each integer $m \geqslant 1$, we let A_m be the annulus $A_m := \left\{ z \in \mathbb{Z}^d \ : \ 2^{m-1}\sqrt{t} \leqslant |z-y| < 3 \cdot 2^m \sqrt{t} \right\}$, we also let $A_0 := B_{\sqrt{t}}(y)$. We then split the proof into two steps:

(i) We first prove the estimate

$$\int_{\mathscr{C}_\infty} \xi\psi^2 w \leqslant C\Xi_1 \left\| \nabla(w\psi) \right\|_{L^2(\mathscr{C}_\infty)} + C\Xi_2 \left\| w\psi \right\|_{L^2(\mathscr{C}_\infty)}, \tag{4.139}$$

where the two quantities Ξ_1, Ξ_2 are defined by the formulas

$$\Xi_1 := \left(\sum_{k=1}^d \sum_{m=0}^\infty \left(2^m\sqrt{t}\right)^d \left\| \widetilde{\mathbf{g}}^*_{e_k} \right\|^2_{\underline{H}^{-1}(\mathscr{C}_\infty \cap A_m)} \left\| \mathcal{D}^2 \bar{p}(t, \cdot - y)\psi \right\|^2_{L^\infty(A_m)} \right)^{\frac{1}{2}},$$
$$\Xi_2 := \left(\sum_{k=1}^d \sum_{m=0}^\infty \left(2^m\sqrt{t}\right)^d \left\| \widetilde{\mathbf{g}}^*_{e_k} \right\|^2_{\underline{H}^{-1}(\mathscr{C}_\infty \cap A_m)} \right. \\ \left. \times \left(\left\| \mathcal{D}^3 \bar{p}(t, \cdot - y)\psi \right\|^2_{L^\infty(A_m)} + \left\| \mathcal{D}^2 \bar{p}(t, \cdot - y)\mathcal{D}\psi \right\|^2_{L^\infty(A_m)} + (3^m\sqrt{t})^{-2} \left\| \mathcal{D}^2 \bar{p}(t, \cdot - y)\psi \right\|^2_{L^\infty(A_m)} \right) \right)^{\frac{1}{2}}; \tag{4.140}$$

(ii) We then prove the estimates

$$\Xi_1 \leqslant Ct^{-\frac{d}{4}-1+\frac{\alpha}{2}} \text{ and } \Xi_2 \leqslant Ct^{-\frac{d}{4}-\frac{3}{2}+\frac{\alpha}{2}}. \tag{4.141}$$

The estimate (4.138) is a consequence of the inequalities (4.139) and (4.141).

We now focus on the proof of the inequality (4.139). The strategy is to use a multiscale analysis. We let η be a smooth cutoff function from $\mathbb{R}^d$ to $\mathbb{R}$ satisfying the properties

$$0 \leqslant \eta \leqslant 1, \qquad |\nabla \eta| \leqslant 1, \qquad \operatorname{supp}(\eta) \subseteq B_3(y), \qquad \eta \equiv 1 \text{ in } B_1(y). \tag{4.142}$$

For an integer m, we define the rescaled version η_m of η according to the formula $\eta_m := \eta\left(\frac{\cdot - y}{2^m\sqrt{t}} + y\right)$, we also set the convention $\eta_{-1,y} \equiv 0$. This function satisfies the property: for each $m \in \mathbb{N}$,

$$|\nabla \eta_m| \leqslant \left(2^m\sqrt{t}\right)^{-1}, \quad \operatorname{supp}(\eta_m) \subseteq B_{2^{m+1}\sqrt{t}}(y), \quad \eta_m \equiv 1 \text{ in } B_{2^m\sqrt{t}}(y), \quad \operatorname{supp}(\eta_m - \eta_{m-1}) \subseteq A_m.$$

We also note that the family of functions η_m can be used as a partition of unity and we have

$$1 = \sum_{m=0}^{\infty} (\eta_m - \eta_{m-1}).$$

With this property, we compute

$$\begin{aligned} \int_{\mathscr{C}_\infty} \xi \psi^2 w &= \sum_{m=0}^{\infty} \int_{\mathscr{C}_\infty} (\eta_m - \eta_{m-1}) \xi \psi^2 w \\ &= \sum_{k=1}^{d} \sum_{m=0}^{\infty} \int_{\mathscr{C}_\infty \cap B_{2^{m+1}\sqrt{t}}(y)} \widetilde{\mathbf{g}}^*_{e_k} \cdot \mathcal{D}^* \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) (\eta_m - \eta_{m-1}) \psi^2 w \\ &\leqslant \sum_{k=1}^{d} \sum_{m=0}^{\infty} (2^m\sqrt{t})^d \left\| \widetilde{\mathbf{g}}^*_{e_k} \right\|_{\underline{H}^{-1}(\mathscr{C}_\infty \cap A_m)} \left\| \mathcal{D}^* \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) (\eta_m - \eta_{m-1}) \psi^2 w \right\|_{\underline{H}^1(\mathscr{C}_\infty \cap A_m)}. \end{aligned} \tag{4.143}$$

Then we calculate the $\underline{H}^1$-norm of the term $\mathcal{D}^* \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) (\eta_m - \eta_{m-1}) \psi^2 w$. We use the fact that the function $(\eta_m - \eta_{m-1})$ is supported in the annulus A_m and write

$$\left\| \mathcal{D}^* \mathcal{D}_{e_k} \bar{p}(t, \cdot - y) (\eta_m - \eta_{m-1}) \psi^2 w \right\|_{\underline{H}^1(\mathscr{C}_\infty \cap B_{2^{m+1}\sqrt{t}}(y))} \leqslant C (I_1 + I_2),$$

where the two terms I_1 and I_2 are defined by the formulas

$$\begin{aligned} I_1 &:= \left\| \mathcal{D}^2 \bar{p}(t, \cdot - y) \psi \right\|_{L^\infty(A_m)} \left\| \nabla (w\psi) \right\|_{\underline{L}^2(\mathscr{C}_\infty \cap A_m)} \\ I_2 &:= \Big(\left\| \mathcal{D}^3 \bar{p}(t, \cdot - y) \psi \right\|_{L^\infty(A_m)} + \left\| \mathcal{D}^2 \bar{p}(t, \cdot - y) \mathcal{D}\psi \right\|_{L^\infty(A_m)} \\ &\quad + (3^m\sqrt{t})^{-1} \left\| \mathcal{D}^2 \bar{p}(t, \cdot - y) \psi \right\|_{L^\infty(A_m)} \Big) \left\| w\psi \right\|_{\underline{L}^2(\mathscr{C}_\infty \cap A_m)}. \end{aligned}$$

We put these equations back into the right-hand side of the estimate (4.143). This gives

$$\int_{\mathscr{C}_\infty} \xi \psi^2 w \leqslant C \sum_{k=1}^{d} \sum_{m=0}^{\infty} (2^m\sqrt{t})^{\frac{d}{2}} \left\| \widetilde{\mathbf{g}}^*_{e_k} \right\|_{\underline{H}^{-1}(\mathscr{C}_\infty \cap A_m)} (I_1 + I_2).$$

We then estimate the two terms on the right side by applying the Cauchy-Schwarz inequality. For the term involving the quantity I_1, we obtain

$$\begin{aligned}
&\sum_{k=1}^{d}\sum_{m=0}^{\infty}(2^m\sqrt{t})^{\frac{d}{2}}\left\|\widetilde{\mathbf{g}}_{e_k}^{*}\right\|_{\underline{H}^{-1}(\mathscr{C}_\infty\cap A_m)} I_1 \\
&\leqslant\left(\sum_{k=1}^{d}\sum_{m=0}^{\infty}(2^m\sqrt{t})^{d}\left\|\widetilde{\mathbf{g}}_{e_k}^{*}\right\|_{\underline{H}^{-1}(\mathscr{C}_\infty\cap B_{2^{m+1}\sqrt{t}}(y))}^2\left\|\mathcal{D}^2\bar{p}(t,\cdot-y)\psi\right\|_{L^\infty(A_m)}^2\right)^{\frac12} \\
&\qquad\times\left(\sum_{k=1}^{d}\sum_{m=0}^{\infty}\left\|\nabla(w\psi)\right\|_{L^2(\mathscr{C}_\infty\cap A_m)}^2\right)^{\frac12} \\
&\leqslant C\Xi_1\left\|\nabla(w\psi)\right\|_{L^2(\mathscr{C}_\infty)},
\end{aligned}$$

where to go from the second to the third line, we used the definition of Ξ_1 given in (4.140) and the inequality $\sum_{m=1}^{\infty}\mathbf{1}_{\{A_m(y)\}}\leqslant 4$. The same argument works for the terms involving the quantities I_2 and Ξ_2, this concludes the proof of the estimate (4.139).

We now prove an estimate on the terms Ξ_1,Ξ_2; precisely we prove the inequality (4.141) which is recalled below

$$\Xi_1\leqslant Ct^{-\frac{d}{4}-1+\frac{\alpha}{2}} \text{ and } \Xi_2\leqslant Ct^{-\frac{d}{4}-\frac{3}{2}+\frac{\alpha}{2}}. \tag{4.144}$$

The proof comes from a direct calculation of the quantities Ξ_1 and Ξ_2. We recall that the function ψ is chosen to be either the constant function equal to 1, or the function $\exp(\Psi_C(t,|\cdot-y|))$, for some large constant C.

We first focus on the estimate of the term Ξ_1; if the constant C in the definition of ψ is chosen large enough, for instance larger than $8\bar{\sigma}^2$, then one has the estimate

$$\left\|\mathcal{D}^2\bar{p}(t,\cdot-y)\psi\right\|_{L^\infty(A_m(y))}\leqslant Ct^{-\frac{d}{2}-1}\exp\left(-\frac{2^{2m}}{8\bar{\sigma}^2}\right).$$

Thanks to the assumption $\sqrt{t}>\mathcal{M}_{\mathrm{flux},\alpha}(y)$, we have the estimate

$$\left\|\widetilde{\mathbf{g}}_{e_k}^{*}\right\|_{\underline{H}^{-1}(\mathscr{C}_\infty\cap A_m)}\leqslant\left\|\widetilde{\mathbf{g}}_{e_k}^{*}\right\|_{\underline{H}^{-1}(\mathscr{C}_\infty\cap B_{2^{m+1}\sqrt{t}}(y))}\leqslant C\left(2^m\sqrt{t}\right)^{\alpha}.$$

Combining these two bounds with the definition of Ξ_1 given in (4.140), we obtain

$$(\Xi_1)^2\leqslant C\sum_{k=1}^{d}\sum_{m=0}^{\infty}\left(2^m\sqrt{t}\right)^{d+2\alpha}t^{-d-2}\exp\left(-\frac{2^{2m}}{4\bar{\sigma}^2}\right)\leqslant Ct^{-\frac{d}{2}-2+\alpha}.$$

The term Ξ_2 can be estimated thanks to a similar strategy and the details are left to the reader. The proof of the estimate (4.144), and thus of the inequality (4.138) is complete.

Step 4 : Quantification of the term $\int_{\mathscr{C}_\infty}\left(tf^2(t,\cdot,y)+|F|^2(t,\cdot,y)\right)\psi^2(t,\cdot,y)$. The goal of this step is to prove the inequality

$$\int_{\mathscr{C}_\infty}\left(tf^2(t,\cdot,y)+|F|^2(t,\cdot,y)\right)\psi^2(t,\cdot,y)\leqslant Ct^{-\frac{d}{2}-2+\alpha}. \tag{4.145}$$

As was the case in the previous step, the function ψ is either the constant function equal to 1 or the function $\exp(\Psi_C(t,|\cdot-y|))$. In the latter case, we assume that the constant C is at least larger than $8\bar{\sigma}^2$.

We first consider the term involving the function f. From the definition of this function given in (4.132), we see that it is the sum of two terms. The first one is the difference of the discrete and the continuous Laplacian of the heat kernel $\bar{p}$; it can be estimated as follows

$$|(\Delta\bar{p}(t,\cdot-y)-(-\mathcal{D}^*\cdot\mathcal{D}\bar{p}(t,\cdot-y)))\,\psi|\leqslant Ct^{-\frac{d}{2}-\frac{3}{2}}\exp\left(-\frac{|\cdot-y|^2}{4\bar{\sigma}^2t}\right).$$

The second term is the quantity $\sum_{k=1}^d \partial_t\mathcal{D}_{e_k}\bar{p}(t,\cdot-y)\phi_{e_k}(\cdot)$. To estimate it, we split the space into different scales using the functions η_m introduced in Step 3. This gives

$$\begin{aligned}
&\int_{\mathscr{C}_\infty} t\sum_{k=1}^d\left(\partial_t\left(\mathcal{D}_{e_k}\bar{p}(t,\cdot-y)\phi_{e_k}\right)\psi\right)^2\\
&=\sum_{k=1}^d\sum_{m=0}^\infty\int_{\mathscr{C}_\infty} t(\eta_m-\eta_{m-1})\left(\partial_t\mathcal{D}_{e_k}\bar{p}(t,\cdot-y)\psi\right)^2\phi_{e_k}^2\\
&\leqslant\sum_{k=1}^d\sum_{m=0}^\infty t\left(2^m\sqrt{t}\right)^d\|\phi_{e_k}\|^2_{\underline{L}^2(\mathscr{C}_\infty\cap B_{2^{m+1}\sqrt{t}}(y))}\|(\partial_t\mathcal{D}_{e_k}\bar{p}(t,\cdot-y)\psi)\|^2_{L^\infty(A_m)}.
\end{aligned}$$

We then use the assumption $\sqrt{t}>\mathcal{M}_{\mathrm{corr},\alpha}(y)$, which implies $\|\phi_{e_k}\|_{\underline{L}^2(\mathscr{C}_\infty\cap B_{2^{m+1}\sqrt{t}}(y))}\leqslant C\left(2^m\sqrt{t}\right)^\alpha$, and the estimate

$$\|(\partial_t\mathcal{D}_{e_k}\bar{p}(t,\cdot-y)\psi)\|_{L^\infty(A_m)}\leqslant Ct^{-\frac{d}{2}-\frac{3}{2}}\exp\left(-\frac{2^{2m}}{8\bar{\sigma}^2}\right),$$

to obtain the inequality

$$\begin{aligned}
\int_{\mathscr{C}_\infty} t\sum_{k=1}^d\left(\partial_t\left(\mathcal{D}_{e_k}\bar{p}(t,\cdot-y)\phi_{e_k}\right)\psi\right)^2&\leqslant C\sum_{k=1}^d\sum_{m=0}^\infty t\left(2^m\sqrt{t}\right)^{d+2\alpha}t^{-d-3}\exp\left(-\frac{2^{2m}}{4\bar{\sigma}^2}\right)\\
&\leqslant Ct^{-\frac{d}{2}-2+\alpha}.
\end{aligned}$$

The estimate for the term involving the function F is similar and we skip its proof.

Step 5 : The conclusion. We collect the results established in the previous steps and complete the proof of Proposition 4.4.1. We first consider the inequality (4.137) in the case $\psi=1$. This gives

$$\int_{\mathscr{C}_\infty}\left(\frac{1}{2}\partial_t w^2+\frac{\lambda}{4}|\nabla w|^2\right)\leqslant\frac{1}{4t}\int_{\mathscr{C}_\infty}w^2+C\left(\int_{\mathscr{C}_\infty}tf^2+|F|^2+\xi w\right),$$

since in this case the constant $\widetilde{C}$ introduced in (4.133) is equal to 1. Applying the main results (4.138) of Step 3 and (4.145) of Step 4, we deduce

$$\int_{\mathscr{C}_\infty}\left(\partial_t w^2+\frac{\lambda}{2}|\nabla w|^2\right)\leqslant\frac{1}{2t}\int_{\mathscr{C}_\infty}w^2+Ct^{-\frac{d}{2}-2+\alpha}+Ct^{-\frac{d}{4}-1+\frac{\alpha}{2}}\left(\|\nabla w\|_{L^2(\mathscr{C}_\infty)}+t^{-\frac{1}{2}}\|w\|_{L^2(\mathscr{C}_\infty)}\right).$$

By Young's inequality, the previous display can be simplified

$$\int_{\mathscr{C}_\infty}\left(\partial_t w^2+\frac{\lambda}{4}|\nabla w|^2\right)\leqslant\frac{1}{t}\int_{\mathscr{C}_\infty}w^2+Ct^{-\frac{d}{2}-2+\alpha},$$

which implies

$$\partial_t\int_{\mathscr{C}_\infty}w^2\leqslant\frac{1}{t}\int_{\mathscr{C}_\infty}w^2+Ct^{-\frac{d}{2}-2+\alpha}.$$

By integrating over the time interval $[\tau, t]$, we obtain that there exists a constant $C(d, \mathfrak{p}, \lambda) < \infty$ such that

$$\int_{\mathscr{C}_\infty} w^2(t, \cdot, s, y) \leqslant C\left(\frac{t}{\tau}\right)\tau^{-\frac{d}{2}-1+\alpha}. \tag{4.146}$$

We now consider the inequality (4.137) in the case $\psi = \exp(\Psi_C(t, \cdot - y))$. This gives

$$\int_{\mathscr{C}_\infty} \left(\frac{1}{2}\partial_t(\psi^2 w^2) + \frac{\lambda}{4}|\nabla w|^2\psi^2\right) \leqslant \int_{\mathscr{C}_\infty} w^2\left((\partial_t\psi)\psi + \frac{1}{8t}\psi^2 + \frac{4\widetilde{C}^2}{\lambda}|\nabla\psi|^2\right) \\ + Ct^{-\frac{d}{2}-2+\alpha} + Ct^{-\frac{d}{4}-1+\frac{\alpha}{2}}\left(\|\nabla(w\psi)\|_{L^2(\mathscr{C}_\infty)} + t^{-\frac{1}{2}}\|w\psi\|_{L^2(\mathscr{C}_\infty)}\right).$$

Applying Young's inequality, the previous display can be simplified and we obtain

$$\int_{\mathscr{C}_\infty} \left(\frac{1}{2}\partial_t(\psi^2 w^2) + \frac{\lambda}{8}|\nabla w|^2\psi^2\right) \leqslant \int_{\mathscr{C}_\infty} w^2\left((\partial_t\psi)\psi + \frac{1}{4t}\psi^2 + \frac{8\widetilde{C}}{\lambda}|\nabla\psi|^2\right) + Ct^{-\frac{d}{2}-2+\alpha}.$$

We then note that if the constant C in the definition of $\psi = \exp(\Psi_C(\cdot, \cdot - y))$ is chosen large enough, then we have

$$(\partial_t\psi)\psi + \frac{1}{4t}\psi^2 + \frac{8\widetilde{C}}{\lambda}|\nabla\psi|^2 \leqslant \frac{C}{t}. \tag{4.147}$$

A combination of the two previous displays shows the differential inequality

$$\partial_t \int_{\mathscr{C}_\infty} \frac{1}{2}\psi^2 w^2 \leqslant \frac{C}{t}\int_{\mathscr{C}_\infty} w^2 + Ct^{-\frac{d}{2}-2+\alpha}.$$

We then apply (4.146) to obtain

$$\partial_t \int_{\mathscr{C}_\infty} \frac{1}{2}\psi^2 w^2 \leqslant Ct^{-1}\left(\frac{t}{\tau}\right)\tau^{-\frac{d}{2}-1+\alpha} + Ct^{-\frac{d}{2}-2+\alpha}.$$

Integrating with respect to the time t and recalling that $w(\tau, \cdot, \tau, y) = 0$, we obtain

$$\int_{\mathscr{C}_\infty} (w\psi)^2 \leqslant C\left(\frac{t}{\tau}\right)\tau^{-\frac{d}{2}-1+\alpha},$$

for some constant $C := C(d, \mathfrak{p}, \lambda) < \infty$. This completes the proof of Proposition 4.4.1. □

Remark. The reason why we choose the function $\psi = \exp(\Psi_C(\cdot, \cdot - y))$, and why the main result (4.130) of Proposition 4.4.1 is stated with this function can be explained by the inequalities (4.133) and (4.147). Indeed, the function $(t, x) \mapsto \exp(\Psi_C(t, x-y))$ is the one which has the fastest growth as x tends to infinity such that the inequalities (4.133) and (4.147) are satisfied. In particular there is an important difference between the discrete setting and the continuous setting: in the latter, one does not need the inequality (4.133) to hold which allows to choose the function $\psi(t, x) := \exp\left(\frac{|x-y|^2}{Ct}\right)$, and to obtain the result with this function (see [25, Lemma 8.22]). This observation is consistent with the asymptotic behavior of the discrete heat kernel on the percolation cluster or on $\mathbb{Z}^d$ which is described by Proposition 4.2.8.

We have now collected all the necessary results to prove Theorem 4.1.1. The following section is devoted to its proof.

4.4.3 Proof of Theorem 1

By translation invariance of the model, it is sufficient to prove the result when $y = 0 \in \mathscr{C}_\infty$. We fix an exponent $\delta > 0$; the objective is to apply the results of Proposition 4.4.1, Lemma 4.4.1 and Lemma 4.4.2 with the mesoscopic time $\tau = t^{1-\kappa}$ and with following values of exponents

$$\alpha := \frac{\delta}{2} \quad \text{and} \quad \kappa := \frac{\delta}{d+2},$$

For later use, we note that with these specific choices of exponents, the following estimates hold

$$(1-\kappa)\left(\frac{1}{2} - \frac{\alpha}{2}\right) > \frac{1}{2} - \delta \quad \text{and} \quad -\kappa + (1-\kappa)\left(\frac{d}{4} + \frac{1}{2} - \frac{\alpha}{2}\right) > \frac{d}{4} + \frac{1}{2} - \delta. \tag{4.148}$$

The proof relies on an induction argument and we give a setup of the proof. We first define the sequence κ_n of real numbers inductively by the formula

$$\kappa_0 = \frac{\kappa}{2} \quad \text{and} \quad \kappa_{n+1} := \min\left((1-\kappa)\kappa_n + \frac{\kappa}{2}, \frac{1}{2} - \delta\right) \tag{4.149}$$

This sequence is increasing and is ultimately constant equal to the value $\frac{1}{2} - \delta$. We let N be the integer

$$N := \inf\left\{n \in \mathbb{N} \ : \ \kappa_n = \frac{1}{2} - \delta\right\},$$

and we note that this integer only depends on the parameters $d, \mathfrak{p}, \lambda$ and δ. For each point $z \in \mathbb{Z}^d$, we define the random time $\mathcal{T}^0_{\mathrm{par}}(z)$ according to the formula

$$\mathcal{T}^0_{\mathrm{par}}(z) := 4\max\left(\mathcal{T}_{\mathrm{approx},\alpha}(z)^{\frac{1}{1-\kappa}}, \mathcal{M}_{\mathrm{corr},\alpha}(z)^{\frac{2}{1-\kappa}}, \mathcal{M}_{\mathrm{flux},\alpha}(z)^2\right)$$

so that for any time $t \geqslant \mathcal{T}^0_{\mathrm{par}}(z)$, all the results of Sections 4.4.1 and 4.4.2 are valid with the value $\tau := t^{1-\kappa}$. We then upgrade the random variable $\mathcal{T}^0_{\mathrm{par}}(z)$ and define

$$\mathcal{T}^1_{\mathrm{par}} := \sup\left\{t \in [1,\infty) \ : \ \exists z \in \mathscr{C}_\infty \text{ such that } |z| \leqslant (N+1)t^{\frac{1}{(1-\kappa)^N}} \text{ and } \mathcal{T}^0_{\mathrm{par}}(z) \geqslant t\right\}, \tag{4.150}$$

so that for any time $t \geqslant \mathcal{T}^1_{\mathrm{par}}$, and any point $z \in \mathscr{C}_\infty$ satisfying $|z| \leqslant (N+1)t^{\frac{1}{(1-\kappa)^N}}$, one has the estimate $t \geqslant \mathcal{T}^0_{\mathrm{par}}(z)$; this implies that all the results of Sections 4.4.1 and 4.4.2 are valid with the value $\tau := t^{1-\kappa}$ for the heat kernel started from the point z. This construction is identical to the used to define the minimal time $\mathcal{T}'_{\mathrm{NA}}(x)$ in (4.102). As it was the case for the random variable $\mathcal{T}'_{\mathrm{NA}}(x)$, an application of Lemma 4.1.1 shows the stochastic integrability estimate

$$\mathcal{T}^1_{\mathrm{par}} \leqslant \mathcal{O}_s(C).$$

For each integer $n \in \{0, \ldots, N\}$, we let H_n be the following statement.

Statement H_n. There exists a constant $C(d,\lambda,n) < \infty$ such that for each time $t \geqslant \left(\mathcal{T}^1_{\mathrm{par}}\right)^{\frac{1}{(1-\kappa)^n}}$, each point $x \in \mathscr{C}_\infty$, and each point $z \in \mathscr{C}_\infty$ satisfying $|z| \leqslant (N-n)t^{\frac{1}{(1-\kappa)^{N-n}}}$, one has the estimate

$$\left|p(t,x,z) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-z)\right| \leqslant Ct^{-\kappa_n}\Phi_C(t,x-z). \tag{4.151}$$

We prove by induction that the statement H_n holds for each integer $n \in \{0, \ldots, N\}$.

The base case. We prove that H_0 holds and first prove the L^2-estimate: for each time $t \geqslant \mathcal{T}_{\mathrm{par}}^1$, and each point $z \in \mathscr{C}_\infty$ satisfying $|z| \leqslant (N+1) t^{\frac{1}{(1-\kappa)^N}}$,

$$\left\| \left(p(t,\cdot,z) - \theta(\mathfrak{p})^{-1}\bar{p}(t,\cdot - z)\right) \exp\left(\Psi_C(t,|\cdot - z|)\right) \right\|_{L^2(\mathscr{C}_\infty)} \leqslant C t^{-\frac{d}{4}-\frac{\kappa}{2}}. \tag{4.152}$$

We recall the definitions of the functions h, q and v stated in (4.107), (4.110), (4.124) respectively as well as the definition of w given by the formula $w := h - v - q$. We write

$$\begin{aligned} p(t,x,z) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-z) = &\left(p(t,x,z) - q(t,x,\tau,z)\right) - v(t,x,\tau,z) + w(t,x,\tau,z) \\ &+ \left(h(t,x,z) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x,z)\right). \end{aligned} \tag{4.153}$$

To prove the estimate (4.152), we split the L^2-norm according to the decomposition (4.153) and estimate each terms thanks to the results established in Sections 4.4.1 and 4.4.2:

- The term $(p(t,x,z) - q(t,x,\tau,z))$ is estimated thanks to Lemma 4.4.1, this term accounts for an error of order
$$\left(\left(\frac{\tau}{t}\right)^{\frac{1}{2}} + \tau^{-\frac{1}{2}+\alpha}\right) t^{-\frac{d}{4}} = \left(t^{-\frac{\kappa}{2}} + t^{(1-\kappa)\left(-\frac{1}{2}+\alpha\right)}\right) t^{-\frac{d}{4}} \leqslant t^{-\frac{d}{4}-\frac{\kappa}{2}};$$
- The term w is estimated thanks to Proposition 4.4.1, this term accounts for an error of order
$$\left(\frac{t}{\tau}\right)^{\frac{1}{2}} \tau^{-\frac{d}{4}-\frac{1}{2}+\frac{\alpha}{2}} \leqslant t^{-\frac{d}{4}-\frac{1}{2}+\delta},$$
where we used the estimate (4.148);
- The term $v(t,x,\tau,z)$ is estimated thanks to Lemma 4.4.2, this term accounts for an error of order
$$t^{-\frac{d}{4}}\tau^{-\frac{1}{2}+\frac{\alpha}{2}} = t^{-\frac{d}{4}+(1-\kappa)\left(-\frac{1}{2}+\frac{\alpha}{2}\right)} \leqslant t^{-\frac{d}{4}-\frac{1}{2}+\delta},$$
where we used the estimate (4.148);
- The term $h(t,x,z) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x,z)$ can be estimated as follows. By the definition of h given in (4.107), we have
$$h(t,x,z) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x,z) = \sum_{k=1}^{d} \mathcal{D}_{e_k}\bar{p}(t,x-z)\phi_{e_k}(x).$$
The term can then be estimated by using the sublinearity of the corrector stated in Proposition 4.2.6 and the assumption $\sqrt{t} \geqslant \mathcal{M}_{\mathrm{corr},\alpha}(z)$. The proof is similar to the one of Lemma 4.4.2 and the details are left to the reader. It accounts for an error of order $t^{-\frac{d}{4}-\frac{1}{2}+\delta}$.

There remains to obtain the pointwise estimate (4.151) in the case $n = 0$ from the L^2-estimate (4.152). To this end, we fix a point $z \in \mathscr{C}_\infty$ such that $|z| \leqslant N t^{\frac{1}{(1-\kappa)^N}}$. We may without loss of generality restrict our attention to the points $x \in \mathscr{C}_\infty$ such that $|x| \leqslant (N+1) t^{\frac{1}{(1-\kappa)^N}}$, otherwise we necessarily have $|x - z| \geqslant t$ and the inequality (4.151) is satisfied by Proposition 4.2.8 and Remark 4.2.4. We then use the semigroup property on the heat kernels p and

$\bar{p}$: for each $x, z \in \mathscr{C}_\infty$, one has

$$\begin{aligned} p(t,x,z) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x-z) &= \underbrace{\int_{\mathscr{C}_\infty} p\left(\frac{t}{2},x,y\right) p\left(\frac{t}{2},y,z\right) - \theta(p)^{-2}\bar{p}\left(\frac{t}{2},x-y\right)\bar{p}\left(\frac{t}{2},y-z\right)dy}_{(4.154)\text{-a}} \\ &+ \theta(p)^{-1}\underbrace{\left(\theta(p)^{-1}\int_{\mathscr{C}_\infty}\bar{p}\left(\frac{t}{2},x-y\right)\bar{p}\left(\frac{t}{2},y-z\right)dy - \bar{p}(t,x-z)\right)}_{(4.154)\text{-b}}. \end{aligned} \tag{4.154}$$

We first treat the part (4.154)-a using the following L^2-estimate

$$|(4.154)\text{-a}| \leqslant (4.154)\text{-a1} + (4.154)\text{-a2},$$

where the two terms (4.154)-a1 and (4.154)-a2 are defined by the formulas

$$\begin{aligned} (4.154)\text{-a1} = &\left\| \left(p\left(\frac{t}{2},x,\cdot\right) - \theta(p)^{-1}\bar{p}\left(\frac{t}{2},x-\cdot\right)\right)\exp\left(\Psi_C\left(\frac{t}{2},|x-\cdot|\right)\right)\right\|_{L^2(\mathscr{C}_\infty)} \\ &\times \left\| p\left(\frac{t}{2},\cdot,z\right)\exp\left(-\Psi_C\left(\frac{t}{2},|x-\cdot|\right)\right)\right\|_{L^2(\mathscr{C}_\infty)} \end{aligned}$$

and

$$\begin{aligned} (4.154)\text{-a2} = &\left\| \left(p\left(\frac{t}{2},\cdot,z\right) - \theta(\mathfrak{p})^{-1}\bar{p}\left(\frac{t}{2},\cdot-z\right)\right)\exp\left(\Psi_C\left(\frac{t}{2},|\cdot-z|\right)\right)\right\|_{L^2(\mathscr{C}_\infty)} \\ &\times \left\| \theta(p)^{-1}\bar{p}\left(\frac{t}{2},x-\cdot\right)\exp\left(-\Psi_C\left(\frac{t}{2},|\cdot-z|\right)\right)\right\|_{L^2(\mathscr{C}_\infty)}. \end{aligned}$$

The term (4.154)-a1 can be estimated by using the three following ingredients:

- The symmetry of the heat kernel p;
- The L^2-estimate (4.152) applied with the point $z = x$ which is valid under the assumption $|x| \leqslant (N+1)t^{\frac{1}{(1-\kappa)^N}}$;
- The upper bound stated in Theorem 4.3.1, which can be applied since we assumed $t \geqslant \mathcal{T}^1_{\mathrm{par}} \geqslant 2\mathcal{T}'_{\mathrm{NA}}(z)$, and reads, by increasing the value of the constant C in the right side if necessary,

$$\left\| p\left(\frac{t}{2},\cdot,z\right)\exp\left(-\Psi_C\left(\frac{t}{2},|x-\cdot|\right)\right)\right\|_{L^2(\mathscr{C}_\infty)} \leqslant t^{\frac{d}{4}}\Phi_C(t,x-z).$$

These arguments imply the estimate

$$(4.154)\text{-a1} \leqslant t^{-\frac{\kappa}{2}}\Phi_C(t,x-y).$$

The term (4.154)-a2 can be treated similarly and we omit the details. There remains to estimate the term (4.154)-b. We note that by an application of the parallelogram law, i.e., the identity $|x-y|^2 + |y-z|^2 = 2\left(\left|\frac{x-z}{2}\right|^2 + \left|y - \frac{x+z}{2}\right|^2\right)$, the function $\bar{p}$ satisfies the following property: for each $t \geqslant 0$, and each $x, y, z \in \mathbb{R}^d$,

$$\bar{p}\left(\frac{t}{2},x-y\right)\bar{p}\left(\frac{t}{2},y-z\right) = \bar{p}\left(t,x-z\right)\bar{p}\left(\frac{t}{4},y-\frac{x+z}{2}\right).$$

By combining this identity with Proposition 4.A.3, we obtain

$$|(4.154)\text{-b}| = \bar{p}(t, x-z) \left| \int_{\mathscr{C}_\infty} \theta(p)^{-1} \bar{p}\left(\frac{t}{4}, y - \frac{x+z}{2}\right) dy - 1 \right| \leqslant Ct^{-\frac{1}{2}+\delta} \bar{p}(t, x-z).$$

This finishes the proof of the base case.

The iteration step. We prove that, for each integer $n \in \mathbb{N}$, the statement H_{n-1} implies the statement H_n. The strategy follows the one of the base case and we first prove the L^2-estimate, under the assumption that the statement H_{n-1} is valid: for each time $t \geqslant \left(\mathcal{T}^1_{\mathrm{par}}\right)^{\frac{1}{(1-\kappa)^n}}$, each point $x \in \mathscr{C}_\infty$, and each point $z \in \mathscr{C}_\infty$ satisfying $|z| \leqslant (N+1-n) t^{\frac{1}{(1-\kappa)^{N-n}}}$,

$$\left\| \left(p(t,\cdot,z) - \theta(\mathfrak{p})^{-1}\bar{p}(t,\cdot - z)\right) \exp\left(\Psi_C(t, |\cdot - z|)\right) \right\|_{L^2(\mathscr{C}_\infty)} \leqslant Ct^{-\frac{d}{4}-\kappa_n}. \tag{4.155}$$

We use the decomposition (4.153) with the same value for the mesoscopic time $\tau = t^{1-\kappa}$. The error introduced by the terms w, v and $h - \theta^{-1}(\mathfrak{p})\bar{p}$ are of order $t^{-\frac{d}{4}-\frac{1}{2}+\delta}$ which is smaller than the value $t^{-\frac{d}{4}-\kappa_n}$ we want to prove in this step. The limiting factor comes from the term $(p(t,x,z) - q(t,x,\tau,z))$ which is estimated in Lemma 4.4.1 and gives an error of order $t^{-\frac{d}{4}-\frac{\kappa}{2}}$. The objective of the induction step is to improve this error by using the statement H_{n-1}.

Under the assumption $t \geqslant \left(\mathcal{T}^1_{\mathrm{par}}\right)^{\frac{1}{(1-\kappa)^n}}$, we have $\tau = t^{1-\kappa} \geqslant \left(\mathcal{T}^1_{\mathrm{par}}\right)^{\frac{1}{(1-\kappa)^{n-1}}}$. We can thus apply the induction hypothesis H_{n-1} with time τ. This gives the inequality, for each point $x \in \mathscr{C}_\infty$, and each point $z \in \mathscr{C}_\infty$ satisfying $|z| \leqslant (N+1-n)\tau^{\frac{1}{(1-\kappa)^{N+1-n}}} = (N+1-n)t^{\frac{1}{(1-\kappa)^{N-n}}}$,

$$\left| p(\tau,x,z) - \theta(\mathfrak{p})^{-1}\bar{p}(\tau, x-z) \right| \leqslant C\tau^{-\frac{d}{4}-\kappa_{n-1}} \Phi_C(\tau, x-z). \tag{4.156}$$

This estimate can be used to improve the result of Lemma 4.4.1 according to the following procedure. We go back to the proof of Lemma 4.4.1 and in the inequality (4.113), instead of using the Nash-Aronson estimate stated in Theorem 4.3.1, we use the homogenization estimate (4.156). We then proceed with the proof and do not make any other modification. This implies the following improved version of Lemma 4.4.1

$$|q(t,x,\tau,z) - p(t,x,z)| \leqslant \left(\tau^{-\kappa_{n-1}} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} + \tau^{-\frac{1}{2}+\alpha} \right) \Phi_C(t, x-z).$$

Once equipped with this estimate, we can prove the L^2-estimate (4.155). The proof is the same as the one presented in the base case, we only use the estimate (4.155) instead of Lemma 4.4.1. We obtain, for any point $z \in \mathscr{C}_\infty$ satisfying $|z| \leqslant (N+1-n)\tau^{\frac{1}{(1-\kappa)^{N-n}}}$,

$$\left\| \left(p(t,\cdot,z) - \theta(\mathfrak{p})^{-1}\bar{p}(t,\cdot - z)\right) \exp\left(\Psi_C(t, |\cdot - z|)\right) \right\|_{L^2(\mathscr{C}_\infty)} \leqslant t^{-\frac{d}{4}} \tau^{-\kappa_{n-1}} \left(\frac{\tau}{t}\right)^{\frac{1}{2}} + t^{-\frac{d}{4}-\frac{1}{2}+\delta}.$$

We then use the equality $\tau = t^{1-\kappa}$ and the inductive definition of the sequence κ_n stated in (4.149) to deduce the estimate

$$\left\| \left(p(t,\cdot,z) - \theta(\mathfrak{p})^{-1}\bar{p}(t,\cdot - z)\right) \exp\left(\Psi_C(t, |\cdot - z|)\right) \right\|_{L^2(\mathscr{C}_\infty)} \leqslant Ct^{-\frac{d}{4}-\kappa_n}.$$

The proof of the pointwise estimate (4.151) is identical to the proof written for the base case and we omit the details. This completes the proof of the induction step.

We then define the minimal time $\mathcal{T}_{\mathrm{par},\delta}(0) := \left(\mathcal{T}^1_{\mathrm{par}}\right)^{\frac{1}{(1-\kappa)^N}}$. Since the statement H_N holds, we have the estimate, for each time $t \geqslant \mathcal{T}_{\mathrm{par},\delta}(0)$,

$$\left|p(t,x,0) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x)\right| \leqslant Ct^{-\frac{1}{2}+\delta}\Phi_C\left(t,x\right).$$

The proof of Theorem 4.1.1 is complete in the case $y = 0$. The proof in the general case is obtained by using the stationarity of the model.

4.5 Quantitative homogenization of the elliptic Green's function

The objective of this section is to present a theorem of quantitative homogenization for the elliptic Green's function on the infinite cluster, i.e., to establish Theorem 4.1.2. This result is a consequence of the quantitative homogenization theorem for the parabolic Green's function, Theorem 4.1.1, established in the previous section: in dimension $d \geqslant 3$, it can be essentially obtained by integrating the heat kernel over time since one has the identity, for each $x, y \in \mathscr{C}_\infty$,

$$g\left(x,y\right) = \int_0^\infty p(t,x,y)\,dt. \tag{4.157}$$

The case of the dimension 2 is more specific and requires some additional attention. In this setting the heat kernel is not integrable as the time t tends to infinity. This difficulty is related to the recurrence of the random walk on $\mathcal{Z}^2$ or to the unbounded behavior of the Green's function in dimension 2. To remedy this, we use a corrected version of the formula (4.157): for each $x, y \in \mathscr{C}_\infty$, one has

$$g(x,y) = \int_0^\infty \left(p(t,x,y) - p(t,y,y)\right)\,dt,$$

where g is the unique elliptic Green's function on the infinite cluster under the environment $\mathbf{a}$ such that $g(y,y) = 0$.

Proof of Theorem 4.1.2. We first treat the case of the dimension $d \geqslant 3$. By the stationarity of the model, we prove the result in the case $y = 0$. To simplify the notation we write $g(x)$ instead of $g(x,0)$. We let $\mathcal{T}_{\mathrm{par},\delta/2}(0)$ be the minimal time provided by Theorem 4.1.1 with exponent $\delta/2$ and define the minimal scale $\mathcal{M}_{\mathrm{ell},\delta}(0)$ according to the formula

$$\mathcal{M}_{\mathrm{ell},\delta}(0) := \mathcal{T}_{\mathrm{par},\delta/2}(0).$$

It is on purpose that we do not respect the parabolic scaling, and we need to have that $\mathcal{M}_{\mathrm{ell},\delta}(0) \gg \sqrt{\mathcal{T}_{\mathrm{par},\delta/2}(0)}$. As was mentioned in the introduction of this section, in dimension $d \geqslant 3$, we use the explicit formula (4.157) and note that Duhamel's principle implies the identity

$$\bar{g}\left(x\right) = \theta(\mathfrak{p})^{-1}\int_0^\infty \bar{p}(t,x)\,dt.$$

We obtain

$$\left|g(x) - \bar{g}(x)\right| \leqslant \int_0^\infty \left|p(t,x,0) - \theta(p)^{-1}\bar{p}(t,x)\right|\,dt.$$

We then split the integral at time $|x|$,

$$\begin{aligned}\left|g(x) - \bar{g}(x)\right| \leqslant & \int_0^{|x|} \left|p(t,x,0) - \theta(p)^{-1}\bar{p}(t,x)\right|\,dt \\ & + \int_{|x|}^\infty \left|p(t,x,0) - \theta(p)^{-1}\bar{p}(t,x)\right|\,dt,\end{aligned} \tag{4.158}$$

and estimate the two terms on the right side separately. The second term is the simplest one, we apply the quantitative estimate (4.7) provided by Theorem 4.1.1 and use the assumption $|x| \geqslant \mathcal{T}_{\mathrm{par},\delta/2}(0)$. This shows

$$\int_{|x|}^{\infty} \left|p(t,x,0) - \theta(p)^{-1}\bar{p}(t,x)\right| dt \leqslant C \int_{|x|}^{\infty} t^{-\frac{d}{2}-\frac{1}{2}+\frac{\delta}{2}} \exp\left(-\frac{|x|^2}{Ct}\right) dt \tag{4.159}$$
$$\leqslant C \int_{0}^{\infty} t^{-\frac{d}{2}-\frac{1}{2}+\frac{\delta}{2}} \exp\left(-\frac{|x|^2}{Ct}\right) dt$$
$$\leqslant C|x|^{-1+\delta}|x|^{2-d}.$$

To treat the first term in the right side of (4.158), we use the first estimate (4.50) of Proposition 4.2.8, which is recalled below, for each $t \in (0,\infty)$, $x \in \mathscr{C}_\infty$ such that $|x| \geqslant t$,

$$p(t,x,0) \leqslant C \exp\left(-C^{-1}|x|\left(1 + \ln\frac{|x|}{t}\right)\right).$$

The same estimate is also valid for the function $\bar{p}$. Therefore, the term $\ln(|x|/t)$ is positive on the interval $(0,|x|]$ and one has the estimate

$$\int_0^{|x|} \left|p(t,x,0) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x)\right| dt \leqslant C \int_0^{|x|} \exp\left(-C^{-1}|x|\left(1 + \ln\frac{|x|}{t}\right)\right) dt$$
$$\leqslant C \int_0^{|x|} \exp\left(-C^{-1}|x|\right) dt$$
$$\leqslant C|x| \exp\left(-C^{-1}|x|\right).$$

By increasing the value of the constant C, one has

$$\int_0^{|x|} \left|p(t,x,0) - \theta(\mathfrak{p})^{-1}\bar{p}(t,x)\right| dt \leqslant C \exp\left(-C^{-1}|x|\right).$$

Combining the previous estimate with (4.159), we deduce

$$|g(x) - \bar{g}(x)| \leqslant C|x|^{-1+\delta}|x|^{2-d} + C\exp\left(-C^{-1}|x|\right)$$
$$\leqslant C|x|^{-1+\delta}|x|^{2-d}.$$

This completes the proof of the estimate (4.10) in dimension larger than 3.

We now focus on the case of the dimension 2. The strategy is similar, but some additional attention is needed due to the fact that the integral (4.157) is ill-defined in dimension 2. We define the elliptic Green's function g on the infinite cluster by the formula

$$g(x) = \int_0^\infty (p(t,x,0) - p(t,0,0))\, dt. \tag{4.160}$$

For the homogenized Green's function, we cannot use the formula (4.160) by replacing the transition kernel p by the homogenized heat kernel $\bar{p}$; indeed the integral

$$\int_0^\infty (\bar{p}(t,x) - \bar{p}(t,0))\, dt, \tag{4.161}$$

is ill-defined as soon as $x \neq y$ since the term $\bar{p}(t,0)$ is of order t^{-1} around 0. To overcome this issue, we introduce the notation $(\bar{p}(t,\cdot))_{B_1} := \fint_{B_1} \bar{p}(t,z)\,dz$ and note that, for each $x \in \mathbb{R}^d$, the integral

$$\int_0^\infty \left(\bar{p}(t,x) - (\bar{p}(t,\cdot))_{B_1}\right) dt$$

is well-defined. Additionally, the function

$$x \mapsto \theta(p)^{-1} \left(\int_0^\infty \bar{p}(t,x) - (\bar{p}(t,\cdot))_{B_1} \, dt \right)$$

is equal to $\bar{g}$ up to a constant (see [98, Chapter 1.8] for detailed discussions). We denote this constant by K_1, i.e., we write, for any $x \in \mathbb{R}^d \setminus \{0\}$,

$$K_1 := \theta(p)^{-1} \left(\int_0^\infty \bar{p}(t,x) - (\bar{p}(t,\cdot))_{B_1} \, dt \right) - \bar{g}(x). \tag{4.162}$$

We note that the value K_1 depends only on the diffusivity $\bar{\sigma}^2$. Using these two integrals, we have

$$\begin{aligned} g(x) - \bar{g}(x) &= \int_0^\infty (p(t,x,0) - p(t,0,0)) - \theta(p)^{-1} \left(\bar{p}(t,x) - (\bar{p}(t,\cdot))_{B_1(y)} \right) dt + K_1 \\ &= \int_0^\infty \left(p(t,x,0) - \theta(p)^{-1} \bar{p}(t,x) \right) dt + K_1 - K_2(0), \end{aligned}$$

where K_2 is defined by the formula

$$K_2(0) := \int_0^\infty p(t,0,0) - \theta(p)^{-1} \left(\bar{p}(t,\cdot) \right)_{B_1} \, dt. \tag{4.163}$$

We now prove that this integral is well-defined, and that the constant K_2 satisfies the stochastic integrability estimate

$$|K_2(0)| \leqslant \mathcal{O}_s(C).$$

The proof relies on Theorem 4.1.1 and on the estimates on the discrete heat kernel $p(t,0,0) \leqslant 1$ and $(\bar{p}(t,\cdot))_{B_1} \leqslant 1$ for all times t. We compute

$$\begin{aligned} |K_2(0)| &\leqslant \int_0^\infty \left| p(t,0,0) - \theta(p)^{-1} \left(\bar{p}(t,\cdot) \right)_{B_1} \right| dt \\ &\leqslant \int_0^{\mathcal{T}_{\mathrm{par},\delta}(0)} \left| p(t,0,0) - \theta(p)^{-1} \left(\bar{p}(t,\cdot) \right)_{B_1} \right| dt + \int_{\mathcal{T}_{\mathrm{par},\delta}(0)}^\infty \left| p(t,0,0) - \theta(p)^{-1} \bar{p}(t,0) \right| dt \\ &\quad + \theta(p)^{-1} \int_{\mathcal{T}_{\mathrm{par},\delta}(0)}^\infty \left| \bar{p}(t,0) - (\bar{p}(t,\cdot))_{B_1} \right| dt \\ &\leqslant \int_0^{\mathcal{T}_{\mathrm{par},\delta}(0)} C \, dt + \int_{\mathcal{T}_{\mathrm{par},\delta}(0)}^\infty C t^{-\frac{3}{2}+\delta} \, dt + \int_{\mathcal{T}_{\mathrm{par},\delta}(0)}^\infty C t^{-\frac{3}{2}} \, dt \\ &\leqslant C \mathcal{T}_{\mathrm{par},\delta}(0) + C. \end{aligned}$$

This implies the estimate $|K_2(0)| \leqslant \mathcal{O}_s(C)$. We define $K(0) := K_1 - K_2(0)$, and by the previous computation, it satisfies the stochastic integrability estimate $|K(0)| \leqslant \mathcal{O}_s(C)$.

To complete the proof Theorem 4.1.2 in dimension 2, it is thus sufficient to control the term $\int_0^\infty \left(p(t,x,0) - \theta(p)^{-1} \bar{p}(t,x) \right) dt$; the argument is the same than in dimension larger than 3 and the details are omitted. □

4.A A concentration inequality for the density of $\mathscr{C}_\infty$

In this appendix, we study the density of the infinite cluster in a cube $\square$, which is defined as the random variable $\frac{|\mathscr{C}_\infty \cap \square|}{|\square|}$. As the size of the cube tends to infinity, an application of the ergodic theorem shows that this random variable converges, almost surely and in L^1, to the value $\theta(\mathfrak{p})$. The objective of the following proposition is to provide a quantitative version of this result.

Proposition 4.A.1. *There exists a positive constant $C(d,\mathfrak{p}) < \infty$ such that for any triadic cube $\square \in \mathcal{T}$ of size 3^m, one has an estimate*

$$\left| \frac{|\mathscr{C}_\infty \cap \square|}{|\square|} - \theta(\mathfrak{p}) \right| \leqslant \mathcal{O}_{\frac{2(d-1)}{3d^2+2d-1}} \left(C 3^{-\frac{dm}{2}} \right). \tag{4.164}$$

As a corollary, we obtain that, for any exponent $\alpha > 0$, there exist a positive constant $C(d,\mathfrak{p},\alpha) < \infty$, an exponent $s(d,\mathfrak{p},\alpha) > 0$, and a minimal scale $\mathcal{M}_{\mathrm{dense},\alpha} \leqslant \mathcal{O}_s(C)$ such that for every $3^m \geqslant \mathcal{M}_{\mathrm{dense},\delta}(y)$,

$$\left| \frac{|\mathscr{C}_\infty \cap \square_m|}{|\square_m|} - \theta(\mathfrak{p}) \right| \leqslant 3^{-\left(\frac{d}{2}-\alpha\right)m}. \tag{4.165}$$

Remark. The stochastic integrability exponent $\frac{2(d-1)}{3d^2+2d-1}$ in the estimate (4.164) is suboptimal and we do not try to reach optimality. The spatial scaling is the one of the central limit theorem and is optimal. We note that a result of large deviation for the concentration of the density of the infinite cluster can be found in the article [201, Theorem 1.2] of Pisztora: for any $\varepsilon > 0$ and $\mathfrak{p} > \mathfrak{p}_c(d)$, there exist two constants $C_1(\mathfrak{p},d,\varepsilon) < \infty, C_2(\mathfrak{p},d,\varepsilon) < \infty$ such that for any cube $\square$ of size 3^m,

$$\mathbb{P}\left(\left| \frac{|\mathscr{C}_\infty \cap \square|}{|\square|} - \theta(\mathfrak{p}) \right| > \varepsilon \right) \leqslant C_1 \exp\left(-C_2 3^{(d-1)m} \right).$$

However, this estimate cannot be used in the setting considered in this article since the dependence of the constants C_1 and C_2 in the variable ε is not explicit.

We prove Proposition 4.164 with an exponential version of the Efron-Stein inequality. A proof of this result can be found in [27, Proposition 2.2]. In the context of supercritical percolation, this inequality was used in [83, Proposition 2.18, Proposition 3.3] to study the corrector and in [134, Proposition 3.2] to study the flux. It is stated in the following proposition and we recall the notations introduced in Section 4.1.6: we denote by $(\Omega,\mathcal{F},\mathbb{P})$ the probability space and by $\mathcal{F}(E_d \backslash \{e\})$ denotes the sigma algebra generated by the collection of random variables $\{\mathbf{a}(e')\}_{e' \in E_d \backslash \{e\}}$.

Proposition 4.A.2 (Exponential Efron-Stein inequality, Proposition 2.2 of [27])**.** *Fix an exponent $\beta \in (0,2)$ and let X be a random variable defined on the probability space $(\Omega,\mathcal{F},\mathbb{P})$. We define the random variables*

$$X_e := \mathbb{E}\left[X | \mathcal{F}(E_d \backslash \{e\}) \right], \qquad \mathbb{V}[X] := \sum_{e \in E_d} (X - X_e)^2. \tag{4.166}$$

There exists a positive constant $C := C(d,\beta) < \infty$ such that

$$\mathbb{E}\left[\exp\left(|X - \mathbb{E}[X]|^\beta \right) \right] \leqslant C \mathbb{E}\left[\exp\left((C \mathbb{V}[X])^{\frac{\beta}{2-\beta}} \right) \right]^{\frac{2-\beta}{2}}. \tag{4.167}$$

We define $X := \frac{|\mathscr{C}_\infty \cap \square|}{|\square|} - \theta(\mathfrak{p})$. To prove Proposition 4.A.1, it suffices to prove the two inequalities

$$\mathbb{E}[X] \leqslant C_1(\mathfrak{p},d) 3^{-\frac{dm}{2}} \quad \text{and} \quad \mathbb{V}[X] \leqslant \mathcal{O}_{s'}\left(C_2(d,\mathfrak{p}) 3^{-dm} \right), \tag{4.168}$$

and to use the estimate (4.167) to deduce that $X \leqslant \mathcal{O}_s(C)$ with the exponent $s = \frac{2s'}{1+s'}$. These two inequalities are natural since they mean that the bias and variance of the random variable satisfy the desired upper bounds. Since the random variable $X = \frac{1}{|\square|}\sum_{x\in\square}\left(\mathbf{1}_{\{\{x\in\mathscr{C}_\infty\}\}} - \theta(\mathfrak{p})\right)$ is centered, we can focus on the term $\mathbb{V}[X]$.

To estimate this term, we consider an independent copy of the environment $\mathbf{a}$ which we denote by $\widetilde{\mathbf{a}}$ (and enlarge the underlying probability space to achieve this if necessary). Given a bond $e \in E_d$, we define $\{\mathbf{a}^e(e')\}_{e'\in E_d}$ "the environment obtained by resampling the conductance at the bond e" by the formula

$$\mathbf{a}^e(e') = \begin{cases} \mathbf{a}(e') & \text{if } e' \neq e, \\ \widetilde{\mathbf{a}}(e') & \text{if } e' = e. \end{cases}$$

We denote by X^e the random variable obtained by resampling the bond e, i.e., $X^e = X(\mathbf{a}^e)$. We also denote by $\mathscr{C}_\infty^e$ the infinite cluster under the environment $\mathbf{a}^e$. We have the following implication

$$\sum_{e\in E_d}(X^e - X)^2 \leqslant \mathcal{O}_{s'}(C3^{-dm}) \Longrightarrow \mathbb{V}[X] \leqslant \mathcal{O}_{s'}(C3^{-dm}), \tag{4.169}$$

whose proof can be found in [134, Lemma 3.1]. We note that since the two environments $\{\mathbf{a}(e')\}_{e'\in E_d}$ and $\{\mathbf{a}^e(e')\}_{e'\in E_d}$ are only different on one bond, the following statement holds $\mathbb{P}$-almost surely

$$\mathscr{C}_\infty^e \subseteq \mathscr{C}_\infty \text{ or } \mathscr{C}_\infty \subseteq \mathscr{C}_\infty^e.$$

We have the following identity

$$|X^e - X| = \frac{1}{|\square|}\left|(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty) \cap \square\right|,$$

where $\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty := (\mathscr{C}_\infty^e \setminus \mathscr{C}_\infty) \cup (\mathscr{C}_\infty \setminus \mathscr{C}_\infty^e)$ denotes the symmetric difference between the two clusters $\mathscr{C}_\infty$ and $\mathscr{C}_\infty^e$. This suggests to study the properties of this quantity and we prove the following lemma.

Lemma 4.A.1. *The following estimates hold:*

1. *There exists a positive constant $C(d,\mathfrak{p}) < \infty$ such that*

$$\forall e \in E_d, \qquad |\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty| \leqslant \mathcal{O}_{\frac{d-1}{d}}(C). \tag{4.170}$$

2. *There exists a positive constant $C(d,\mathfrak{p}) < \infty$ such that*

$$\forall e \in E_d \setminus E_d(3\square), \qquad \left|(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty) \cap \square\right|^2 \leqslant \mathcal{O}_{\frac{d-1}{(3d+1)d}}\left(\frac{C}{\operatorname{dist}(e,\square)^{d+1}}\right), \tag{4.171}$$

where we recall the notation $3\square$ introduced in (4.27). As a corollary, we have that

$$\sum_{e\in E_d\setminus E_d(3\square)} \left|(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty) \cap \square\right|^2 \leqslant \mathcal{O}_{\frac{d-1}{(3d+1)d}}(C). \tag{4.172}$$

We first show how to obtain Proposition 4.A.1 from Lemma 4.A.1 and then prove Lemma 4.A.1.

Proof of Proposition 4.A.1. The result is a consequence of the estimate (4.23) and Lemma 4.A.1. We have

$$\begin{aligned}\sum_{e\in E_d}(X^e-X)^2 &= \frac{1}{|\square|^2}\sum_{e\in E_d}\left(|\mathscr{C}_\infty^e\cap\square|-|\mathscr{C}_\infty\cap\square|\right)^2\\ &= 3^{-2dm}\underbrace{\sum_{e\in E_d(3\square)}|(\mathscr{C}_\infty^e\,\triangle\,\mathscr{C}_\infty)\cap\square|^2}_{\leqslant 3^{d(m+1)}\times\mathcal{O}_{\frac{d-1}{2d}}(C)}+3^{-2dm}\underbrace{\sum_{e\in E_d\setminus E_d(3\square)}|(\mathscr{C}_\infty^e\,\triangle\,\mathscr{C}_\infty)\cap\square|^2}_{\leqslant\mathcal{O}_{\frac{d-1}{(3d+1)d}}(C)}\\ &\leqslant \mathcal{O}_{\frac{d-1}{(3d+1)d}}\left(C3^{-dm}\right).\end{aligned}$$

We then complete the proof by applying the implication (4.169) and Proposition 4.A.2 with the exponent $s=\frac{2(d-1)}{3d^2+2d-1}$. □

We now prove Lemma 4.A.1. The argument relies on the upper and lower bounds on the tail of the distribution of the finite clusters in supercritical percolation. The result is stated below, was proved by Kesten and Zhang in [150] for the upper bound and by Aizenman, Delyon and Souillard in [1] for the lower bound. We also refer to the monograph [131, Section 8.6] for related discussions.

Theorem 4.A.1 (Sub-exponential decay of cluster size distribution [150, 1])**.** *For any supercritical probability $\mathfrak{p}\in(\mathfrak{p}_c(d),1]$, there exist positive constants $0<c_1(d,\mathfrak{p}),c_2(d,\mathfrak{p})<\infty$ such that, if we denote by $\mathscr{C}(0)$ the cluster containing 0 and let n be a strictly positive integer, then we have the estimate*

$$\forall n\in\mathbb{N}^+,\qquad \exp\left(-c_1n^{\frac{d-1}{d}}\right)\leqslant\mathbb{P}\left[|\mathscr{C}(0)|=n\right]\leqslant\exp\left(-c_2n^{\frac{d-1}{d}}\right).\tag{4.173}$$

Remark. With the notation $\mathcal{O}_s$, one can reformulate the upper bound as $|\mathscr{C}(0)|\leqslant\mathcal{O}_{\frac{d-1}{d}}(C)$.

Remark. The estimate (4.173) implies the inequality $\mathbb{P}\left[n\leqslant|\mathscr{C}(0)|<\infty\right]\leqslant\exp\left(-c_3n^{\frac{d-1}{d}}\right)$.

Proof of Lemma 4.A.1. We first prove the inequality (4.170). We use the definition of $\mathcal{O}_s$ notation and prove the estimate

$$\mathbb{E}\left[\exp\left(\left(\frac{|\mathscr{C}_\infty^e\,\triangle\,\mathscr{C}_\infty|}{C}\right)^{\frac{d-1}{d}}\right)\right]\leqslant 2,$$

for some constant $C(d,\mathfrak{p})<\infty$. By symmetry, it suffices to consider the case $\{\mathbf{a}^e(e)>0,\mathbf{a}(e)=0\}$ and we have the identity

$$\mathbb{E}\left[\exp\left(\left(\frac{|\mathscr{C}_\infty^e\,\triangle\,\mathscr{C}_\infty|}{C}\right)^{\frac{d-1}{d}}\right)\right]=1+2\mathbb{E}\left[\exp\left(\left(\frac{|\mathscr{C}_\infty^e\,\triangle\,\mathscr{C}_\infty|}{C}\right)^{\frac{d-1}{d}}\right)\mathbf{1}_{\{\{\mathbf{a}^e(e)>0,\mathbf{a}(e)=0\}\}}\right].$$

We then notice that, under the condition $\{\mathbf{a}^e(e)>0,\mathbf{a}(e)=0\}$, we have the equality $\mathscr{C}_\infty^e\,\triangle\,\mathscr{C}_\infty=\mathscr{C}_\infty^e\setminus\mathscr{C}_\infty$. We then distinguish two cases:

- Either there exists a finite cluster connected to the bond e in the environment $\{\mathbf{a}(e')\}_{e'\in E_d}$. In that case, we denote this cluster by $\mathscr{C}(e)$ and we have the identity

$$\mathscr{C}(e)=\mathscr{C}_\infty^e\,\triangle\,\mathscr{C}_\infty;$$

- Or both ends of the bond e are connected to the infinite cluster $\mathscr{C}_\infty$ under the environment $\{\mathbf{a}(e')\}_{e' \in E_d}$. In that case, we have the equality $\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty = \varnothing$.

We then use Theorem 4.A.1 to estimate the volume of the cluster $\mathscr{C}(e)$ and we obtain

$$\mathbb{E}\left[\exp\left(\left(\frac{|\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty|}{C}\right)^{\frac{d-1}{d}}\right)\mathbf{1}_{\{\{\mathbf{a}^e(e)>0,\mathbf{a}(e)=0\}\}}\right] \leqslant \sum_{n=0}^{\infty} \exp\left(\frac{n^{\frac{d-1}{d}}}{C^{\frac{d-1}{d}}}\right)\exp\left(-c_1 n^{\frac{d-1}{d}}\right).$$

Then, we can choose a constant C depending on the parameters d and $\mathfrak{p}$ such that

$$\mathbb{E}\left[\exp\left(\frac{|\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty|}{C}\right)^{\frac{d-1}{d}}\right] \leqslant 2.$$

This implies that $|\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty| \leqslant \mathcal{O}_{\frac{d-1}{d}}(C)$.

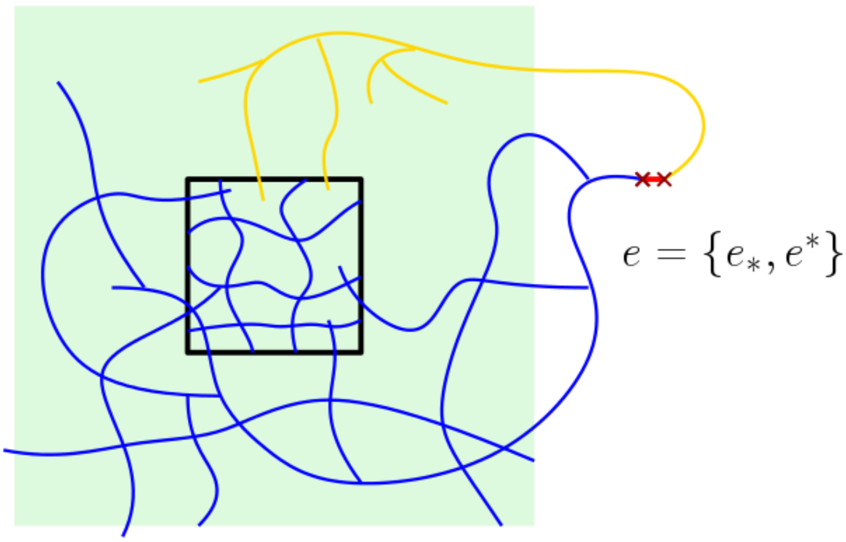

Figure 4.5: The figure illustrates the situation when the set $(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty) \cap \square$ is nonempty under the condition $\mathbf{a}^e(e) > 0$, $\mathbf{a}(e) = 0$. The blue cluster is the infinite cluster $\mathscr{C}_\infty$ under the environment $\mathbf{a}$, and the yellow cluster is the finite cluster connecting the bond e to the cube $\square$. The green square represents the cube $3\square$. The probability of the event depicted in the picture becomes exponentially small when the sizes of the cubes are large.

We now prove the estimate (4.171). It relies on the following observation: when the bond e is far away from the cube $\square$, the set $(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty) \cap \square$ is non-empty with exponentially small probability. More precisely, if we denote by $l = \operatorname{dist}(e, \square)$, then we have the estimate

$$\begin{aligned}\mathbb{P}\left[(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty) \cap \square \neq \varnothing\right] &= 2\mathbb{P}\left[\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty \neq \varnothing \text{ and } \mathscr{C}(e) \cap \square \neq \varnothing \text{ and } \mathbf{a}^e(e) > 0, \mathbf{a}(e) = 0\right] \quad (4.174)\\ &\leqslant 2\mathbb{P}[\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty \neq \varnothing \text{ and } |\mathscr{C}(e)| > l \text{ and } \mathbf{a}^e(e) > 0, \mathbf{a}(e) = 0]\\ &\leqslant 2\exp\left(-c_3 l^{\frac{d-1}{d}}\right).\end{aligned}$$

We also note that, since the bond e lies outside the cube $3\square$, we have the estimate $l \geqslant 3^m$. This implies the almost sure inequalities

$$|(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty) \cap \square| \leqslant 3^{dm} \leqslant l^d.$$

Then, we can calculate the expectation

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\left(\frac{l^{d+1}\left|(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty)\cap\square\right|^2}{C}\right)^{\frac{d-1}{(3d+1)d}}\right)\right] \\
=&\mathbb{E}\left[\mathbf{1}_{\{\{(\mathscr{C}_\infty^e\triangle\mathscr{C}_\infty)\cap\square=\varnothing\}\}}\right] + \mathbb{E}\left[\exp\left(\left(\frac{l^{d+1}\left|(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty)\cap\square\right|^2}{C}\right)^{\frac{d-1}{(3d+1)d}}\right)\mathbf{1}_{\{\{(\mathscr{C}_\infty^e\triangle\mathscr{C}_\infty)\cap\square\neq\varnothing\}\}}\right] \\
\leqslant&1 + \exp\left(\left(\frac{l^{3d+1}}{C}\right)^{\frac{d-1}{(3d+1)d}}\right)\mathbb{P}[(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty)\cap\square\neq\varnothing] \\
\leqslant&1 + \exp\left(\frac{l^{\frac{d-1}{d}}}{C^{\frac{d-1}{(3d+1)d}}}\right)\mathbb{P}[(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty)\cap\square\neq\varnothing].
\end{aligned}$$

We use the estimate (4.174) and select a constant C large enough such that

$$\mathbb{E}\left[\exp\left(\left(\frac{l^{d+1}\left|(\mathscr{C}_\infty^e \triangle \mathscr{C}_\infty)\cap\square\right|^2}{C}\right)^{\frac{d-1}{(3d+1)d}}\right)\right] \leqslant 1 + 2\exp\left(\frac{l^{\frac{d-1}{d}}}{C^{\frac{d-1}{(3d+1)d}}}\right)\exp\left(-c_4 l^{\frac{d-1}{d}}\right) \leqslant 2.$$

This completes the proof of the inequality (4.171). The estimate (4.172), is then a consequence of the inequality (4.23), by noting that the sum $\sum_{e\in E_d\setminus E_d(3\square)}\operatorname{dist}(e,\square)^{-d-1}$ is finite. Finally, we define

$$\mathcal{M}_{\mathrm{dense},\alpha} := \sup\left\{3^m \in \mathbb{N} : 3^{\left(\frac{d}{2}-\alpha\right)m}\left|\frac{|\mathscr{C}_\infty\cap\square_m|}{|\square_m|} - \theta(\mathfrak{p})\right| \geqslant 1\right\},$$

and use Lemma 4.1.1 to obtain that this random variable satisfies the stochastic integrability estimate $\mathcal{M}_{\mathrm{dense},\alpha} \leqslant \mathcal{O}_s(C)$. □

We complete this section by stating and proving a version of the concentration estimate of Lemma 4.A.1 involving the homogenized heat-kernel. This result is used in Lemma 4.4.1.

Proposition 4.A.3. *There exists a positive constant $C(d,\mathfrak{p}) < \infty$ such that, for any time $t > 0$, and any vertex $y \in \mathbb{Z}^d$, one has the estimate*

$$\left|\int_{\mathscr{C}_\infty} \bar{p}(t, x-y)\,dx - \theta(\mathfrak{p})\right| \leqslant \mathcal{O}_{\frac{2(d-1)}{3d^2+2d-1}}\left(Ct^{-\frac{1}{2}}\right). \tag{4.175}$$

As a corollary, for any $\alpha > 0$ and $y \in \mathbb{Z}^d$, there exist a positive constant $C(d,\mathfrak{p},\alpha) < \infty$, an exponent $s(d,\mathfrak{p},\alpha) > 0$, and a minimal time $\mathcal{T}_{\mathrm{dense},\alpha}(y) \leqslant \mathcal{O}_s(C)$ such that, for every time $t > \mathcal{T}_{\mathrm{dense},\alpha}(y)$, we have

$$\left|\int_{\mathscr{C}_\infty} \bar{p}(t, x-y)\,dx - \theta(\mathfrak{p})\right| \leqslant Ct^{-\left(\frac{1}{2}-\alpha\right)}. \tag{4.176}$$

Proof. Without loss of generality, we suppose that $y = 0$. The strategy is similar to the one of the proof of Proposition 4.A.1. We denote by $X := \int_{\mathscr{C}_\infty} \bar{p}(t,x)\,dx - \theta(\mathfrak{p})$, apply the concentration inequality stated in Proposition 4.A.2 and verify the two conditions (4.168) and (4.169). For the term involving the expectation, we have

$$|\mathbb{E}[X]| = \left|\int_{\mathbb{Z}^d} \bar{p}(t,x)\mathbf{1}_{\{\{x\in\mathscr{C}_\infty\}\}}\,dx - \theta(\mathfrak{p})\right| = \theta(p)\left|\int_{\mathbb{Z}^d}\bar{p}(t,x)\,dx - \int_{\mathbb{R}^d}\bar{p}(t,x)\,dx\right| \leqslant \frac{C}{\sqrt{t}},$$

by the estimate on the gradient of the heat kernel. We then focus on the variation, i.e.,

$$\sum_{e\in E_d}(X^e - X)^2 = \sum_{e\in E_d}\left(\int_{\mathbb{Z}^d}\bar{p}(t,x)\left(\mathbf{1}_{\{\{x\in\mathscr{C}_\infty^e\}\}} - \mathbf{1}_{\{\{x\in\mathscr{C}_\infty\}\}}\right)dx\right)^2.$$

We apply a multiscale analysis: we define the balls and annuli

$$B_{-1} := \varnothing, \qquad \forall n \geqslant 1,\ B_n := \{x\in\mathbb{Z}^d : |x| \leqslant 3^n\sqrt{t}\}, \qquad \forall n\geqslant 0,\ A_n := B_n \setminus B_{n-1}.$$

We also define, for any subset $A \subseteq \mathbb{Z}^d$, $I_A^e := \int_A \bar{p}(t,x)\left(\mathbf{1}_{\{\{x\in\mathscr{C}_\infty^e\}\}} - \mathbf{1}_{\{\{x\in\mathscr{C}_\infty\}\}}\right)dx$. This notation is useful to localize the random variables $(Y^e - Y)$. We write

$$\sum_{e\in E_d}(Y^e - Y)^2 = \sum_{e\in E_d}\left(I_{\mathbb{Z}^d}^e\right)^2 = \sum_{e\in E_d}\left(\sum_{n=0}^\infty I_{A_n}^e\right)^2.$$

Then, we use the Cauchy-Schwarz inequality to factorize the sum

$$\left(\sum_{n=0}^\infty I_{A_n}^e\right)^2 = \left(\sum_{n=0}^\infty I_{A_n}^e 3^n \times 3^{-n}\right)^2 \leqslant \left(\sum_{n=0}^\infty 3^{2n}\left(I_{A_n}^e\right)^2\right)\left(\sum_{n=0}^\infty 3^{-2n}\right) \leqslant 2\sum_{n=0}^\infty 3^{2n}\left(I_{A_n}^e\right)^2.$$

With Fubini's theorem, we obtain

$$\sum_{e\in E_d}(Y^e - Y)^2 \leqslant 2\sum_{n=0}^\infty 3^{2n}\sum_{e\in E_d}\left(I_{A_n}^e\right)^2. \tag{4.177}$$

We fix an integer $n\in\mathbb{N}$ and estimate the quantity $\sum_{e\in E_d}\left(I_{A_n}^e\right)^2$. The strategy is similar to the proof of Proposition 4.A.1 and we adapt the proof of Lemma 4.A.1 from the case of cubes to the case of balls

$$\begin{aligned}
\sum_{e\in E_d}\left(I_{A_n}^e\right)^2 &= \sum_{e\in E_d(B_{n+1})}\left(I_{A_n}^e\right)^2 + \sum_{e\in E_d\setminus E_d(B_{n+1})}\left(I_{A_n}^e\right)^2 \\
&\leqslant \left(\max_{A_n}\bar{p}(t,x)\right)^2\left(\sum_{e\in E_d(B_{n+1})}\left|\left(\mathscr{C}_\infty^e \bigtriangleup \mathscr{C}_\infty\right)\cap A_n\right|^2 + \sum_{e\in E_d\setminus E_d(B_{n+1})}\left|\left(\mathscr{C}_\infty^e \bigtriangleup \mathscr{C}_\infty\right)\cap A_n\right|^2\right) \\
&\leqslant \frac{1}{(2\pi t\bar{\sigma}^2)^{d/2}}\exp\left(-\frac{3^{2n}}{2\bar{\sigma}^2}\right)\left(\sum_{e\in E_d(B_{n+1})}\left|\left(\mathscr{C}_\infty^e \bigtriangleup \mathscr{C}_\infty\right)\cap B_n\right|^2 + \sum_{e\in E_d\setminus E_d(B_{n+1})}\left|\left(\mathscr{C}_\infty^e \bigtriangleup \mathscr{C}_\infty\right)\cap B_n\right|^2\right) \\
&\leqslant \mathcal{O}_{\frac{d-1}{(3d+1)d}}\left(C3^{dn}t^{-\frac{d}{2}}\exp\left(-\frac{3^{2n}}{2\bar{\sigma}^2}\right)\right).
\end{aligned}$$

We put this inequality back in equation (4.177) and use the estimate (4.23) to conclude

$$\sum_{e\in E_d}(Y^e - Y)^2 \leqslant \mathcal{O}_{\frac{d-1}{(3d+1)d}}\left(Ct^{-\frac{d}{2}}\left(\sum_{n=0}^\infty 3^{(d+2)n}\exp\left(-\frac{3^{2n}}{2\bar{\sigma}^2}\right)\right)\right) \leqslant \mathcal{O}_{\frac{d-1}{(3d+1)d}}\left(Ct^{-\frac{d}{2}}\right).$$

Finally, for any exponent $\alpha > 0$, we define

$$\mathcal{T}_{\mathrm{dense},\alpha}(0) := \sup\left\{t\in(0,\infty) : t^{\left(\frac{1}{2}-\alpha\right)}\left|\int_{\mathscr{C}_\infty}\bar{p}(t,x)\,dx - \theta(\mathfrak{p})\right| \geqslant 1\right\},$$

and apply Lemma 4.1.1 to conclude Proposition 4.A.3. □

4.B Quantification of the weak norm of the flux on $\mathscr{C}_\infty$

In this appendix, we prove a quantification of the $\underline{H}^{-1}$-norm of the flux on the cluster. We recall that the flux on the cluster associated to the direction e_k is defined by

$$\widetilde{\mathbf{g}}_{e_k} : \mathscr{C}_\infty \to \mathbb{R}^d, \qquad \widetilde{\mathbf{g}}_{e_k} = \mathbf{a}(\mathcal{D}\phi_{e_k} + e_k) - \frac{1}{2}\bar{\sigma}^2 e_k. \tag{4.178}$$

The main estimate is stated in the following proposition.

Proposition 4.B.1. *Fix a point $y \in \mathbb{Z}^d$, for each exponent $\alpha > 0$, there exist a positive constant $C := C(\lambda, d, \mathfrak{p}, \alpha) < \infty$, an exponent $s := s(\lambda, d, \mathfrak{p}, \alpha) > 0$, and a random variable $\mathcal{M}_{\mathrm{flux},\alpha}(y)$ satisfying the stochastic integrability estimate*

$$\mathcal{M}_{\mathrm{flux},\alpha}(y) \leqslant \mathcal{O}_s(C),$$

such that, for every radius $r \geqslant \mathcal{M}_{\mathrm{flux},\alpha}(y)$, one has

$$\sum_{k=1}^d \|\widetilde{\mathbf{g}}_{e_k}\|_{\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r(y))} \leqslant C r^\alpha. \tag{4.179}$$

Without loss of generality, we assume $y = 0$. The strategy of the proof is to make use of another centered flux defined on the entire space $\mathbb{Z}^d$,

$$\mathbf{g}_{e_k} : \mathbb{Z}^d \to \mathbb{R}^d, \qquad \mathbf{g}_{e_k} = \mathbf{a}(\mathcal{D}\phi_{e_k} + e_k) - \bar{\mathbf{a}} e_k.$$

The homogenized conductance $\bar{\mathbf{a}}$ is defined in [19, Definition 5.1] by the formula: for each $p \in \mathbb{R}^d$,

$$\frac{1}{2} p \cdot \bar{\mathbf{a}} p := \lim_{n\to\infty} \mathbb{E}\left[\nu\left(\square_m, p\right)\right],$$

where the energy $\nu(\square_m, p)$ is defined by

$$\nu\left(\square_m, p\right) := \inf_{u \in l_p + C_0(\mathscr{C}_\infty \cap \square_m)} \frac{1}{2|\square_m|} \int_{\square_m} \nabla u \cdot \mathbf{a} \nabla u, \tag{4.180}$$

where the notation l_p denotes the affine function of slope p (i.e., for each point $x \in \mathbb{Z}^d$, $l_p(x) = p \cdot x$) and the symbol $C_0(\mathscr{C}_\infty \cap \square_m)$ denotes the set of functions defined on the set $\mathscr{C}_\infty \cap \square_m$, valued in $\mathbb{R}$, which are equal to 0 on the boundary $\mathscr{C}_\infty \cap \partial \square_m$. The reason we introduce this quantity is that, building upon the results of [19], we can prove the following $\underline{H}^{-1}$-estimate: there exists a non-negative random variable $\mathcal{M}_{\mathrm{flux}-\mathbb{Z}^d,\alpha}$ satisfying the stochastic integrability estimate $\mathcal{M}_{\mathrm{flux}-\mathbb{Z}^d,\alpha} \leqslant \mathcal{O}_s(C)$ such that, for every $r \geqslant \mathcal{M}_{\mathrm{flux}-\mathbb{Z}^d,\alpha}$,

$$\sum_{k=1}^d \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \leqslant C r^\alpha, \tag{4.181}$$

where the $\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)$-norm is defined by the formula

$$\|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} = \sup_{\|\varphi\|_{\underline{H}^1(\mathbb{Z}^d \cap B_r)} \leqslant 1} \frac{1}{|\mathbb{Z}^d \cap B_r|} \int_{\mathbb{Z}^d \cap B_r} \varphi \mathbf{g}_{e_k},$$

and where the $\underline{H}^1(\mathbb{Z}^d \cap B_r)$-norm of a function $\varphi : \mathbb{Z}^d \cap B_r \to \mathbb{R}$, denotes the discrete normalized Sobolev norm defined by the formula

$$\|\varphi\|^2_{\underline{H}^1(\mathbb{Z}^d \cap B_r)} := r^{-1} \|\varphi\|_{\underline{L}^2(\mathbb{Z}^d \cap B_r)} + \|\nabla \varphi\|_{\underline{L}^2(\mathbb{Z}^d \cap B_r)}.$$

Once this result is established, we set the value

$$\bar{\sigma}^2 := 2\theta(\mathfrak{p})^{-1}\bar{\mathbf{a}}, \tag{4.182}$$

and deduce Proposition 4.B.1 from the estimate. The main difference between the estimates (4.179) and (4.181) is that in the former estimate, the $\underline{H}^{-1}$-norm is computed on the ball B_r while in the latter is computed on the intersection $\mathscr{C}_\infty \cap B_r$. This makes an important difference and motivates the introduction of the diffusivity $\bar{\sigma}^2$ in (4.182) and of the new flux $\widetilde{\mathbf{g}}_{e_k}$ in (4.178). In the following paragraph, we give an heuristic argument explaining why we expect Proposition 4.B.1 to hold assuming that the estimate (4.181) is valid.

We start by using the constant test function equal to 1 in the definition of the $\underline{H}^{-1}(B_r)$ norm in the estimate (4.181) shows

$$\int_{\mathbb{Z}^d \cap B_r} \mathbf{a}(\mathcal{D}\phi_{e_k} + e_k) \simeq \int_{\mathbb{Z}^d \cap B_r} \bar{\mathbf{a}} e_k, \tag{4.183}$$

where the symbol $\simeq$ means that the two quantities on the left and right sides differ by a small term, which by (4.181) is of order $r^{-(1-\alpha)}$. Since the function $\mathbf{a}(\mathcal{D}\phi_{e_k} + e_k)$ is defined to be equal to 0 outside the infinite cluster, the left side of (4.183) can be rewritten

$$\int_{\mathbb{Z}^d \cap B_r} \mathbf{a}(\mathcal{D}\phi_{e_k} + e_k) = \int_{\mathscr{C}_\infty \cap B_r} \mathbf{a}(\mathcal{D}\phi_{e_k} + e_k).$$

For the right-hand side of (4.183), using that the density of the cluster has density $\theta(\mathfrak{p})$, one expects

$$\int_{\mathbb{Z}^d \cap B_r} \bar{\mathbf{a}} e_k \simeq \int_{\mathscr{C}_\infty \cap B_r} \theta(\mathfrak{p})^{-1} \bar{\mathbf{a}} e_k.$$

This shows

$$\int_{\mathscr{C}_\infty \cap B_r} \mathbf{a}(\mathcal{D}\phi_{e_k} + e_k) \simeq \int_{\mathscr{C}_\infty \cap B_r} \theta(\mathfrak{p})^{-1} \bar{\mathbf{a}} e_k.$$

Thus if we want the estimate (4.179) to hold, the only admissible value for the coefficient $\bar{\sigma}^2$ is $2\theta(\mathfrak{p})^{-1}\bar{\mathbf{a}}$, indeed testing the constant function equal to 1 in the definition of the $\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r)$-norm of the estimate (4.179) shows

$$\int_{\mathscr{C}_\infty \cap B_r} \mathbf{a}(\mathcal{D}\phi_{e_k} + e_k) \simeq \int_{\mathscr{C}_\infty \cap B_r} \frac{1}{2}\bar{\sigma}^2 e_k.$$

Remark. We note that the identity (4.182) is the definition of the diffusivity $\bar{\sigma}^2$ used in this article: thanks to this definition and the result of Proposition 4.B.1, we are able to prove Theorem 4.1.1, and then to recover the invariance principle stated in (4.4).

The rest of this section is organized as follows. We first explain how to prove the estimate (4.181) by using the results of [134] and the strategies of stochastic homogenization in the uniformly elliptic setting presented in [25]. We then show how to deduce Proposition 4.B.1 from the inequality (4.181).

Proof of the estimate (4.181). We first extend the function $\mathbf{g}_{e_k}$ from $\mathbb{Z}^d$ to $\mathbb{R}^d$ and let $[\mathbf{g}_{e_k}]$ be the function defined on $\mathbb{R}^d$, which is equal to $\mathbf{g}_{e_k}$ on $\mathbb{Z}^d$ and which is piecewise constant on the unit cubes $z + \left[-\frac{1}{2}, \frac{1}{2}\right)^d$. We have the identity $\|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \simeq C \,\|[\mathbf{g}_{e_k}]\|_{\underline{H}^{-1}(B_r)}$ up to a constant C depending only on the dimension, where $\underline{H}^{-1}(B_r)$ is the standard Sobolev norm. We then want to control the continuous $\underline{H}^{-1}(B_r)$ norm of $[\mathbf{g}_{e_k}]$. The strategy is to apply the multiscale Poincaré inequality stated in [25, Remark D.6, equation (D.28)]. Its rescaled version reads

$$\|[\mathbf{g}_{e_k}]\|_{\underline{H}^{-1}(B_r))} \leqslant Cr\left(\int_0^1 \left(\int_{\mathbb{R}^d} r^{-d} e^{-\frac{|x|}{r}} |\Phi_{r^2 t} \star [\mathbf{g}_{e_k}]|^2(x)\,dx\right) dt\right)^{\frac{1}{2}}, \tag{4.184}$$

where the function Φ_t is the standard heat kernel defined by $\Phi_t := \frac{1}{(2\pi t)^{\frac{d}{2}}} \exp\left(-\frac{|x|^2}{2t}\right)$ and the operator $\star$ is the standard convolution on $\mathbb{R}^d$. We then apply the following results:

- The spatial average of the flux decays: one has the estimate, for each $t > 0$,

$$|\Phi_t \star [\mathbf{g}_{e_k}]| \leqslant \mathcal{O}_s\left(C t^{-\frac{d}{4}}\right). \tag{4.185}$$

  A proof of this result can be found in [134, Section 3.1, Proposition 1.1] (see Proposition 3.1.2).

- The flux is essentially bounded: one has the estimate, for each $t > 0$,

$$|\Phi_t \star [\mathbf{g}_{e_k}]| \leqslant \mathcal{O}_s(C), \tag{4.186}$$

  To prove this estimate, we first note that the bound on the corrector stated in (4.46) imply the following Lipschitz estimate on the corrector (by choosing x and y to be two neighboring points): for each vector $p \in B_1$, and each edge $e \in E_d$,

$$|\nabla \phi_p(e)| \leqslant \mathcal{O}_s\left(C|p|\right). \tag{4.187}$$

  This estimate is also stated in [19, Remark 1.1]. The inequality (4.186) is then a consequence of the estimate (4.187) and the property (4.23) of the $\mathcal{O}_s$ notation.

We then truncate the integral in the right side of (4.184) at the value $t = r^{-2}$ and obtain

$$\begin{aligned}\|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \leqslant Cr\left(\int_0^{r^{-2}} \left(\int_{\mathbb{R}^d} r^{-d} e^{-\frac{|x|}{r}} |\Phi_{r^2 t} \star [\mathbf{g}_{e_k}]|^2(x)\,dx\right) dt\right)^{\frac{1}{2}} \\ + Cr\left(\int_{r^{-2}}^1 \left(\int_{\mathbb{R}^d} r^{-d} e^{-\frac{|x|}{r}} |\Phi_{r^2 t} \star [\mathbf{g}_{e_k}]|^2(x)\,dx\right) dt\right)^{\frac{1}{2}}.\end{aligned}$$

To estimate the first term in the right side, we apply the estimate (4.186) and, to estimate the second term, we apply the estimate (4.185). Together with the property (4.23) of the $\mathcal{O}_s$ notation, this gives

$$\|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \leqslant \begin{cases} \mathcal{O}_s(C) & \text{if } d \geqslant 3, \\ \mathcal{O}_s\left(\log^{\frac{1}{2}}(1+r)\right) & \text{if } d = 2. \end{cases} \tag{4.188}$$

Finally, for every exponent $\alpha > 0$, we set

$$\mathcal{M}_{\text{flux}-\mathbb{Z}^d,\alpha} := \sup\left\{ r \in \mathbb{R}^+ : r^{-\alpha} \sum_{k=1}^d \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \geqslant 1 \right\},$$

and apply Lemma 4.1.1. This completes the proof of the estimate (4.181). □

Proof of Proposition 4.B.1. We fix an exponent $\alpha > 0$. We define the exponent $q := \max\left(\frac{(2d-1)}{\alpha}, 2d\right)$ and split the proof into 3 steps.

Step 1. In this step, we establish the inequality, for any radius $r \geqslant \mathcal{M}_q(\mathcal{P})$,

$$\|\widetilde{\mathbf{g}}_{e_k}\|_{\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r)} \leqslant Cr^{\frac{\alpha}{2}}\left(\|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d\cap B_r)} + \left\|\frac{1}{2}\bar{\sigma}^2 e_k(\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p}))\right\|_{\underline{H}^{-1}(\mathbb{Z}^d\cap B_r)}\right), \tag{4.189}$$

where C is a constant depending only on the parameters $\lambda, d, \mathfrak{p}$. We recall the definition of the $\underline{H}^{-1}$-norm on the infinite cluster

$$\|\widetilde{\mathbf{g}}_{e_k}\|_{\underline{H}^{-1}(\mathscr{C}_\infty\cap B_r)} = \sup_{\|\varphi\|_{\underline{H}^1(\mathscr{C}_\infty\cap B_r)}\leqslant 1} \frac{1}{|\mathscr{C}_\infty\cap B_r|}\int_{\mathscr{C}_\infty\cap B_r}\varphi\widetilde{\mathbf{g}}_{e_k}.$$

We fix a function $\varphi : \mathscr{C}_\infty \cap B_r \to \mathbb{R}$ such that $\|\varphi\|_{\underline{H}^1(\mathscr{C}_\infty\cap B_r)} \leqslant 1$. The main idea is to extend the function φ from the infinite cluster to $\mathbb{Z}^d$. To this end, we use the coarsened function $[\varphi]_\mathcal{P}$ introduced in Section 4.2.1. We extend the function $\widetilde{\mathbf{g}}_{e_k}$ by 0 outside the infinite cluster so that we have

$$\int_{\mathscr{C}_\infty\cap B_r}\varphi\widetilde{\mathbf{g}}_{e_k} = \int_{\mathbb{Z}^d\cap B_r}[\varphi]_\mathcal{P}\widetilde{\mathbf{g}}_{e_k}.$$

Since the radius r is assumed to be larger than the minimal scale $\mathcal{M}_q(\mathcal{P})$, the ratio $\frac{|\mathbb{Z}^d\cap B_r|}{|\mathscr{C}_\infty\cap B_r|}$ is bounded from above by a constant $C(d,\mathfrak{p})$. Then, we compute

$$\begin{aligned}
&\frac{1}{|\mathscr{C}_\infty\cap B_r|}\int_{\mathscr{C}_\infty\cap B_r}\varphi\widetilde{\mathbf{g}}_{e_k} \\
&\quad= \frac{1}{|\mathscr{C}_\infty\cap B_r|}\int_{\mathbb{Z}^d\cap B_r}[\varphi]_\mathcal{P}(\widetilde{\mathbf{g}}_{e_k} - \mathbf{g}_{e_k}) + \frac{1}{|\mathscr{C}_\infty\cap B_r|}\int_{\mathbb{Z}^d\cap B_r}[\varphi]_\mathcal{P}\mathbf{g}_{e_k} \\
&\quad\leqslant \left(\frac{|\mathbb{Z}^d\cap B_r|}{|\mathscr{C}_\infty\cap B_r|}\right)\|[\varphi]_\mathcal{P}\|_{\underline{H}^1(\mathbb{Z}^d\cap B_r)}\left(\|\widetilde{\mathbf{g}}_{e_k} - \mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d\cap B_r)} + \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d\cap B_r)}\right) \\
&\quad\leqslant C(d,\mathfrak{p})\,\|[\varphi]_\mathcal{P}\|_{\underline{H}^1(\mathbb{Z}^d\cap B_r)}\left(\left\|\frac{1}{2}\bar{\sigma}^2 e_k(\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p}))\right\|_{\underline{H}^{-1}(\mathbb{Z}^d\cap B_r)} + \|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d\cap B_r)}\right),
\end{aligned} \tag{4.190}$$

where we used the equation $\bar{\mathbf{a}} = \frac{1}{2}\theta(\mathfrak{p})\bar{\sigma}^2$ to go from the second line to the third line. We then use the estimates (4.42) and (4.43) to estimate the $\underline{H}^1$-norm of the coarsened function φ in terms of the $\underline{H}^1$-norm of the function φ, and the assumption $r \geqslant \mathcal{M}_q(\mathcal{P})$ to estimate the size of the cubes of the partition. This gives

$$\|[\varphi]_\mathcal{P}\|_{\underline{H}^1(\mathbb{Z}^d\cap B_r)} \leqslant Cr^{\frac{\alpha}{2}}\|\varphi\|_{\underline{H}^1(\mathscr{C}_\infty\cap B_r)}.$$

Combining the previous estimate with the inequality (4.190) completes the proof of Step 1.

Step 2: Control over the quantity $\left\|\bar{\sigma}^2 e_k(\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p}))\right\|_{\underline{H}^{-1}(B_r)}$. We let $\square_m$ be the triadic cube such that $\square_{m-1} \subseteq B_r \subseteq \square_m$. We note that

$$\left\|\bar{\sigma}^2 e_k(\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p}))\right\|_{\underline{H}^{-1}(\mathbb{Z}^d\cap B_r)} \leqslant C\left\|\bar{\sigma}^2 e_k(\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p}))\right\|_{\underline{H}^{-1}(\square_m)},$$

where the constant C depends only on the dimension d. We apply another version of the multiscale Poincaré inequality, which is stated in [25, Proposition 1.7] (in the continuous setting, the extension to the discrete setting considered here does not affect the proof) and reads

$$\left\| \bar\sigma^2 e_k (\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\square_m)} \leqslant C \sum_{n=0}^{m-1} 3^n \left(\frac{1}{|3^n \mathbb{Z}^d \cap \square_m|} \sum_{y \in 3^n \mathbb{Z}^d \cap \square_m} \bar\sigma^4 (\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p}))^2_{y+\square_n} \right)^{\frac12},$$

where we recall the notation $(f)_{y+\square_n} = \frac{1}{|\square_n|} \sum_{x\in y+\square_n} f(x)$. We apply Proposition 4.A.1

$$\left(\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p})\right)_{y+\square_n} \leqslant \mathcal{O}_s\left(C 3^{-\frac{dn}{2}}\right).$$

Using that the dimension is larger than 2 and the property (4.23) of the $\mathcal{O}_s$ notation, we obtain

$$\left\| \bar\sigma^2 e_k (\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\square_m)} \begin{cases} \mathcal{O}_s(C) & \text{if } d \geqslant 3, \\ \mathcal{O}_s(Cm) & \text{if } d = 2. \end{cases}$$

We then apply Lemma 4.1.1 to the collection of random variables

$$X_m := 3^{-\frac{\alpha m}{2}} \left\| \bar\sigma^2 e_k (\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\square_m)},$$

to construct a minimal scale $\mathcal{M}_{\mathrm{cluster},\frac{\alpha}{2}}$ such that, for any radius $r \geqslant \mathcal{M}_{\mathrm{cluster},\frac{\alpha}{2}}$,

$$\left\| \bar\sigma^2 e_k (\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \leqslant C r^{\frac{\alpha}{2}}. \tag{4.191}$$

Step 3: The conclusion. We let $\mathcal{M}_{\mathrm{flux}-\mathbb{Z}^d,\frac{\alpha}{2}}$ be the minimal scale provided by equation (4.181) with the exponent $\frac{\alpha}{2}$. We define the random variable $\mathcal{M}_{\mathrm{flux},\alpha}(0)$ according to the formula

$$\mathcal{M}_{\mathrm{flux},\alpha}(0) := \max\left(\mathcal{M}_{\mathrm{cluster},\frac{\alpha}{2}}, \mathcal{M}_{\mathrm{flux}-\mathbb{Z}^d,\frac{\alpha}{2}}, \mathcal{M}_q(\mathcal{P})\right).$$

Combining the main results (4.189) of Step 1 and (4.191) of Step 2 shows, for any $r \geqslant \mathcal{M}_{\mathrm{flux},\alpha}(0)$,

$$\begin{aligned} \|\widetilde{\mathbf{g}}_{e_k}\|_{\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r)} &\leqslant C r^{\frac{\alpha}{2}} \left(\|\mathbf{g}_{e_k}\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} + \left\| \frac12 \bar\sigma^2 e_k (\mathbf{1}_{\{\mathscr{C}_\infty\}} - \theta(\mathfrak{p})) \right\|_{\underline{H}^{-1}(\mathbb{Z}^d \cap B_r)} \right) \\ &\leqslant C r^{\frac{\alpha}{2}} \left(r^{\frac{\alpha}{2}} + r^{\frac{\alpha}{2}} \right) \\ &\leqslant C r^{\alpha}. \end{aligned}$$

Thus, we obtain the main result (4.179) of Proposition 4.B.1. □

Remark. One result used in this article is a variation of Proposition 4.B.1. We are interested in another function $\widetilde{\mathbf{g}}^*_{e_k} : \mathscr{C}_\infty \to \mathbb{R}^d$, satisfying the identity, for each function $u : \mathscr{C}_\infty \to \mathbb{R}$,

$$\mathcal{D}^* \cdot \left(u \left(\mathbf{a} \left(\mathcal{D}\phi_{e_k} + e_k \right) - \frac12 \bar\sigma^2 e_k \right) \right) = \mathcal{D}^* u \cdot \widetilde{\mathbf{g}}^*_{e_k}. \tag{4.192}$$

One can check that the quantity $\widetilde{\mathbf{g}}^*_{e_k}$ is different from $\widetilde{\mathbf{g}}_{e_k}$ with an exact formula

$$\widetilde{\mathbf{g}}^*_{e_k} := \begin{pmatrix} T_{-e_1}\left[\mathbf{a}\left(\mathcal{D}\phi_{e_k} + e_k\right) - \frac12 \bar\sigma^2 e_k\right]_1 \\ \vdots \\ T_{-e_d}\left[\mathbf{a}\left(\mathcal{D}\phi_{e_k} + e_k\right) - \frac12 \bar\sigma^2 e_k\right]_d \end{pmatrix},$$

where the (minor) difference comes from the translation when applying the finite difference operator. The $\underline{H}^{-1}$-norm of the function $\widetilde{\mathbf{g}}^*_{e_k}$ can also be controlled and one has the following property: for any exponent $\alpha > 0$, any vertex $y \in \mathbb{Z}^d$, and any radius $r \geqslant \mathcal{M}_{\mathrm{flux},\alpha}(y)$, one has the estimate

$$\sum_{k=1}^{d} \left\| \widetilde{\mathbf{g}}^*_{e_k} \right\|_{\underline{H}^{-1}(\mathscr{C}_\infty \cap B_r(y))} \leqslant C r^{\alpha}.$$

The proof is identical to the proof of Proposition 4.B.1 and the details are left to the reader.

Chapter 5

Decay of the semigroup for particle systems

We show the heat kernel type variance decay $t^{-\frac{d}{2}}$, up to a logarithmic correction, for the semigroup of an infinite particle system on $\mathbb{R}^d$, where every particle evolves following a divergence-form operator with diffusivity coefficient that depends on the local configuration of particles. The proof relies on the strategy from [142], and generalizes the localization estimate to the continuum configuration space introduced by S. Albeverio, Y.G. Kondratiev and M. Röckner.

This chapter corresponds to the article [135].

5.1 Introduction

In this work, we study an interacting diffusive particle system in $\mathbb{R}^d$ and the heat kernel type estimate for its semigroup. Let us give an informal introduction to the model and main result at first. We denote by $\mathcal{M}_\delta(\mathbb{R}^d)$ the set of point measures of type $\mu = \sum_{i=1}^{\infty} \delta_{x_i}$ on $\mathbb{R}^d$, which we call *configurations* of particles, by $\mathcal{F}_U$ the σ-algebra generated by $\mu(V)$ tested with all the Borel set $V \subseteq U$, and use the shorthand $\mathcal{F} := \mathcal{F}_{\mathbb{R}^d}$. Let $\mathbb{P}_\rho$ be the Poisson point process of density $\rho \in (0,\infty)$ as the law for the configuration μ, with $\mathbb{E}_\rho$, $\mathbb{V}\mathrm{ar}_\rho$ the associated expectation and variance. We have $\mathbf{a}_\circ : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}^{d\times d}_{sym}$ an $\mathcal{F}_{B_1}$-measurable symmetric matrix, i.e. it only depends on the configuration in the unit ball B_1, and $|\xi|^2 \leqslant \xi \cdot \mathbf{a}_\circ \xi \leqslant \Lambda |\xi|^2$ for any $\xi \in \mathbb{R}^d$. Then let $\mathbf{a}(\mu, x) := \mathbf{a}_\circ(\tau_{-x}\mu)$ be the diffusive coefficient with local interaction at x, where τ_{-x} represents the transport operation by the direction $-x$. Denoting by $\mu_t := \sum_{i=1}^{\infty} \delta_{x_{i,t}}$ the configuration at time $t \geqslant 0$, our model can be informally described as an infinite-dimensional system with local interaction such that every particle $x_{i,t}$ evolves as a diffusion associated to the divergence-form operator $-\nabla \cdot \mathbf{a}(\mu_t, x_{i,t})\nabla$. More precisely, it is a Markov process $(\Omega, (\mathscr{F}_t)_{t\geqslant 0}, \mathbb{P}_\rho)$ defined by the *Dirichlet form*

$$\mathcal{E}^{\mathbf{a}}(f,f) := \mathbb{E}_\rho\left[\int_{\mathbb{R}^d} \nabla f(\mu,x)\cdot \mathbf{a}(\mu,x)\nabla f(\mu,x)\,\mathrm{d}\mu(x)\right], \tag{5.1}$$

where the directional derivative $e_k \cdot \nabla f(\mu, x) := \lim_{h\to 0} \frac{1}{h}(f(\mu - \delta_x + \delta_{x+he_k}) - f(\mu))$ along the canonical direction $\{e_k\}_{1\leqslant k\leqslant d}$ is defined for a family of suitable functions and $x \in \operatorname{supp}(\mu)$.

One may expect that the diffusion follows the heat kernel estimate established by the pioneering work of John Nash [192], as every single particle is a diffusion of divergence type. This is the object of our main theorem. We denote by $\mathscr{C}_c^\infty(\mathbb{R}^d)$ functions smooth with respect to the transport of every particle (see Section 5.2.1 for its rigorous definition), $\overline{Q}_{l_u} := \left[-\frac{l_u}{2}, \frac{l_u}{2}\right]^d$ the closed cube of side length l_u, $\mathscr{L}^\infty := L^\infty(\mathcal{M}_\delta(\mathbb{R}^d), \mathcal{F}, \mathbb{P}_\rho)$, and let $u_t := \mathbb{E}_\rho[u(\mu_t)|\mathscr{F}_0]$, then we have the following estimate.

Theorem 5.1.1 (Decay of variance)**.** *There exist two finite positive constants $\gamma := \gamma(\rho, d, \Lambda)$, $C := C(\rho, d, \Lambda)$ such that for any $u \in \mathscr{C}_c^\infty(\mathbb{R}^d) \cap \mathscr{L}^\infty$ which is $\mathcal{F}_{\overline{Q}_{l_u}}$-measurable, then we have*

$$\mathbb{V}\mathrm{ar}_\rho[u_t] \leqslant C(1 + |\log t|)^\gamma \left(\frac{1 + l_u}{\sqrt{t}}\right)^d \|u\|^2_{\mathscr{L}^\infty}. \tag{5.2}$$

Interacting particle systems remain an active research topic, and it is hard to list all the references. We refer to the excellent monographs [152, 157, 172, 215] for a panorama of the field. In recent years, many works in probability and stochastic processes illustrate the diffusion universality in various models: a well-understood model is the *random conductance model*, see [53] for a survey, and especially the *heat kernel bound* and *invariance principle* is established for the percolation clusters in [49, 180, 211, 179, 39, 41, 209]; from the view point of *stochastic homogenization*, the *quantitative* results are also proved in a series of work [30, 23, 31, 24, 123, 124, 120, 121], and the monograph [25], and these techiques also apply on the percolation clusters setting, as shown in [19, 83, 134, 85]; for the system of hard-spheres, Bodineau, Gallagher and Saint-Raymond prove that Brownian motion is the Boltzmann-Grad limit of a tagged particle in [57, 56, 58]. All these works make us believe that the model in this work should also have diffusive behavior in large scale or long time.

Notice that our model is of *non-gradient type*, and our result is established in the continuum configuration space rather than a function space on $\mathbb{R}^d$. In previous works, the construction of similar diffusion processes is studied by Albeverio, Kondratiev and Röckner using Dirichlet forms in [2, 3, 4, 5]; see also the survey [206]. To the best of our knowledge, we do not find Theorem 5.1.1 in the literature. While in the lattice side, let us remark one important work [142] by Janvresse, Landim, Quastel and Yau, where the decay of variance is proved in the $\mathbb{Z}^d$ zero range model, which is of *gradient type*. Since our research is inspired by [142] and also uses some of their techniques, we point out our contributions in the following.

Firstly, we give an explicit bound with respect to the size of the support of the local function u, that is uniform over t; the bound $\left(\frac{l_u}{\sqrt{t}}\right)^d$ captures the correct typical scale. For comparison, [142, Theorem 1.1] states the result

$$\mathbb{V}\mathrm{ar}_\rho[u_t] = \frac{[\widetilde{u}'(\rho)]^2\chi(\rho)}{[8\pi\phi'(\rho)t]^{\frac{d}{2}}} + o\left(t^{-\frac{d}{2}}\right), \tag{5.3}$$

which should be considered as the asymptotic behavior in long time, and the term $o\left(t^{-\frac{d}{2}}\right)$ is of type $(l_u)^{5d}t^{-(\frac{d}{2}+\varepsilon)}$ if one tracks carefully the dependence of l_u in the steps of the proof of [142, Theorem 1.1]. To get the typical scale $\left(\frac{l_u}{\sqrt{t}}\right)^d$, we do some combinatorial improvement in the intermediate coarse-graining argument in eq. (5.27); see also Figure 5.1 for illustration. On the other hand, we also wonder if we could establish a similar result as eq. (5.3) to identify the diffusive constant in the long time behavior. This an interesting question and one perspective in future research, but a major difficulty here is to characterize the effective diffusion constant, because the zero range model satisfies the *gradient condition* while our model does not. We believe that it is related to the *bulk diffusion coefficient* and the equilibrium density fluctuation in the lattice nongradient model as indicated in [215, eq.(2.14), Proposition 2.1].

Secondly, we extend a localization estimate to the continuum configuration space: under the same context of Theorem 5.1.1, and recalling that $\mathcal{F}_{\overline{Q}_K}$ represents the information of μ in the closed cube $\overline{Q}_K = \left[-\frac{K}{2}, \frac{K}{2}\right]^d$, we define $\mathsf{A}_K u_t := \mathbb{E}_\rho[u_t|\mathcal{F}_{\overline{Q}_K}]$, and show that for every $t \geqslant \max\left\{(l_u)^2, 16\Lambda^2\right\}$ and $K \geqslant \sqrt{t}$

$$\mathbb{E}_\rho\left[(u_t - \mathsf{A}_K u_t)^2\right] \leqslant C(\Lambda)\exp\left(-\frac{K}{\sqrt{t}}\right)\mathbb{E}_\rho\left[u^2\right]. \tag{5.4}$$

This is a key estimate appearing in [142, Proposition 3.1], and is also natural as $\sqrt{t}$ is the typical scale of diffusion, thus when $K \gg \sqrt{t}$ one gets very good approximation in eq. (5.4). Its generalization in the continuum configuration space is non-trivial, since in the proof of [142, Proposition 3.1], one tests the Dirichlet form with $\mathsf{A}_K u_t$, but in our model it is not in the domain of Dirichlet form $\mathcal{D}(\mathcal{E}^{\mathbf{a}})$ and one cannot put $\mathsf{A}_K u_t$ directly in the Dirichlet form eq. (5.1). This is one essential difference between our model and a lattice model. To solve it, we have to apply some regularization steps which we present in Theorem 5.4.1.

Finally, we remark kindly a minor error in the proof in [142] and fix it when revisiting the paper. This will be presented in Section 5.3.1 and Section 5.3.1.

The rest of this article is organized as follows. In Section 5.2, we define all the notations and the rigorous construction of our model. Section 5.3 is the main part of the proof of Theorem 5.1.1, where Section 5.3.1 gives its outline and we fix the minor error in [142] mentioned above. The proof of some technical estimates used in Section 5.3 are put in the last two sections, where Section 5.4 proves the localization estimate eq. (5.4) in continuum configuration space, and Section 5.5 serves as a toolbox of other estimates including spectral inequality, perturbation estimate and calculation of the entropy.

5.2 Preliminaries

5.2.1 Notations

In this part, we introduce the notations used in this paper. We write $\mathbb{R}^d$ for the d-dimensional Euclidean space, $B_r(x)$ for the ball of radius r centered at x. We denote by $Q_s(x) := x + \left(-\frac{s}{2}, \frac{s}{2}\right)^d$

as the cube of edge length s centered at x, and $\overline{Q}_s(x)$ for its closure. We also denote by B_r and Q_s respectively short for $B_r(0)$ and $Q_s(0)$. The lattice set is defined by $\mathcal{Z}_s := \mathbb{Z}^d \cap Q_s$.

Continuum configuration space

For any metric space (E,d), we denote by $\mathcal{M}(E)$ the set of Radon measures on E. For every Borel set $U \subseteq E$, we denote by $\mathcal{F}_U$ the smallest σ-algebra such that for every Borel subset $V \subseteq U$, the mapping $\mu \in \mathcal{M}(E) \mapsto \mu(V)$ is measurable. For a $\mathcal{F}_U$-measurable function $f : \mathcal{M}(E) \to \mathbb{R}$, we say that f supported in U i.e. $\mathrm{supp}(f) \subseteq U$. In the case $\mu \in \mathcal{M}(E)$ is of finite total mass, we write

$$⨍ f \, \mathrm{d}\mu := \frac{\int f \, \mathrm{d}\mu}{\int \mathrm{d}\mu}. \tag{5.5}$$

We also define the collection of point measure $\mathcal{M}_\delta(E) \subseteq \mathcal{M}(E)$

$$\mathcal{M}_\delta(E) := \left\{ \mu \in \mathcal{M}(E) : \mu = \sum_{i \in I} \delta_{x_i} \text{ for some } I \text{ finite or countable, and } x_i \in E \text{ for any } i \in I \right\},$$

which serves as the *continuum configuration space* where each Dirac measure stands the position of a particle. In this work we will mainly focus on the Euclidean space $\mathbb{R}^d$ and its associated point measure space $\mathcal{M}_\delta(\mathbb{R}^d)$, and use the shorthand notation $\mathcal{F} := \mathcal{F}_{\mathbb{R}^d}$.

We define two operations for elements in $\mathcal{M}_\delta(\mathbb{R}^d)$: *restriction* and *transport*.

- For every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$ and Borel set $U \subseteq \mathbb{R}^d$, we define the restriction operation $\mu \llcorner U$, such that for every Borel set $V \subseteq \mathbb{R}^d$, $(\mu \llcorner U)(V) = \mu(U \cap V)$. Then for a function $f : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}$ which is $\mathcal{F}_U$-measurable, we have $f(\mu) = f(\mu \llcorner U)$.

- The transport on the set is defined as

$$\forall h \in \mathbb{R}^d, U \subseteq \mathbb{R}^d, \tau_h U := \{y + h : y \in U\}.$$

  Then for every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$ and $h \in \mathbb{R}^d$, we define the transport operation $\tau_h \mu$ such that for every Borel set U, we have

$$\tau_h \mu(U) := \mu(\tau_{-h} U). \tag{5.6}$$

  For f an $\mathcal{F}_V$-measurable function, we also define the transport operation $\tau_h f$ as a pullback that

$$\tau_h f(\mu) := f(\tau_{-h} \mu), \tag{5.7}$$

  which is an $\mathcal{F}_{\tau_h V}$-measurable function.

Notice that the restriction operation can be defined similarly in $\mathcal{M}(E)$ for a metric space, but the transport operation requires that E is at least a vector space.

We fix $\rho > 0$ once and for all, and define $\mathbb{P}_\rho$ a probability measure on $(\mathcal{M}_\delta(\mathbb{R}^d), \mathcal{F})$, to be the Poisson measure on $\mathbb{R}^d$ with density ρ (see [151]). We denote by $\mathbb{E}_\rho$ the expectation, $\mathbb{V}\mathrm{ar}_\rho$ the variance associated with the law $\mathbb{P}_\rho$, and by μ the canonical $\mathcal{M}_\delta(\mathbb{R}^d)$-valued random variable on the probability space $(\mathcal{M}_\delta(\mathbb{R}^d), \mathcal{F}, \mathbb{P}_\rho)$. In the case $U \subseteq \mathbb{R}^d$ a bounded Borel set and f a $\mathcal{F}_U$-measurable function, we can rewrite the expectation $\mathbb{E}_\rho[f]$ in an explicit expression

$$\mathbb{E}_\rho[f] = \sum_{N=0}^{+\infty} e^{-\rho|U|} \frac{(\rho|U|)^N}{N!} ⨍_{U^N} f\left(\sum_{i=1}^N \delta_{x_i}\right) \mathrm{d}x_1 \cdots \mathrm{d}x_N. \tag{5.8}$$

For instance, for every bounded Borel set $U \subseteq \mathbb{R}^d$ and bounded measurable function $g : U \to \mathbb{R}$, we can write

$$\mathbb{E}_\rho \left[\int_U g(x)\, \mathrm{d}\mu(x) \right] = \rho \int_U g(x)\, \mathrm{d}x.$$

Notice that the measure μ is a Poisson point process under $\mathbb{P}_\rho$. In particular, the measures $\mu \llcorner U$ and $\mu \llcorner (\mathbb{R}^d \setminus U)$ are independent, and the conditional expectation $\mathbb{E}_\rho \left[\cdot | \mathcal{F}_{(\mathbb{R}^d \setminus U)} \right]$ can thus be described equivalently as an averaging over the law of $\mu \llcorner U$.

For any $1 \leqslant p < \infty$, we denote by $\mathscr{L}^p$ the set of $\mathcal{F}$-measurable functions $f : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}$ such that the norm

$$\|f\|_{\mathscr{L}^p} := \left(\mathbb{E}_\rho \left[|f|^p\right]\right)^{\frac{1}{p}},$$

is finite. We denote by $\mathscr{L}^\infty$ the norm defined by essential upper bound under $\mathbb{P}_\rho$.

Derivative and $\mathscr{C}_c^\infty(U)$

We define the directional derivative for a $\mathcal{F}$-measurable function $f : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}$. Let $\{e_k\}_{1 \leqslant k \leqslant n}$ be d canonical directions, for $x \in \operatorname{supp}(\mu)$, we define

$$\partial_k f(\mu, x) := \lim_{h \to 0} \frac{1}{h} \left(f(\mu - \delta_x + \delta_{x + he_k}) - f(\mu) \right),$$

if the limit exists, and the gradient as a vector

$$\nabla f(\mu, x) := (\partial_1 f(\mu, x), \partial_2 f(\mu, x), \cdots \partial_d f(\mu, x)).$$

One can define the function with higher derivative iteratively, but here we use a more natural way: for every Borel set $U \subseteq \mathbb{R}^d$ and $N \in \mathbb{N}$, let $\mathcal{M}_\delta(U, N) \subseteq \mathcal{M}_\delta(E)$ be defined as

$$\mathcal{M}_\delta(U, N) := \left\{ \mu \in \mathcal{M}_\delta(\mathbb{R}^d) : \mu = \sum_{i=1}^N \delta_{x_i}, x_i \in U \text{ for every } 1 \leqslant i \leqslant N \right\}.$$

Then a function $f : \mathcal{M}_\delta(U, N) \to \mathbb{R}$ can be identified with a function $\widetilde{f} : U^N \to \mathbb{R}$ by setting

$$\widetilde{f}(x) = \widetilde{f}(x_1, \ldots, x_N) := f\left(\sum_{i=1}^N \delta_{x_i} \right). \tag{5.9}$$

The function $\widetilde{f}$ is invariant under permutations of its N coordinates. Conversely, any function satisfying this symmetry can be identified with a function from $\mathcal{M}_\delta(U, N)$ to $\mathbb{R}$. We denote by $C^\infty(\mathcal{M}_\delta(U, N))$ the set of functions $f : \mathcal{M}_\delta(U, N) \to \mathbb{R}$ such that $\widetilde{f}$ is infinitely differentiable. For every $f \in C^\infty(\mathcal{M}_\delta(U, N))$ and $x_1, \ldots, x_N \in U$, the gradient at x_1 coincides with the its canonical sense for the coordinate x_1.

$$\nabla f\left(\sum_{i=1}^N \delta_{x_i}, x_1 \right) = \nabla_{x_1} \widetilde{f}(x_1, \ldots, x_N). \tag{5.10}$$

We denote by $\mathscr{C}_c^\infty(U)$ the set of functions $f : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}$ that satisfy:

1. there exists a compact Borel set $V \subseteq U$ such that f is $\mathcal{F}_V$-measurable;
2. for every $N \in \mathbb{N}$,

$$\text{the mapping } \left\{ \begin{array}{ccc} \mathcal{M}_\delta(U, N) & \to & \mathbb{R} \\ \mu & \mapsto & f(\mu) \end{array} \right. \text{ belongs to } C^\infty(\mathcal{M}_\delta(U, N)).$$

A more heuristic description for $f \in \mathscr{C}_c^\infty(U)$ is a function depending only on the information in a compact subset $V \subseteq U$, and when we do projection $f(\mu) = f(\mu \llcorner V)$ it can be identified as a function C^∞ with finite coordinate, and also smooth when the number of particles in V changes. However, as the number of particles can be arbitrarily large, $\mathscr{C}_c^\infty(U)$ is not a subset of any $\mathscr{L}^p$ with $p \geqslant 1$.

Sobolev space on $\mathcal{M}_\delta(U)$

We define the $\mathscr{H}^1(U)$ norm by

$$\|f\|_{\mathscr{H}^1(U)} := \left(\|f\|_{\mathscr{L}^2}^2 + \mathbb{E}_\rho\left[\int_U |\nabla f|^2 \, \mathrm{d}\mu \right] \right)^{\frac{1}{2}},$$

and denote by $\mathscr{H}^1(U)$ the Sobolev space of $\mathcal{F}$-measurable functions with finite norm. Let $\mathscr{H}_0^1(U)$ denote the completion with respect to this norm of the space

$$\left\{ f \in \mathscr{C}_c^\infty(U) \ : \ \|f\|_{\mathscr{H}^1(U)} < \infty \right\}.$$

We remark that $\mathscr{H}^1(U)$ is not necessarily $\mathcal{F}_U$-measurable.

5.2.2 Construction of model

Diffusion coefficient

In this part, we define the coefficient field of the diffusion. We give ourselves a symmetric matrix valued function $\mathbf{a}_\circ : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}^{d\times d}_{sym}$ which satisfies the following properties:

- uniform ellipticity: there exists $\Lambda \in [1, +\infty)$ such that for every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$ and every $\xi \in \mathbb{R}^d$,
$$|\xi|^2 \leqslant \xi \cdot \mathbf{a}_\circ(\mu)\xi \leqslant \Lambda |\xi|^2 ; \tag{5.11}$$
- locality: for every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$, $\mathbf{a}_\circ(\mu) = \mathbf{a}_\circ(\mu \llcorner B_1)$.

We extend $\mathbf{a}_\circ$ by stationarity using the transport operation defined in eq. (5.7): for every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$,

$$\mathbf{a}(\mu, x) := \tau_x \mathbf{a}_\circ(\mu) = \mathbf{a}_\circ(\tau_{-x}\mu).$$

A typical example of a coefficient field $\mathbf{a}$ of interest is $\mathbf{a}_\circ(\mu) := (1 + \mathbf{1}_{\{\{\mu(B_1)=1\}\}})\mathbf{Id}$ whose extension is given by $\mathbf{a}(\mu, x) := (1 + \mathbf{1}_{\{\{\mu(B_1(x))=1\}\}})\mathbf{Id}$. In words, for $x \in \mathrm{supp}(\mu)$, the quantity $\mathbf{a}(\mu, x)$ is equal to 2 whenever there is no other point than x in the unit ball around x, and is equal to 1 otherwise.

Markov process defined by Dirichlet form

In this part, we construct our infinite particle system on $\mathcal{M}_\delta(\mathbb{R}^d)$ by Dirichlet form (see [109, 177] for the notations). We define at first the non-negative bilinear symmetric form

$$\mathcal{E}^{\mathbf{a}}(f, g) := \mathbb{E}_\rho\left[\int_{\mathbb{R}^d} \nabla f(\mu, x) \cdot \mathbf{a}(\mu, x) \nabla g(\mu, x) \, \mathrm{d}\mu(x) \right],$$

on its *domain* $\mathcal{D}(\mathcal{E}^{\mathbf{a}})$ that

$$\mathcal{D}(\mathcal{E}^{\mathbf{a}}) := \mathscr{H}_0^1(\mathbb{R}^d).$$

We also use $\mathcal{E}^{\mathbf{a}}(f) := \mathcal{E}^{\mathbf{a}}(f,f)$ for short. It is clear that $\mathcal{E}^{\mathbf{a}}$ is *closed* and *Markovian* thus it is a *Dirichlet form*, so it defines the correspondence between the Dirichlet form and the generator $\mathcal{L}$ that

$$\mathcal{E}^{\mathbf{a}}(f,g) = \mathbb{E}_\rho\left[f(-\mathcal{L})g\right], \qquad \mathcal{D}(\mathcal{E}^{\mathbf{a}}) = \mathcal{D}(-\mathcal{L}).$$

and a $\mathscr{L}^2$ strongly continuous Markov semigroup $(P_t)_{t\geqslant 0}$. We denote by $(\mathscr{F}_t)_{t\geqslant 0}$ its filtration and $(\mu_t)_{t\geqslant 0}$ the associated $\mathcal{M}_\delta(\mathbb{R}^d)$-valued Markov process which stands the configuration of the particles, then for any $u \in \mathscr{L}^2$,

$$u_t(\mu) := P_t u(\mu) = \mathbb{E}_\rho[u(\mu_t)|\mathscr{F}_0],$$

is an element in $\mathcal{D}(\mathcal{E}^{\mathbf{a}})$ and is characterized by the parabolic equation on $\mathcal{M}_\delta(\mathbb{R}^d)$ that for any $v \in \mathcal{D}(\mathcal{E}^{\mathbf{a}})$

$$\mathbb{E}_\rho[u_t v] - \mathbb{E}_\rho[uv] = -\int_0^t \mathcal{E}^{\mathbf{a}}(u_s, v)\,\mathrm{d}s. \tag{5.12}$$

Finally, we remark that the average is conserved for u_t as we test eq. (5.12) by constant 1 that

$$\mathbb{E}_\rho[u_t] - \mathbb{E}_\rho[u] = -\int_0^t \mathbb{E}_\rho\left[\int_{\mathbb{R}^d} \nabla 1 \cdot \mathbf{a}(\mu,x)\nabla u_s(\mu,x)\,\mathrm{d}\mu\right]\mathrm{d}s = 0. \tag{5.13}$$

In this work, we focus more on the quantitative property of P_t; see [206] for more details about the trajectory property of similar type of process.

5.2.3 A solvable case

We propose a solvable model to illustrate that the behavior of this process is close to the diffusion and the rate of decay is the best one that we can expect.

In the following, we suppose that $\mathbf{a} = \frac{1}{2}$ which means that in fact every particle evolves as a Brownian motion i.e. $\mu = \sum_{i=1}^\infty \delta_{x_i}$, $\mu_t = \sum_{i=1}^\infty \delta_{B_t^{(i)}}$ that $\left(B_t^{(i)}\right)_{t\geqslant 0}$ is a Brownian motion issued from x_i and $\left(B_\cdot^{(i)}\right)_{i\in\mathbb{N}}$ is independent.

Example 5.2.1. *Let $u(\mu) := \int_{\mathbb{R}^d} f\,\mathrm{d}\mu$ with $f \in C_c^\infty(\mathbb{R}^d)$. In this case, we have*

$$u_t(\mu) = P_t u(\mu) = \mathbb{E}_\rho\left[u(\mu_t)|\mathscr{F}_0\right] = \mathbb{E}_\rho\left[\sum_{i\in\mathbb{N}} f\left(B_t^{(i)}\right)\middle|\mathscr{F}_0\right] = \int_{\mathbb{R}^d} f_t(x)\,\mathrm{d}\mu(x),$$

where $f_t \in C^\infty(\mathbb{R}^d)$ is the solution of the Cauchy problem of the standard heat equation: $\Phi_t(x) = \frac{1}{(2\pi t)^{\frac{d}{2}}}\exp\left(-\frac{|x|^2}{2t}\right)$ and $f_t(x) = \Phi_t \star f(x)$. Then we use the formula of variance for Poisson process

$$\begin{aligned}\mathbb{V}\mathrm{ar}_\rho\left[u\right] &= \rho\int_{\mathbb{R}^d} f^2(x)\,\mathrm{d}x = \rho\|f\|^2_{L^2(\mathbb{R}^d)},\\ \mathbb{V}\mathrm{ar}_\rho\left[u_t\right] &= \rho\int_{\mathbb{R}^d} f_t^2(x)\,\mathrm{d}x = \rho\|f_t\|^2_{L^2(\mathbb{R}^d)}.\end{aligned}$$

By the heat kernel estimate for the standard heat equation, we known that $\|f_t\|^2_{L^2(\mathbb{R}^d)} \simeq C(d)t^{-\frac{d}{2}}\|f\|^2_{L^1(\mathbb{R}^d)}$, thus the scale $t^{-\frac{d}{2}}$ is the best one that we can obtain. Moreover, if we take $f = \mathbf{1}_{\{Q_r\}}$, and

$t = r^{2(1-\varepsilon)}$ for a small $\varepsilon > 0$, then we see that the typical scale of diffusion is a ball of size $r^{1-\varepsilon}$. So for every $x \in Q_{r\left(1-r^{-\frac{\varepsilon}{2}}\right)}$, the value $f_t(x) \simeq 1 - e^{-r^{\frac{\varepsilon}{2}}}$ and we have

$$\mathbb{V}\mathrm{ar}_\rho\left[u_t\right] = \rho \int_{\mathbb{R}^d} f_t^2(x)\,\mathrm{d}x \geqslant \rho r^d (1 - r^{-\frac{\varepsilon}{2}}) = (1 - r^{-\frac{\varepsilon}{2}})\mathbb{V}\mathrm{ar}_\rho\left[u\right].$$

It illustrates that before the scale $t = r^2$, the decay is very slow so in the Theorem 5.1.1 the factor $\left(\frac{l_u}{\sqrt{t}}\right)^d$ is reasonable.

5.3 Strategy of proof

In this part, we state the strategy of the proof of Theorem 5.1.1. We will give a short outline in Section 5.3.1, which can be seen as an "approximation-variance decomposition", and then focus on the term approximation in Section 5.3.2. Several technical estimates will be used in this procedure and their proofs will be postponed in Section 5.4 and Section 5.5.

5.3.1 Outline

As mentionned, this work is inspired from [142], and we revisit the strategy here. We pick a centered $u \in \mathscr{C}_c^\infty(\mathbb{R}^d) \cap \mathscr{L}^\infty$ supported in Q_{l_u} such that $\mathbb{E}_\rho[u] = 0$ and this implies $\mathbb{E}_\rho[u_t] = 0$ from eq. (5.13). Then we set a multi-scale $\{t_n\}_{n \geqslant 0}, t_{n+1} = Rt_n$, where $R > 1$ is a scale factor to be fixed later. It suffices to prove that eq. (5.2) for every t_n, then for $t \in [t_n, t_{n+1}]$, one can use the decay of $\mathscr{L}^2$ that

$$\mathbb{E}_\rho[(u_t)^2] \leqslant \mathbb{E}_\rho[(u_{t_n})^2] \leqslant C(1 + \log(t_n))^\gamma \left(\frac{1 + l_u}{\sqrt{t_n}}\right)^d \|u\|_{\mathscr{L}^\infty}^2 \leqslant CR^{\frac{d}{2}}(1 + \log t)^\gamma \left(\frac{1 + l_u}{\sqrt{t}}\right)^d \|u\|_{\mathscr{L}^\infty}^2,$$

then by resetting the constant C one concludes the main theorem. Another ingredient of the proof is an "approximation-variance type decomposition":

$$\begin{aligned} u_t &= v_t + w_t, \\ v_t &:= u_t - \frac{1}{|\mathcal{Z}_K|} \sum_{y \in \mathcal{Z}_K} \tau_y u_t, \\ w_t &:= \frac{1}{|\mathcal{Z}_K|} \sum_{y \in \mathcal{Z}_K} \tau_y u_t, \end{aligned} \tag{5.14}$$

where we recall $\mathcal{Z}_K = Q_K \cap \mathbb{Z}^d$ is the lattice set of scale K. The philosophy of this decomposition is that in long time, the information in a local scale K is mixed, thus w_t as a spatial average is a good approximation of u_t and v_t is the error term. Thus, the following control Proposition 5.3.1 and Proposition 5.5.2 of the two terms w_t and v_t proves the main theorem Theorem 5.1.1.

Proposition 5.3.1. *There exists a finite positive number $C := C(d)$ such that for any $u \in \mathscr{C}_c^\infty(\mathbb{R}^d) \cap \mathscr{L}^\infty$ which is $\mathcal{F}_{\overline{Q}_{l_u}}$-measurable and with mean zero, and any $K \geqslant l_u$, we have*

$$\mathbb{V}\mathrm{ar}_\rho\left[\left(\frac{1}{|\mathcal{Z}_K|} \sum_{y \in \mathcal{Z}_K} \tau_y u_t\right)^2\right] \leqslant C(d) \left(\frac{l_u}{K}\right)^d \mathbb{E}_\rho[u^2]. \tag{5.15}$$

Proof. Then we can estimate the variance simply by $\mathscr{L}^2$ decay that

$$\mathbb{E}_\rho[(w_t)^2] = \mathbb{E}_\rho\left[\left(P_t\left(\frac{1}{|\mathcal{Z}_K|}\sum_{y\in\mathcal{Z}_K}\tau_y u\right)\right)^2\right] \leqslant \mathbb{E}_\rho\left[\left(\frac{1}{|\mathcal{Z}_K|}\sum_{y\in\mathcal{Z}_K}\tau_y u\right)^2\right] = \frac{1}{|\mathcal{Z}_K|^2}\sum_{x,y\in\mathcal{Z}_K}\mathbb{E}_\rho\left[(\tau_{x-y}u)u\right].$$

We know that for $|x-y| \geqslant l_u$, then the term $\tau_{x-y}u$ and u is independent so $\mathbb{E}_\rho\left[(\tau_{x-y}u)u\right] = 0$. This concludes eq. (5.15). □

Proposition 5.3.2. *There exist two finite positive numbers $C := C(d,\rho), \gamma := \gamma(d,\rho)$ such that for any $u \in \mathscr{C}_c^\infty(\mathbb{R}^d) \cap \mathscr{L}^\infty$ which is $\mathcal{F}_{Q_{l_u}}$-measurable, $K \geqslant l_u$ and v_t defined in eq. (5.14), $t_n \geqslant \max\left\{l_u^2, 16\Lambda^2\right\}$, $t_{n+1} = Rt_n$ with $R > 1$ we have*

$$(t_{n+1})^{\frac{d+2}{2}}\mathbb{E}_\rho[(v_{t_{n+1}})^2] - (t_n)^{\frac{d+2}{2}}\mathbb{E}_\rho[(v_{t_n})^2] \leqslant C(\log(t_{n+1}))^\gamma K^2(l_u)^d\|u\|^2_{\mathscr{L}^\infty} + \mathbb{E}_\rho[u^2]. \quad (5.16)$$

Proof of Theorem 5.1.1 from Proposition 5.3.2 and Proposition 5.3.1. For the case $t \leqslant (l_u)^2$ or $t \leqslant 16\Lambda^2$, the right hand side of eq. (5.2) is larger than $\mathbb{E}_\rho[u^2]$ and we can use the $\mathscr{L}^2$ decay to prove the theorem. Thus without loss of generality, we set $t_0 := \max\left\{l_u^2, 16\Lambda^2\right\}$ and put eq. (5.16) into eq. (5.2) by setting $K := \sqrt{t_{n+1}}$ that

$$\begin{aligned}
&\mathbb{E}_\rho[(u_{t_{n+1}})^2] \\
\leqslant& 2\mathbb{E}_\rho[(v_{t_{n+1}})^2] + 2\mathbb{E}_\rho[(w_{t_{n+1}})^2] \\
\leqslant& 2\left(\frac{t_n}{t_{n+1}}\right)^{\frac{d+2}{2}}\mathbb{E}_\rho[(v_{t_n})^2] + 2(t_{n+1})^{-\frac{d+2}{2}}\left(C(\log(t_{n+1}))^\gamma t_{n+1}(l_u)^d\|u\|^2_{\mathscr{L}^\infty} + \mathbb{E}_\rho[u^2]\right) \\
&\quad + 2\mathbb{E}_\rho[(w_{t_{n+1}})^2] \\
\leqslant& 4\left(\frac{t_n}{t_{n+1}}\right)^{\frac{d+2}{2}}\mathbb{E}_\rho[(u_{t_n})^2] + 2(t_{n+1})^{-\frac{d+2}{2}}\left(C(\log(t_{n+1}))^\gamma t_{n+1}(l_u)^d\|u\|^2_{\mathscr{L}^\infty} + \mathbb{E}_\rho[u^2]\right) \\
&\quad + 4\left(\frac{t_n}{t_{n+1}}\right)^{\frac{d+2}{2}}\mathbb{E}_\rho[(w_{t_n})^2] + 2\mathbb{E}_\rho[(w_{t_{n+1}})^2].
\end{aligned} \quad (5.17)$$

We set $U_n = (t_n)^{\frac{d}{2}}\mathbb{E}_\rho[(u_{t_n})^2]$ and put eq. (5.15) into the equation above, we have

$$U_{n+1} \leqslant \theta U_n + C_2\left((\log(t_{n+1}))^\gamma(l_u)^d\|u\|^2_{\mathscr{L}^\infty} + (t_{n+1})^{-1}\mathbb{E}_\rho[u^2]\right) + C_3(l_u)^d\mathbb{E}_\rho[u^2],$$

where $\theta = 4R^{-1}$. By choosing R large such that $\theta \in (0,1)$, we do a iteration for the equation above to obtain that

$$\begin{aligned}
U_{n+1} &\leqslant \sum_{k=1}^{n}\left(C_2\left((\log(t_{n+1}))^\gamma(l_u)^d\|u\|^2_{\mathscr{L}^\infty} + \mathbb{E}_\rho[u^2]\right) + C_3(l_u)^d\mathbb{E}_\rho[u^2]\right)\theta^{n-k} + U_0\theta^{n+1} \\
&\leqslant \frac{1}{1-\theta}\left(C_2\left((\log(t_{n+1}))^\gamma(l_u)^d\|u\|^2_{\mathscr{L}^\infty} + \mathbb{E}_\rho[u^2]\right) + C_3(l_u)^d\mathbb{E}_\rho[u^2]\right) + (l_u)^d\mathbb{E}_\rho[u^2] \\
\Longrightarrow& \mathbb{E}_\rho[(u_{t_{n+1}})^2] \leqslant C_4(\log(t_{n+1}))^\gamma\left(\frac{l_u}{\sqrt{t_{n+1}}}\right)^d\|u\|^2_{\mathscr{L}^\infty}.
\end{aligned}$$

□

Remark. We remark that there is a small error in the similar argument in [142, Proof of Proposition 2.2]: the authors apply eq. (5.16) from t_0 to t_n, and they neglect the change of scale in K at the endpoints $\{t_n\}_{n\geqslant 0}$. However, it does not harm the whole proof and we fix it here: we add one more step of decomposition in eq. (5.17), and put the iteration directly in u_t instead of v_t, which avoids the problem of the changes of K.

5.3.2 Error for the approximation

In this part, we prove Proposition 5.3.2. The proof can be divided into 6 steps.

Proof of Proposition 5.3.2. . *Step 1: Setting up.* To shorten the equation, we define

$$\Delta_n := (t_{n+1})^{\frac{d+2}{2}} \mathbb{E}_\rho[(v_{t_{n+1}})^2] - (t_n)^{\frac{d+2}{2}} \mathbb{E}_\rho[(v_{t_n})^2], \tag{5.18}$$

and it is the goal of the whole subsection. In the step setting up, we do derivative for the flow $t^{\frac{d+2}{2}} \mathbb{E}_\rho[(v_t)^2]$ that

$$\Delta_n = \int_{t_n}^{t_{n+1}} \left(\frac{d+2}{2}\right) t^{\frac{d}{2}} \mathbb{E}_\rho[(v_t)^2] - 2t^{\frac{d+2}{2}} \mathbb{E}_\rho[v_t(-\mathcal{L}v_t)] \,\mathrm{d}t. \tag{5.19}$$

Step 2: Localization. We set $\mathsf{A}_L v_t = \mathbb{E}\left[v_t | \mathcal{F}_{\overline{Q}_L}\right]$ and use it to approximate v_t in $\mathscr{L}^2$. Since it is a diffusion process, one can guess naturally a scale larger than $\sqrt{t}$ will have enough information for this approximation. In Theorem 5.4.1 we prove an estimate

$$\mathbb{E}_\rho\left[(v_t - \mathsf{A}_L v_t)^2\right] \leqslant C(\Lambda) \exp\left(-\frac{L}{\sqrt{t}}\right) \mathbb{E}_\rho\left[(v_0)^2\right],$$

and we choose $L = \lfloor \gamma \log(t_{n+1}) \rfloor \sqrt{t_{n+1}}, \gamma > \frac{d+4}{2}$ here, and put it back to eq. (5.19) to obtain

$$\begin{aligned} \Delta_n &\leqslant \int_{t_n}^{t_{n+1}} (d+2) t^{\frac{d}{2}} \mathbb{E}_\rho\left[(\mathsf{A}_L v_t)^2\right] + (d+2) t^{\frac{d}{2}-\gamma} \mathbb{E}_\rho\left[(v_0)^2\right] - 2t^{\frac{d+2}{2}} \mathbb{E}_\rho[v_t(-\mathcal{L}v_t)] \,\mathrm{d}t \\ &\leqslant \mathbb{E}_\rho[(u_0)^2] + \int_{t_n}^{t_{n+1}} (d+2) t^{\frac{d}{2}} \mathbb{E}_\rho\left[(\mathsf{A}_L v_t)^2\right] - 2t^{\frac{d+2}{2}} \mathbb{E}_\rho[v_t(-\mathcal{L}v_t)] \,\mathrm{d}t. \end{aligned} \tag{5.20}$$

Step 3: Approximation by density. We apply a second approximation: we choose another scale $l > 0$, whose value will be fixed but $L/l \in \mathbb{N}$ and $l \simeq \sqrt{t_{n+1}}$. We denote by $q = (L/l)^d$ and $\mathbf{M}_{L,l} = (\mathbf{M}_1, \mathbf{M}_2 \cdots \mathbf{M}_q)$ a random vector, where $\mathbf{M}_i$ is the number of the particle in i-th cube of scale l. Then we define an operator

$$\mathsf{B}_{L,l} v_t := \mathbb{E}_\rho\left[v_t | \mathbf{M}_{L,l}\right].$$

The main idea here is that the random vector $\mathbf{M}_{L,l}$ captures the information of convergence, once we know the density in every cube of scale $l \simeq \sqrt{t_{n+1}}$ converges to ρ. In Proposition 5.5.1 we will prove a spectral inequality that

$$\mathbb{E}_\rho\left[(\mathsf{A}_L v_t - \mathsf{B}_{L,l} v_t)^2\right] \leqslant R_0 l^2 \mathbb{E}_\rho\left[v_t(-\mathcal{L}v_t)\right].$$

We put this estimate into eq. (5.20)

$$\begin{aligned} \Delta_n &\leqslant \mathbb{E}_\rho[(u_0)^2] + \int_{t_n}^{t_{n+1}} 2(d+2) t^{\frac{d}{2}} \mathbb{E}_\rho\left[\left(\mathsf{B}_{L,l} v_t\right)^2\right] + 2t^{\frac{d}{2}}((d+2)R_0 l^2 - t) \mathbb{E}_\rho[v_t(-\mathcal{L}v_t)] \,\mathrm{d}t \\ &\leqslant \mathbb{E}_\rho[(u_0)^2] + \int_{t_n}^{t_{n+1}} 2(d+2) t^{\frac{d}{2}} \mathbb{E}_\rho\left[\left(\mathsf{B}_{L,l} v_t\right)^2\right] \mathrm{d}t, \end{aligned}$$

where we obtain the last line by choosing a scale $l = c\sqrt{t_{n+1}}$ such that $(d+2)R_0 l^2 \leqslant t_n$ and $L/l \in \mathbb{N}$.

It remains to estimate how small $\mathbb{E}_\rho\left[\left(\mathsf{B}_{L,l} v_t\right)^2\right]$ is. The typical case is that the density is close to ρ in every cube of scale l in Q_L. Let us define $M = (M_1, M_2, \cdots M_q)$, and we have

$$\mathsf{B}_{L,l} v_t(M) = \mathbb{E}_\rho\left[v_t | \mathbf{M}_{L,l} = M\right].$$

Then we call $\mathcal{C}_{L,l,\rho,\delta}$ the δ-good configuration that

$$\mathcal{C}_{L,l,\rho,\delta} := \left\{ M \in \mathbb{N}^q \,\middle|\, \forall 1 \leqslant i \leqslant q, \left| \frac{M_i}{\rho |Q_l|} - 1 \right| \leqslant \delta \right\}. \tag{5.21}$$

We can use standard Chernoff bound and union bound to prove the upper bound of $\mathbb{P}_\rho\left[\mathbf{M}_{L,l} \notin \mathcal{C}_{L,l,\rho,\delta}\right]$: for any $\lambda > 0$, we have

$$\begin{aligned}
\mathbb{P}_\rho\left[\exists \leqslant i \leqslant q, \frac{M_i}{\rho|Q_l|} \geqslant 1+\delta\right] &\leqslant \left(\frac{L}{l}\right)^d \exp(-\lambda(1+\delta))\mathbb{E}_\rho\left[\exp\left(\frac{\lambda\mu(Q_l)}{\rho|Q_l|}\right)\right] \\
&= \left(\frac{L}{l}\right)^d \exp\left(-\lambda(1+\delta) + \rho|Q_l|\left(e^{\frac{\lambda}{\rho|Q_l|}} - 1\right)\right) \\
&\leqslant \left(\frac{L}{l}\right)^d \exp\left(-\lambda\delta + \frac{\lambda^2}{\rho|Q_l|}\right).
\end{aligned}$$

In the second line we use the exact Laplace transform for $\mu(Q_l)$ as we know $\mu(Q_l) \overset{\text{law}}{\sim} \text{Poisson}(\rho|Q_l|)$. Then we do optimization by choosing $\lambda = \frac{\delta\rho|Q_l|}{2}$. The other side is similar and we conclude

$$\mathbb{P}_\rho\left[\mathbf{M}_{L,l} \notin \mathcal{C}_{L,l,\rho,\delta}\right] \leqslant (\gamma \log(t_{n+1}))^d \exp\left(-\frac{\rho|Q_l|\delta^2}{4}\right). \tag{5.22}$$

For the case $M \notin \mathcal{C}_{L,l,\rho,\delta}$, we can bound $\mathsf{B}_{L,l} v_t(M)$ naively by $|\mathsf{B}_{L,l} v_t(M)| \leqslant C\|u_0\|_{\mathscr{L}^\infty}$, thus we have

$$\mathbb{E}_\rho\left[\left(\mathsf{B}_{L,l} v_t\right)^2\right] \leqslant \sum_{M \in \mathcal{C}_{L,l,\rho,\delta}} \mathbb{P}_\rho[\mathbf{M}_{L,l} = M](\mathsf{B}_{L,l} v_t(M))^2 + (\gamma \log(t_{n+1}))^d \exp\left(-\frac{\rho|Q_l|\delta^2}{4}\right) \|u_0\|^2_{\mathscr{L}^\infty}$$

and we finish this step by

$$\begin{aligned}
\Delta_n \leqslant &\mathbb{E}_\rho[(u_0)^2] + (t_{n+1})^{\frac{d+2}{2}} (\gamma \log(t_{n+1}))^d \exp\left(-\frac{\rho|Q_l|\delta^2}{4}\right) \|u_0\|^2_{\mathscr{L}^\infty} \\
&+ \sum_{M \in \mathcal{C}_{L,l,\rho,\delta}} \mathbb{P}_\rho[\mathbf{M}_{L,l} = M] \int_{t_n}^{t_{n+1}} 2(d+2) t^{\frac{d}{2}} (\mathsf{B}_{L,l} v_t(M))^2 \, \mathrm{d}t.
\end{aligned} \tag{5.23}$$

We remark that the parameter $\delta > 0$ will be fixed at the end of the proof.

Step 4: Perturbation estimate. It remains to estimate the term $(\mathsf{B}_{L,l} v_t(M))^2$ for the the δ-good configuration. Now we put the expression of v_t in and obtain

$$(\mathsf{B}_{L,l} v_t(M))^2 = \left(\frac{1}{|\mathcal{Z}_K|} \sum_{y \in \mathcal{Z}_K} (\mathsf{B}_{L,l}(u_t - \tau_y u_t))(M)\right)^2,$$

and our aim is to control

$$\int_{t_n}^{t_{n+1}} 2(d+2) t^{\frac{d}{2}} \left(\frac{1}{|\mathcal{Z}_K|} \sum_{y \in \mathcal{Z}_K} (\mathsf{B}_{L,l}(u_t - \tau_y u_t))(M)\right)^2 \mathrm{d}t. \tag{5.24}$$

To treat eq. (5.24), we calculate the Radon-Nikodym derivative that

$$g_M := \frac{\mathrm{d}\mathbb{P}_\rho[\cdot|\mathbf{M}_{L,l} = M]}{\mathrm{d}\mathbb{P}_\rho} = \frac{1}{\mathbb{P}_\rho[\mathbf{M}_{L,l} = M]]} \mathbf{1}_{\{\mathbf{M}_{L,l} = M]\}}. \tag{5.25}$$

Then we use the reversibility of the semigroup P_t and denote by $g_{M,t} := P_t g_M$

$$\mathsf{B}_{L,l}(u_t - \tau_y u_t)(M) = \mathbb{E}_\rho[g_M(u_t - \tau_y u_t)] = \mathbb{E}_\rho\left[g_{M,t}(u - \tau_y u)\right].$$

Then we would like to apply the a perturbation estimate Proposition 5.5.2 to control it: let $l_k := l_u + 2k$ then for any $|y| \leqslant k$, we have

$$\mathbb{E}_\rho[g_M(u_t - \tau_y u_t)] \leqslant C(d)(l_k \|u\|_{L^\infty})^2 \mathcal{E}_{Q_{l_k}}(\sqrt{g_M}),$$

where $\mathcal{E}_{Q_{l_k}}(\sqrt{g_M})$ is a localized Dirichlet form defined in eq. (5.59). A heuristic analysis of order is $\mathcal{E}_{Q_{l_k}}(\sqrt{g_M}) \simeq O\left((l_k)^d\right)$ since it is a Dirichlet form on Q_{l_k}. If we choose $k = K$ here to cover all the term, the bound will be of order $O(K^d)$, which is big when $K \simeq \sqrt{t} \geqslant l_u$. Therefore, we apply a coarse-graining argument: let $\overline{[0,y]}_k := \{z_i\}_{0 \leqslant i \leqslant n(y)}$ be a lattice path that of scale k, $z_0 = 0, z_{n(y)} = y, \{z_i\}_{1 \leqslant i < n(y)} \in (k\mathbb{Z})^d$ so the length of path is the shortest one. (See Figure 5.1 for illustration.) Then we have

$$(u - \tau_y u) = \sum_{i=0}^{n(y)-1} (\tau_{z_i} u - \tau_{z_{i+1}} u) = \sum_{i=0}^{n(y)-1} \tau_{z_i}(u - \tau_{h_{z_i}} u),$$

where $h_{z_i} = z_{i+1} - z_i$ the vector connecting the two and $|h_{z_i}| \leqslant k$. This expression with the transport invariant law of Poisson point process, Cauchy-Schwartz inequality implies

$$\begin{aligned}
\left(\mathsf{B}_{L,l}(u_t - \tau_y u_t)(M)\right)^2 &= \left(\sum_{z \in \overline{[0,y]}_k} \mathbb{E}_\rho\left[g_{M,t} \tau_z (u - \tau_{h_z} u)\right]\right)^2 \\
&= \left(\sum_{z \in \overline{[0,y]}_k} \mathbb{E}_\rho\left[(\tau_{-z} g_{M,t})\,(u - \tau_{h_z} u)\right]\right)^2 \\
&\leqslant C(d) n(y) \sum_{z \in \overline{[0,y]}_k} \left(\mathbb{E}_\rho\left[(\tau_{-z} g_{M,t})\,(u - \tau_{h_z} u)\right]\right)^2.
\end{aligned} \tag{5.26}$$

This term appears a perturbation estimate, which will be proved in Proposition 5.5.2 that

$$\begin{aligned}
\left(\mathbb{E}_\rho\left[(\tau_{-z} g_{M,t})\,(u - \tau_{h_z} u)\right]\right)^2 &\leqslant C(d)(l_k \|u\|_{L^\infty})^2 \mathcal{E}_{Q_{l_k}}\left(\sqrt{\tau_{-z} g_{M,t}}\right) \\
&= C(d)(l_k \|u\|_{L^\infty})^2 \mathcal{E}_{\tau_z Q_{l_k}}\left(\sqrt{g_{M,t}}\right),
\end{aligned}$$

where in the last step we use the transport invariant property of Poisson point process. Now we turn to the choice of the scale k. By the heuristic analysis that every $\mathcal{E}_{Q_{l_k}}$ contributes order $O((l_k)^d)$ and taking in account $n(y) \leqslant K/k$ we have in eq. (5.26)

$$\left(\mathsf{B}_{L,l}(u_t - \tau_y u_t)(M)\right)^2 \simeq O\left(\left(\frac{K}{k}\right)^2 (l_k)^{d+2}\right) \simeq O\left(\left(\frac{K}{k}\right)^2 (l_u + 2k)^{d+2}\right).$$

From this we see that a good scale should be $k = l_u$ so the term above is of order $O(K^2 (l_u)^d)$. We put these estimate back to eq. (5.24)

$$\text{eq. (5.24)} \leqslant \|u\|_{L^\infty}^2 \int_{t_n}^{t_{n+1}} 2(d+2) t^{\frac{d}{2}} K l_u \left(\frac{1}{|\mathcal{Z}_K|} \sum_{y \in \mathcal{Z}_K} \sum_{z \in \overline{[0,y]}_{l_u}} \mathcal{E}_{\tau_z Q_{3 l_u}}\left(\sqrt{g_{M,t}}\right)\right) \mathrm{d}t. \tag{5.27}$$

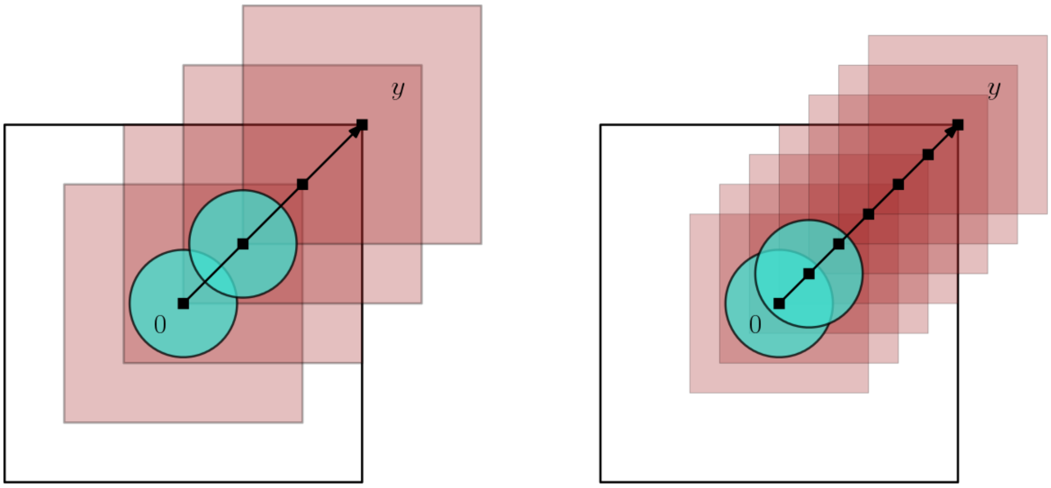

Figure 5.1: The illustration of the coarse-graining argument, where we take a lattice path of scale k to connect 0 and y. The ball in blue is the support of u and the box in red is Q_{l_k}. For the one on the left, the scale is $k = l_u$; the one on the right the scale is finer and we see that the coarse-graining is too dense.

Step 5: Covering argument. In this step, we calculate the right hand side of eq. (5.27), where we notice one essential problem: there are totally about K^{d+1}/l_u terms of Dirichlet form $\mathcal{E}_{\tau_z Q_{3l_u}}\left(\sqrt{g_{M,t}}\right)$ in the sum $\sum_{y\in\mathcal{Z}_K}\sum_{z\in\overline{[0,y]}_{l_u}}\mathcal{E}_{\tau_z Q_{l_u}}\left(\sqrt{g_{M,t}}\right)$, but the one with z close to 0 are counted of order K^d times, while the one with z near $\partial\mathcal{Z}_K$ are counted only constant times. To solve this problem, we have to reaverage the sum: by the transport invariant property of Poisson point process, at the beginning of the Step 1, we can write

$$\Delta_n = \frac{1}{|\mathcal{Z}_l|}\sum_{x\in\mathcal{Z}_l}\left((t_{n+1})^{\frac{d+2}{2}}\mathbb{E}_\rho[(\tau_x v_{t_{n+1}})^2] - (t_n)^{\frac{d+2}{2}}\mathbb{E}_\rho[(\tau_x v_{t_n})^2]\right).$$

Then all estimates works in Step 1, Step 2 and Step 3 work by replacing $v_t \mapsto \tau_x v_t$ and $u_t \mapsto \tau_x u_t$. In the Step 4, this operation will change our object term eq. (5.24)

$$\text{eq. (5.24)-avg} = \int_{t_n}^{t_{n+1}} 2(d+2)t^{\frac{d}{2}}\left(\frac{1}{|\mathcal{Z}_l|}\sum_{w\in\mathcal{Z}_l}\frac{1}{|\mathcal{Z}_K|}\sum_{y\in\mathcal{Z}_K}(\mathsf{B}_{L,l}\tau_w(u_t - \tau_y u_t)(M))^2\right)\mathrm{d}t,$$

and the perturbation argument Proposition 5.5.2 reduces the problem as

$$\begin{aligned}\text{eq. (5.24)-avg} \leqslant \|u\|^2_{\mathscr{L}^\infty}&\int_{t_n}^{t_{n+1}} 2(d+2)t^{\frac{d}{2}}Kl_u \\ &\times\left(\frac{1}{|\mathcal{Z}_l|}\sum_{w\in\mathcal{Z}_l}\frac{1}{|\mathcal{Z}_K|}\sum_{y\in\mathcal{Z}_K}\sum_{z\in\overline{[0,y]}_{l_u}}\mathcal{E}_{\tau_{w+z}Q_{3l_u}}\left(\sqrt{g_{M,t}}\right)\right)\mathrm{d}t.\end{aligned} \tag{5.28}$$

Now we can apply the Fubini's lemma

$$\frac{1}{|\mathcal{Z}_l|}\sum_{w\in\mathcal{Z}_l}\frac{1}{|\mathcal{Z}_K|}\sum_{y\in\mathcal{Z}_K}\sum_{z\in\overline{[0,y]}_{l_u}}\mathcal{E}_{\tau_{w+z}Q_{3l_u}}\left(\sqrt{g_{M,t}}\right)$$
$$=\frac{1}{|\mathcal{Z}_l|}\frac{1}{|\mathcal{Z}_K|}\mathbb{E}_\rho\left[\int_{\mathbb{R}^d}\left(\sum_{w\in\mathcal{Z}_l}\sum_{y\in\mathcal{Z}_K}\sum_{z\in\overline{[0,y]}_{l_u}}\mathbf{1}_{\{x\in\tau_{w+z}Q_{3l_u}\}}\right)\nabla\sqrt{g_{M,t}}(\mu,x)\cdot\nabla\sqrt{g_{M,t}}(\mu,x)\,\mathrm{d}\mu(x)\right],$$

while we notice that

$$\sum_{w\in\mathcal{Z}_l}\sum_{y\in\mathcal{Z}_K}\sum_{z\in\overline{[0,y]}_{l_u}}\mathbf{1}_{\{x\in\tau_{w+z}Q_{3l_u}\}}=\sum_{y\in\mathcal{Z}_K}\sum_{z\in\overline{[0,y]}_{l_u}}\underbrace{\sum_{w\in\mathcal{Z}_l}\mathbf{1}_{\{x-w\in\tau_z Q_{3l_u}\}}}_{\leqslant|Q_{3l_u}|}$$
$$\leqslant\sum_{y\in\mathcal{Z}_K}\sum_{z\in\overline{[0,y]}_{l_u}}(3l_u)^d\leqslant C(d)(l_u)^{d-1}K^{d+1},$$

so we have

$$\frac{1}{|\mathcal{Z}_l|}\sum_{w\in\mathcal{Z}_l}\frac{1}{|\mathcal{Z}_K|}\sum_{y\in\mathcal{Z}_K}\sum_{z\in\overline{[0,y]}_{l_u}}\mathcal{E}_{\tau_{w+z}Q_{3l_u}}\left(\sqrt{g_{M,t}}\right)\leqslant\frac{C(d)(l_u)^{d-1}K}{|\mathcal{Z}_l|}\mathcal{E}(\sqrt{g_{M,t}}).$$

We put this estimate to eq. (5.28) and use $l=c\sqrt{t_{n+1}}$,

$$\begin{aligned}\text{eq. (5.24)-avg}&\leqslant C(d)\|u\|^2_{\mathscr{L}^\infty}K^2(l_u)^d\int_{t_n}^{t_{n+1}}\left(\frac{t^{\frac{1}{2}}}{l}\right)^d\mathcal{E}(\sqrt{g_{M,t}})\,\mathrm{d}t\\&\leqslant C(d)\|u\|^2_{\mathscr{L}^\infty}K^2(l_u)^d\int_{t_n}^{t_{n+1}}\mathcal{E}(\sqrt{g_{M,t}})\,\mathrm{d}t.\end{aligned}\tag{5.29}$$

We put eq. (5.29) back to eq. (5.24) and eq. (5.23) and conclude

$$\begin{aligned}\Delta_n\leqslant&\mathbb{E}_\rho[(u_0)^2]+(t_{n+1})^{\frac{d+2}{2}}(\gamma\log(t_{n+1}))^d\exp\left(-\frac{\rho|Q_l|\delta^2}{4}\right)\|u_0\|^2_{\mathscr{L}^\infty}\\&+C(d)\|u\|^2_{\mathscr{L}^\infty}K^2(l_u)^d\sum_{M\in\mathcal{C}_{L,l,\rho,\delta}}\mathbb{P}_\rho[\mathbf{M}_{L,l}=M]\int_{t_n}^{t_{n+1}}\mathcal{E}(\sqrt{g_{M,t}})\,\mathrm{d}t.\end{aligned}\tag{5.30}$$

Step 6: Entropy inequality. In this step, we analyze the quantity $\int_{t_n}^{t_{n+1}}\mathcal{E}(\sqrt{g_{M,t}})\,\mathrm{d}t$. We recall the definition of the entropy inequality: let $H(g_M)=\mathbb{E}_\rho[g_M\log(g_M)]$, then

$$H(g_{M,t})=H(g_M)-4\int_0^t\mathbb{E}_\rho[\sqrt{g_{M,s}}(-\mathcal{L}\sqrt{g_{M,s}})]\,\mathrm{d}s,\tag{5.31}$$

we have

$$\int_{t_n}^{t_{n+1}}\mathcal{E}(\sqrt{g_{M,t}})\,\mathrm{d}t\leqslant\int_{t_n}^{t_{n+1}}\mathbb{E}_\rho[\sqrt{g_{M,t}}(-\mathcal{L}\sqrt{g_{M,t}})]\,\mathrm{d}t\leqslant H(g_{M,t_{n+1}})\leqslant H(g_M).$$

For any $M\in\mathcal{C}_{L,l,\rho,\delta}$, one can calculate the bound of the entropy and we prove it in Lemma 5.5.2

$$H(g_M)\leqslant C(d,\rho)\left(\frac{L}{l}\right)^d\left(\log(l)+l^d\delta^2\right).$$

This helps us conclude that

$$\Delta_n \leqslant \mathbb{E}_\rho[(u_0)^2] + \|u_0\|^2_{L^\infty} (\gamma \log(t_{n+1}))^d \left((t_{n+1})^{\frac{d+2}{2}} \exp\left(-\frac{\rho|Q_l|\delta^2}{4}\right) + K^2(l_u)^d \left(\log(l) + l^d\delta^2\right)\right).$$

To make the bound small, we choose a parameter $\delta = c(d,\rho)(\log t_{n+1})^{\frac{1}{2}}(t_{n+1})^{-\frac{d}{2}}$, where $c(d,\rho)$ is a positive number large enough to compensate the term $(t_{n+1})^{\frac{d+2}{2}}$ and this proves eq. (5.16)
□

5.4 Localization estimate

In this part, we prove the key localization estimate: we recall our notation of conditional expectation here that $\mathsf{A}_s f = \mathbb{E}\left[f|\mathcal{F}_{\overline{Q}_s}\right]$ for $\overline{Q}_s$ a closed cube $\left[-\frac{s}{2},\frac{s}{2}\right]^d$.

Theorem 5.4.1. *For $u \in \mathscr{L}^2$ which is $\mathcal{F}_{\overline{Q}_{l_u}}$-measurable, any $t \geqslant \max\left\{l_u^2, 16\Lambda^2\right\}$, $K \geqslant \sqrt{t}$, and u_t the function associated to the generator $\mathcal{L}$ at time t, then we have the estimate*

$$\mathbb{E}_\rho\left[(u_t - \mathsf{A}_K u_t)^2\right] \leqslant C(\Lambda)\exp\left(-\frac{K}{\sqrt{t}}\right)\mathbb{E}_\rho\left[u^2\right]. \tag{5.32}$$

This is an important inequality which allows us to pay some error to localize the function, and it is introduced in [142] and also used in [116]. The main idea to prove it is to use a multi-scale functional and analyze its evolution with respect to the time. Let us introduce its continuous version: for any $f \in \mathscr{L}^2$, $f \mapsto (\mathsf{A}_s f)_{s\geqslant 0}$ is a càdlàg $\mathscr{L}^2$-martingale with respect to $\left(\Omega, \left(\mathcal{F}_{\overline{Q}_s}\right)_{s\geqslant 0}, \mathbb{P}\right)$.

Our multi-scale functional for $f \in \mathscr{H}_0^1(\mathbb{R}^d)$ is defined as

$$S_{k,K,\beta}(f) = \alpha_k \mathbb{E}_\rho\left[(\mathsf{A}_k f)^2\right] + \int_k^K \alpha_s \, d\mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right] + \alpha_K \mathbb{E}_\rho\left[(f - \mathsf{A}_K f)^2\right], \tag{5.33}$$

with $\alpha_s = \exp\left(\frac{s}{\beta}\right), \beta > 0$. We can apply the integration by part formula for the Lebesgue-Stieltjes integral and obtain

$$S_{k,K,\beta}(f) = \alpha_K \mathbb{E}_\rho\left[f^2\right] - \int_k^K \alpha_s' \mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right] ds, \tag{5.34}$$

where α_s' is the derivative with respect to s. The main idea is to put u_t in eq. (5.34) and then study its derivative $\frac{d}{dt} S_{k,K,\beta}(u_t)$ and use it to prove Theorem 5.4.1. In this procedure, we will use the Dirichlet form for $\mathsf{A}_s u_t$, but we have to remark that in fact we do not know a priori this is a function in $\mathscr{H}_0^1(\mathbb{R}^d)$. We will give a counter example to make it clearer in the next section and introduce a regularized version of $\mathsf{A}_s f$ to pass this difficulty.

5.4.1 Conditional expectation, spatial martingale and its regularization

$(\mathsf{A}_s f)_{s\geqslant 0}$ has nice property: we can treat it as a localized function or a martingale. Thus we use the notation

$$\mathscr{M}_s^f := \mathsf{A}_s f, \tag{5.35}$$

which is a more canonical notation in martingale theory. In this subsection, we would like to understand the regularity of the closed martingale $\left(\mathscr{M}_s^f\right)_{s\geqslant 0}$. We will see it is a càdlàg $\mathscr{L}^2$-martingale and the jump happens when there is particles on the boundary ∂Q_s. At first, we remark a useful property for Poisson point process.

Lemma 5.4.1. *With probability* 1, *for any* $0 < s < \infty$, *there is at most one particle one the boundary* ∂Q_s.

Proof. We denote by

$$\mathcal{N} := \{\mu : \exists 0 < s < \infty, \text{ there exist more than two particles on } \partial Q_s\}.$$

Then we choose an increasing sequence $\{s_k^\varepsilon\}_{k\geqslant 0}$ with $s_0^\varepsilon = 0$, such that

$$\mathbb{R}^d = \bigsqcup_{k=1}^\infty C_{s_k^\varepsilon}, \qquad C_{s_k^\varepsilon} := Q_{s_k^\varepsilon} \backslash Q_{s_{k-1}^\varepsilon}, \qquad |C_{s_k^\varepsilon}| = \frac{\varepsilon}{k}.$$

Then we have that

$$\begin{aligned}
\mathbb{P}_\rho[\mathcal{N}] &\leqslant \mathbb{P}_\rho\left[\exists k, \mu(C_{s_k^\varepsilon}) \geqslant 2\right] \\
&\leqslant \sum_{k=1}^\infty \mathbb{P}_\rho\left[\mu(C_{s_k^\varepsilon}) \geqslant 2\right] \\
&\leqslant \sum_{k=1}^\infty \left(\rho|C_{s_k^\varepsilon}|\right)^2 \\
&\leqslant (\rho\varepsilon)^2.
\end{aligned}$$

We we let ε go down to 0 and prove that $\mathbb{P}_\rho[\mathcal{N}] = 0$. □

For this reason, in the following, we can do modification of the probability space and always suppose that there is at most one particle on the boundary. This helps us to prove the following regularity property for $\left(\mathscr{M}_s^f\right)_{s\geqslant 0}$.

Lemma 5.4.2. *After a modification, for any* $f \in \mathscr{L}^2$ *the process* $\left(\mathscr{M}_s^f\right)_{s\geqslant 0}$ *is a càdlàg* $\mathscr{L}^2$-*martingale, and the discontinuity point occurs for* s *such that* $\mu(\partial Q_s) = 1$.

Proof. By the classical martingale theory, we know that $\{\mathcal{F}_{\overline{Q}_s}\}_{s\geqslant 0}$ is a right continuous filtration, thus after a modification the process is càdlàg. Moreover, from Lemma 5.4.1 we can modify the value to 0 on a negligible set so that $\mu(\partial Q_s) \leqslant 1$ for all positive s. It remains to prove that if $\mu(\partial Q_s) = 0$, then the process is also left continuous. In this case, there exists a $0 < \varepsilon_0 < s$ such that for any $0 < \varepsilon < \varepsilon_0$, we have

$$\mu \llcorner \overline{Q}_s = \mu \llcorner \overline{Q}_{s-\varepsilon} = \mu \llcorner \overline{Q}_{s-\varepsilon_0}. \tag{5.36}$$

We use $\mu \llcorner \overline{Q}_s = (\mu \llcorner \overline{Q}_{s-\varepsilon}) + (\mu \llcorner (\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}))$, then $\mathsf{A}_{s-\varepsilon} f(\mu)$ has an expression

$$\begin{aligned}
\mathsf{A}_{s-\varepsilon} f(\mu) &= \int_{\mathcal{M}_\delta(\mathbb{R}^d)} \mathsf{A}_s f(\mu \llcorner \overline{Q}_{s-\varepsilon} + \mu' \llcorner (\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}))\, \mathrm{d}\mathbb{P}_\rho(\mu') \\
&= \int_{\mathcal{M}_\delta(\mathbb{R}^d)} \mathsf{A}_s f(\mu \llcorner \overline{Q}_{s-\varepsilon} + \mu' \llcorner (\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}))\mathbf{1}_{\{\mu'(\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon})=0\}}\, \mathrm{d}\mathbb{P}_\rho(\mu') \\
&\quad + \int_{\mathcal{M}_\delta(\mathbb{R}^d)} \mathsf{A}_s f(\mu \llcorner \overline{Q}_{s-\varepsilon} + \mu' \llcorner (\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}))\mathbf{1}_{\{\mu'(\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon})\geqslant 1\}}\, \mathrm{d}\mathbb{P}_\rho(\mu') \\
&= e^{-\rho|\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}|}\mathsf{A}_s f(\mu \llcorner \overline{Q}_{s-\varepsilon}) \\
&\quad + \int_{\mathcal{M}_\delta(\mathbb{R}^d)} \mathsf{A}_s f(\mu \llcorner \overline{Q}_{s-\varepsilon} + \mu' \llcorner (\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}))\mathbf{1}_{\{\mu'(\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon})\geqslant 1\}}\, \mathrm{d}\mathbb{P}_\rho(\mu').
\end{aligned}$$

This helps us estimate

$$
\begin{aligned}
&|\mathsf{A}_{s-\varepsilon} f(\mu) - \mathsf{A}_s f(\mu)| \\
&\leqslant \left(1 - e^{-\rho|\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}|}\right) |\mathsf{A}_s f(\mu \llcorner \overline{Q}_{s-\varepsilon})| \\
&\quad + \left(1 - e^{-\rho|\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}|}\right)^{\frac{1}{2}} \left(\int_{\mathcal{M}_\delta(\mathbb{R}^d)} |\mathsf{A}_s f(\mu \llcorner \overline{Q}_{s-\varepsilon} + \mu' \llcorner (\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}))|^2 \, \mathrm{d}\mathbb{P}_\rho(\mu')\right)^{\frac{1}{2}} \\
&\leqslant \left(1 - e^{-\rho|\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}|}\right) |\mathsf{A}_s f(\mu \llcorner \overline{Q}_{s-\varepsilon})| + \left(1 - e^{-\rho|\overline{Q}_s \backslash \overline{Q}_{s-\varepsilon}|}\right)^{\frac{1}{2}} \left(\mathsf{A}_{s-\varepsilon}(f^2)(\mu \llcorner \overline{Q}_{s-\varepsilon})\right)^{\frac{1}{2}},
\end{aligned}
$$

where we use Jensen's inequality at the last line. Notice that $f \in \mathscr{L}^2$, so $(\mathsf{A}_s(f^2))_{s\geqslant 0}$ is also a càdlàg martingale. Thus $\lim_{\varepsilon \searrow 0} \mathsf{A}_{s-\varepsilon}(f^2)$ admits a left limit $\mathsf{A}_{s-}(f^2)$. Moreover, we use eq. (5.36) and locally we have

$$
|\mathsf{A}_{s-\varepsilon} f(\mu) - \mathsf{A}_s f(\mu)| \leqslant C\rho\varepsilon s^{d-1} |\mathsf{A}_s f(\mu \llcorner \overline{Q}_{s-\varepsilon_0})| + C\rho^{\frac{1}{2}} \varepsilon^{\frac{1}{2}} s^{\frac{d-1}{2}} \left|\mathsf{A}_{s-}(f^2)(\mu \llcorner \overline{Q}_{s-\varepsilon_0})\right|^{\frac{1}{2}}.
$$

This implies that almost surely the càdlàg martingale $\left(\mathscr{M}_s^f\right)_{s\geqslant 0}$ is also continuous at s for $\mu(\partial Q_s) = 0$. □

The following corollaries are simple applications of the result above.

Corollary 5.4.1. *For $f \in \mathscr{L}^2$, we can define a bracket process for $\left(\mathscr{M}_s^f\right)_{s\geqslant 0}$: we define that*

$$
\left[\mathscr{M}^f\right]_s := \sum_{0<\tau\leqslant s} \left(\Delta \mathscr{M}_\tau^f\right)^2, \qquad \Delta \mathscr{M}_\tau^f = \mathscr{M}_\tau^f - \mathscr{M}_{\tau-}^f, \qquad \tau \text{ is jump point.} \tag{5.37}
$$

Then $\left(\left(\mathscr{M}_s^f\right)^2 - \left[\mathscr{M}^f\right]_s\right)_{s\geqslant 0}$ is a martingale with respect to $\left(\Omega, \left(\mathcal{F}_{\overline{Q}_s}\right)_{s\geqslant 0}, \mathbb{P}_\rho\right)$.

Proof. This is a direct result from jump process; see [141, Chapter 4e]. □

Corollary 5.4.2. *Let $x \in \operatorname{supp}(\mu)$, and we define a stopping time for x*

$$
\tau(x) := \min\{s \geqslant 0 | x \in Q_s\}, \tag{5.38}
$$

and the normal direction $\overrightarrow{\mathbf{n}}(x)$ and we define

$$
\mathsf{A}_{\tau(x)-} f(\mu - \delta_x + \delta_{x-}) := \lim_{\varepsilon \searrow 0} \mathsf{A}_{\tau(x)-\varepsilon} f(\mu - \delta_x + \delta_{x-\varepsilon \overrightarrow{\mathbf{n}}(x)}). \tag{5.39}
$$

Then for any $f \in \mathscr{L}^2$ we have almost surely

$$
\mathsf{A}_{\tau(x)-} f(\mu) = \mathsf{A}_{\tau(x)} f(\mu - \delta_x), \qquad \mathsf{A}_{\tau(x)-} f(\mu - \delta_x + \delta_{x-}) = \mathsf{A}_{\tau(x)} f(\mu). \tag{5.40}
$$

Proof. The equation $\mathsf{A}_{\tau(x)-} f(\mu) = \mathsf{A}_{\tau(x)} f(\mu - \delta_x)$ is the result of left continuous: from Lemma 5.4.1 we know with probability 1 there is only x on $\partial Q_{\tau(x)}$ and $\mu - \delta_x$ does not have particle on the boundary so we apply Lemma 5.4.2 and obtain this equation.

For the second equation, we have

$$
\begin{aligned}
\mathsf{A}_{\tau(x)} f(\mu) &= \lim_{\varepsilon_1 \searrow 0} \mathsf{A}_{\tau(x)} f(\mu - \delta_x + \delta_{x-\varepsilon_1 \overrightarrow{\mathbf{n}}(x)}) \\
&= \lim_{\varepsilon_1 \searrow 0} \mathsf{A}_{\tau(x)} f(\mu - \delta_x + \delta_{x-\varepsilon_1 \overrightarrow{\mathbf{n}}(x)}) \\
&= \lim_{\varepsilon_2 \searrow 0} \lim_{\varepsilon_1 \searrow \varepsilon_2} \mathsf{A}_{\tau(x)-\varepsilon_2} f(\mu - \delta_x + \delta_{x-\varepsilon_1 \overrightarrow{\mathbf{n}}(x)}) \\
&= \lim_{\varepsilon \searrow 0} \mathsf{A}_{\tau(x)-\varepsilon} f(\mu - \delta_x + \delta_{x-\varepsilon \overrightarrow{\mathbf{n}}(x)}).
\end{aligned}
$$

In the last step, we use the uniformly left continuous for $\mathsf{A}_s f$ and the continuity with respect to x. □

One important remark about the conditional expectation is that in fact for $f \in \mathscr{C}_c^\infty(\mathbb{R}^d)$, we may have $\mathsf{A}_L f \notin \mathscr{C}_c^\infty(\mathbb{R}^d)$. The reason is that the conditional expectation creates a small gap at the boundary for the function. Here we give an example of the conditional expectation for $\mathbb{E}_\rho[f|\mathcal{F}_{B_r}]$, which is easier to state but it shares the same property of $\mathsf{A}_L f$.

Example 5.4.1. *Let $\eta \in C_c^\infty(\mathbb{R}^d)$ be a plateau function:*

$$\mathrm{supp}(\eta) \subseteq B_1, 0 \leqslant \eta \leqslant 1, \eta \equiv 1 \textit{ in } B_{\frac{1}{2}}, \eta(x) = \eta(|x|) \textit{ decreasing with respect to } |x|.$$

and we define our function

$$f(\mu) = \left(\int_{\mathbb{R}^d} \eta(x)\, \mathrm{d}\mu(x)\right) \wedge 3.$$

We define the level set B_r such that

$$B_r := \left\{ x \in \mathbb{R}^d \,\middle|\, \frac{1}{2} \leqslant \eta(x) \leqslant 1 \right\}.$$

Then, we have $\mathbb{E}_\rho[f|\mathcal{F}_{B_r}] \notin \mathscr{C}_c^\infty(\mathbb{R}^d)$.

Proof. Let $\mu_1 = \mu \llcorner B_r, \mu_2 = \mu \llcorner (B_1 \backslash B_r)$, then since $\mathrm{supp}(f) \subseteq B_1$, we have that

$$\mathbb{E}_\rho[f|\mathcal{F}_{B_r}] = (\mu_1(\eta) + \mu_2(\eta)) \wedge 3.$$

Let us choose a specific configuration to see that $\mathbb{E}_\rho[f|\mathcal{F}_{B_r}](\mu)$ is not even continuous:

$$\mu_1 = \delta_{x_1} + \delta_{x_2} + \delta_{x_3}, \text{ where } x_1, x_2 \in B_{\frac{1}{2}}, x_3 \in B_r \backslash B_{\frac{1}{2}}.$$

Then we can calculate that $2.5 \leqslant \mu_1(\eta) < 3$ and $2.5 \leqslant \mathbb{E}_\rho[f|\mathcal{F}_{B_r}](\mu) < 3$. However, if we take another μ_1 that

$$\mu_1 = \delta_{x_1} + \delta_{x_2} + \delta_{x_3} + \delta_{x_4}, \text{ where } x_1, x_2 \in B_{\frac{1}{2}}, x_3 \in B_r \backslash B_{\frac{1}{2}}, x_4 \in B_r.$$

Then we see that $\mu_1(\eta) > 3$ and we have $\mathbb{E}_\rho[f|\mathcal{F}_{B_r}](\mu) = 3$. Therefore, once the 4-th particle x_4 enters the ball B_r, the value of the function will jump to 3. From this we conclude that $\mathbb{E}_\rho[f|\mathcal{F}_{B_r}] \notin \mathscr{C}_c^\infty(\mathbb{R}^d)$. □

To make the conditional expectation more regular, we introduce its regularized version: for any $0 < \varepsilon < \infty$, we define

$$\mathsf{A}_{s,\varepsilon} f := \frac{1}{\varepsilon} \int_0^\varepsilon \mathsf{A}_{s+t} f \, dt, \tag{5.41}$$

Then we have the following properties.

Proposition 5.4.1. *For any $f \in \mathscr{H}_0^1(\mathbb{R}^d)$, the function $\mathsf{A}_{s,\varepsilon} f \in \mathscr{H}_0^1(\mathbb{R}^d)$ and $\left(\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} f)^2\right]\right)_{s\geqslant 0}$ a C^1 increasing process.*

Proof. We calculate the formula for $\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} f)^2\right]$:

$$\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} f)^2\right] = \frac{1}{\varepsilon^2} \int_0^\varepsilon \int_0^\varepsilon \mathbb{E}_\rho\left[\mathsf{A}_{s+t_1} f \mathsf{A}_{s+t_2} f\right] \mathrm{d}t_1 \mathrm{d}t_2.$$

As we know that $\mathbb{E}_\rho\left[\mathsf{A}_{s+t_1} f \mathsf{A}_{s+t_2} f\right] = \mathbb{E}_\rho\left[(\mathsf{A}_{s+(t_1 \wedge t_2)} f)^2\right]$, we obtain that

$$\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} f)^2\right] = \frac{2}{\varepsilon^2}\int_0^\varepsilon (\varepsilon - t)\mathbb{E}_\rho\left[(\mathsf{A}_{s+t} f)^2\right] \mathrm{d}t. \tag{5.42}$$

Then we calculate its derivative that for $0 < h < \varepsilon$

$$\begin{aligned}
&\lim_{h \searrow 0} \frac{1}{h}\left(\mathbb{E}_\rho\left[(\mathsf{A}_{s+h,\varepsilon} f)^2\right] - \mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} f)^2\right]\right) \\
=& \lim_{h \searrow 0} \frac{2}{h\varepsilon^2}\left(\int_\varepsilon^{\varepsilon+h} (\varepsilon + h - t)\mathbb{E}_\rho\left[(\mathsf{A}_{s+t} f)^2\right] \mathrm{d}t - \int_0^h (\varepsilon - t)\mathbb{E}_\rho\left[(\mathsf{A}_{s+t} f)^2\right] \mathrm{d}t + \int_h^\varepsilon h\mathbb{E}_\rho\left[(\mathsf{A}_{s+t} f)^2\right] \mathrm{d}t\right) \\
=& \frac{2}{\varepsilon^2}\int_0^\varepsilon \mathbb{E}_\rho\left[(\mathsf{A}_{s+t} f)^2\right] - \mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right] \mathrm{d}t.
\end{aligned}$$

In the last step, we use the right continuity and this proves that

$$\frac{d}{ds}\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} f)^2\right] = \frac{2}{\varepsilon^2}\int_0^\varepsilon \mathbb{E}_\rho\left[(\mathsf{A}_{s+t} f)^2\right] - \mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right] \mathrm{d}t. \tag{5.43}$$

Then we calculate the partial derivative. We use the formula that

$$e_k \cdot \nabla \mathsf{A}_{s,\varepsilon} f(\mu, x) = \lim_{h \to 0} \frac{1}{h}\left(\frac{1}{\varepsilon}\int_0^\varepsilon \mathsf{A}_{s+t} f(\mu - \delta_x + \delta_{x+he_k}) - \mathsf{A}_{s+t} f(\mu)\, dt\right). \tag{5.44}$$

We study this derivative case by case.

1. Case $x \in \overline{Q}^c_{s+\varepsilon}$. In this case, in eq. (5.44), for a h small enough, for any $t \in [0, \varepsilon]$, neither x nor $x + he_k$ is in Q_{t+s}, so we have $\mathsf{A}_{s+t} f(\mu - \delta_x + \delta_{x+he_k}) = \mathsf{A}_{s+t} f(\mu \llcorner \overline{Q}_{s+t})$. This implies that eq. (5.44) is 0 in this case.

2. Case $x \in Q_s$. In this case, for a h small enough, for any $t \in [0, \varepsilon]$, both x and $x + he_k$ is in Q_{t+s}, then we have

$$\begin{aligned}
e_k \cdot \nabla \mathsf{A}_{s,\varepsilon} f(\mu, x) &= \lim_{h \to 0} \frac{1}{h}\left(\frac{1}{\varepsilon}\int_0^\varepsilon \mathsf{A}_{s+t} f(\mu - \delta_x + \delta_{x+he_k}) - \mathsf{A}_{s+t} f(\mu)\, dt\right) \\
&= \frac{1}{\varepsilon}\int_0^\varepsilon \lim_{h \to 0} \frac{1}{h}\left(\mathsf{A}_{s+t} f(\mu - \delta_x + \delta_{x+he_k}) - \mathsf{A}_{s+t} f(\mu)\right) dt \\
&= \mathsf{A}_{s,\varepsilon}\left(e_k \cdot \nabla f(\mu, x)\right).
\end{aligned}$$

3. Case $x \in \left(\overline{Q}_{s+\varepsilon} \backslash Q_s\right)$, e_k is the normal direction. In this case, we study at first the situation $\vec{\mathbf{n}}(x)$ and $h \searrow 0$. We divide eq. (5.44) in three terms:

$$\begin{aligned}
e_k \cdot \nabla \mathsf{A}_{s,\varepsilon} f(\mu, x) &= \mathbf{I} + \mathbf{II} + \mathbf{III} \\
\mathbf{I} &= \frac{1}{\varepsilon}\int_0^\varepsilon \mathbf{1}_{\{s+t<\tau(x)\}}\frac{1}{h}\left(\mathsf{A}_{s+t} f(\mu - \delta_x + \delta_{x+he_k}) - \mathsf{A}_{s+t} f(\mu)\right) dt \\
\mathbf{II} &= \frac{1}{\varepsilon}\int_0^\varepsilon \mathbf{1}_{\{s+t\geqslant\tau(x)+h\}}\frac{1}{h}\left(\mathsf{A}_{s+t} f(\mu - \delta_x + \delta_{x+he_k}) - \mathsf{A}_{s+t} f(\mu)\right) dt \\
\mathbf{III} &= \frac{1}{\varepsilon}\int_0^\varepsilon \mathbf{1}_{\{\tau(x)\leqslant s+t<\tau(x)+h\}}\frac{1}{h}\left(\mathsf{A}_{s+t} f(\mu - \delta_x + \delta_{x+he_k}) - \mathsf{A}_{s+t} f(\mu)\right) dt.
\end{aligned}$$

The term **I** and **II** are similar as we have discussed above and we have

$$\lim_{h \searrow 0} \mathbf{I} + \mathbf{II} = \frac{1}{\varepsilon}\int_0^\varepsilon \mathbf{1}_{\{s+t>\tau(x)\}}\mathsf{A}_{s+t}\left(e_k \cdot \nabla f(\mu, x)\right) dt.$$

For the term **III**, since $x + he_k \notin \overline{Q}_{s+t}$, we have $\mathsf{A}_{s+t}f(\mu - \delta_x + \delta_{x+he_k}) = \mathsf{A}_{s+t}f(\mu - \delta_x)$. Then, we use the right continuity of $\mathsf{A}_s f$

$$\begin{aligned}\lim_{h\searrow 0}\mathbf{III} &= \lim_{h\searrow 0}\frac{1}{h\varepsilon}\int_{\tau(x)-s}^{\tau(x)-s+h}\mathsf{A}_{s+t}f(\mu-\delta_x) - \mathsf{A}_{s+t}f(\mu)\,dt\\ &= \frac{1}{\varepsilon}\left(\mathsf{A}_{\tau(x)}f(\mu-\delta_x) - \mathsf{A}_{\tau(x)}f(\mu)\right).\end{aligned}$$

We should also remark that is is also the case we do partial derivative from left, in this case we should pay attention on the term **III** which is

$$\begin{aligned}\lim_{h\searrow 0}\mathbf{III}' &= \lim_{h\searrow 0}\frac{1}{h\varepsilon}\int_0^{\varepsilon}\mathbf{1}_{\{\tau(x)-h\leqslant s+t<\tau(x)\}}\left(\mathsf{A}_{s+t}f(\mu-\delta_x) - \mathsf{A}_{s+t}f(\mu-\delta_x+\delta_{x-he_k})\right)dt\\ &= \frac{1}{\varepsilon}\left(\mathsf{A}_{\tau(x)-}f(\mu-\delta_x) - \mathsf{A}_{\tau(x)-}f(\mu-\delta_x+\delta_{x-})\right).\end{aligned}$$

In the last step, we use the left continuity of $\mathsf{A}_{\tau(x)}f$ when the particle on the boundary is removed. Thanks to Corollary 5.4.2, we know this limit coincide with that of **III**. In conclusion, we could use the notation eq. (5.37)

$$\Delta\mathsf{A}_{\tau(x)}f = \mathsf{A}_{\tau(x)}f - \mathsf{A}_{\tau(x)-}f, \tag{5.45}$$

to unify the two. Thus we see it is nothing but the jump of the càdlàg martingale.

4. Case $x \in \left(\overline{Q}_{s+\varepsilon}\backslash Q_s\right)$, e_k is not the normal direction. This case is simpler than e_k is normal direction, where we do not have to consider the term **III** in the discussion above.

In summary, we obtain the formula that for any $x \in \operatorname{supp}(\mu)$

$$\nabla\mathsf{A}_{s,\varepsilon}f(\mu,x) = \begin{cases}\mathsf{A}_{s,\varepsilon}(\nabla f(\mu,x)) & x\in Q_s;\\ \frac{1}{\varepsilon}\int_{\tau(x)-s}^{\varepsilon}\mathsf{A}_{s+t}(\nabla f(\mu,x))\,dt - \frac{\vec{\mathbf{n}}(x)}{\varepsilon}\Delta\mathsf{A}_{\tau(x)}f & x\in\left(\overline{Q}_{s+\varepsilon}\backslash Q_s\right);\\ 0 & x\in\overline{Q}_{s+\varepsilon}^c.\end{cases} \tag{5.46}$$

Finally, we prove that $\mathsf{A}_{s,\varepsilon}f \in \mathscr{H}_0^1(\mathbb{R}^d)$. It is clear that $\mathsf{A}_{s,\varepsilon}f \in \mathscr{L}^2$ by Jensen's inequality for conditional expectation. For its gradient, we have

$$\begin{aligned}\mathbb{E}_\rho\left[\int_{\mathbb{R}^d}|\nabla\mathsf{A}_{s,\varepsilon}|^2\,\mathrm{d}\mu\right] \leqslant \mathbb{E}_\rho\left[\int_{Q_s}|\mathsf{A}_{s,\varepsilon}(\nabla f)|^2\,\mathrm{d}\mu\right] + 2\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\backslash Q_s}\left|\frac{1}{\varepsilon}\int_{\tau(x)-s}^{\varepsilon}\mathsf{A}_{s+t}(\nabla f)\,dt\right|^2\mathrm{d}\mu\right]\\ + \frac{2}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\backslash Q_s}|\Delta\mathsf{A}_{\tau(x)}f|^2\,\mathrm{d}\mu\right].\end{aligned}$$

For the first and second term in the equation above, we use Jensen's inequality for conditional expectation and Cauchy's inequality that

$$\begin{aligned}\mathbb{E}_\rho\left[\int_{Q_s}|\mathsf{A}_{s,\varepsilon}(\nabla f)|^2\,\mathrm{d}\mu\right] + 2\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\backslash Q_s}\left|\frac{1}{\varepsilon}\int_{\tau(x)-s}^{\varepsilon}\mathsf{A}_{s+t}(\nabla f)\,dt\right|^2\mathrm{d}\mu\right]\\ \leqslant \mathbb{E}_\rho\left[\int_{Q_s}|\nabla f|^2\,\mathrm{d}\mu\right] + \frac{2}{\varepsilon}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\backslash Q_s}|\nabla f|^2\,\mathrm{d}\mu\right].\end{aligned}$$

For the third term, it is in fact the sum of square of the jump part in the martingale $\left(\mathscr{M}_s^f\right)_{s\geqslant 0}$, so we use Corollary 5.4.1 that

$$\begin{aligned}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\backslash Q_s} |\Delta \mathsf{A}_{\tau(x)} f|^2 \, \mathrm{d}\mu\right] &= \mathbb{E}_\rho\left[\sum_{s\leqslant \tau\leqslant s+\varepsilon} |\Delta \mathscr{M}_\tau^f|^2\right] = \mathbb{E}_\rho\left[[\mathscr{M}^f]_{s+\varepsilon} - [\mathscr{M}^f]_s\right] \\ &= \mathbb{E}_\rho\left[\left(\mathscr{M}_{s+\varepsilon}^f\right)^2 - \left(\mathscr{M}_s^f\right)^2\right] = \mathbb{E}_\rho\left[(A_{s+\varepsilon} f)^2 - (A_s f)^2\right],\end{aligned}$$

where in the last step we also use the L^2 isometry for martingale. This concludes the desired result $\mathsf{A}_{s,\varepsilon} f \in \mathscr{H}_0^1(\mathbb{R}^d)$. □

5.4.2 Proof of Theorem 5.4.1

In this part, we prove Theorem 5.4.1 in three steps.

Proof. Step 1: Setting up. We propose a regularized multi-scale functional of eq. (5.33)

$$S_{k,K,\beta,\varepsilon}(f) = \alpha_k \mathbb{E}_\rho\left[(\mathsf{A}_{k,\varepsilon} f)^2\right] + \int_k^K \alpha_s \left(\frac{d}{ds}\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} f)^2\right]\right) ds + \alpha_K \mathbb{E}_\rho\left[f^2 - (\mathsf{A}_{K,\varepsilon} f)^2\right], \quad (5.47)$$

where we recall that $\alpha_s = \exp\left(\frac{s}{\beta}\right)$. The advantage is that $\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} f)^2\right]$ is C^1 for s from eq. (5.43), we can treat it as usual Riemann integral and apply integration by part to obtain an equivalent definition

$$S_{k,K,\beta,\varepsilon}(f) = \alpha_K \mathbb{E}_\rho\left[f^2\right] - \int_k^K \alpha_s' \mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} f)^2\right] ds. \quad (5.48)$$

Our object is to calculate $\frac{d}{dt} S_{k,K,\beta,\varepsilon}(u_t)$, and we pay attention to $\frac{d}{dt}\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} u_t)^2\right]$. We use the formula from eq. (5.42)

$$\begin{aligned}\frac{d}{dt}\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} u_t)^2\right] &= \frac{d}{dt}\frac{2}{\varepsilon^2}\int_0^\varepsilon (\varepsilon - r)\mathbb{E}_\rho\left[(\mathsf{A}_{s+r} u_t)^2\right] \mathrm{d}r \\ &= \frac{d}{dt}\frac{2}{\varepsilon^2}\int_0^\varepsilon (\varepsilon - r)\mathbb{E}_\rho\left[(\mathsf{A}_{s+r} u_t) u_t\right] \mathrm{d}r.\end{aligned}$$

We define that

$$\widetilde{\mathsf{A}_{s,\varepsilon}} f := \frac{2}{\varepsilon^2}\int_0^\varepsilon (\varepsilon - r)\mathsf{A}_{s+r} f \, \mathrm{d}r, \quad (5.49)$$

and it satisfies similar property as $\mathsf{A}_{s,\varepsilon} f$. For example, we have also the formula

$$\nabla \widetilde{\mathsf{A}_{s,\varepsilon}} f(\mu, x) = \begin{cases} \widetilde{\mathsf{A}_{s,\varepsilon}}\left(\nabla f(\mu,x)\right) & x \in Q_s; \\ \frac{2}{\varepsilon^2}\left(\int_{\tau(x)-s}^\varepsilon (\varepsilon - r)\mathsf{A}_{s+r}\left(\nabla f(\mu,x)\right) dr - (s+\varepsilon-\tau(x))\Delta \mathsf{A}_{\tau(x)} f \overrightarrow{\mathbf{n}}(x)\right) & x \in \left(\overline{Q}_{s+\varepsilon}\backslash Q_s\right); \\ 0 & x \in \overline{Q}_{s+\varepsilon}^c. \end{cases} \quad (5.50)$$

then we have

$$\frac{d}{dt}\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon} u_t)^2\right] = \frac{d}{dt}\mathbb{E}_\rho\left[\left(\widetilde{\mathsf{A}_{s,\varepsilon}} u_t\right) u_t\right] = \mathbb{E}_\rho\left[\left(\frac{d}{dt}\widetilde{\mathsf{A}_{s,\varepsilon}} u_t\right) u_t\right] + \mathbb{E}_\rho\left[\widetilde{\mathsf{A}_{s,\varepsilon}} u_t (\mathcal{L} u_t)\right]. \quad (5.51)$$

We study at first the semi-group. For a function $g \in \mathscr{H}_0^1(\mathbb{R}^d)$, we recall the definition that

$$g_t(\mu) = P_t g(\mu) := \mathbb{E}_\rho \left[g(\mu_t) | \mathscr{F}_0 \right].$$

We also know its semi-group that

$$\frac{d}{dt} P_t g(\mu) = \mathcal{L} P_t g(\mu) \Rightarrow \partial_t g_t(\mu) = \mathcal{L} g_t(\mu).$$

Now in our question we propose that $g = \widetilde{\mathsf{A}_{s,\varepsilon}} u_0$, then we have

$$\begin{aligned} g_t(\mu) &= P_t \left(\frac{2}{\varepsilon^2} \int_0^\varepsilon (\varepsilon - r) \mathbb{E}_\rho [u(\mu) | \mathcal{F}_{Q_{s+r}}] \, \mathrm{d}r \right) \\ &= \mathbb{E}_\rho \left[\left(\frac{2}{\varepsilon^2} \int_0^\varepsilon (\varepsilon - r) \mathbb{E}_\rho [u(\mu_t) | \mathcal{F}_{Q_{s+r}}] \, \mathrm{d}r \right) | \mathscr{F}_0 \right] \\ &= \frac{2}{\varepsilon^2} \int_0^\varepsilon (\varepsilon - r) \mathbb{E}_\rho \left[\mathbb{E}_\rho \left[u(\mu_t) \,|\mathscr{F}_0 \right] |\mathcal{F}_{Q_{s+r}} \right] \mathrm{d}r \\ &= \widetilde{\mathsf{A}_{s,\varepsilon}} u_t(\mu). \end{aligned}$$

Therefore, we have $\frac{d}{dt} \widetilde{\mathsf{A}_{s,\varepsilon}} u_t(\mu) = \mathcal{L} \widetilde{\mathsf{A}_{s,\varepsilon}} u_t(\mu)$ and put it back to eq. (5.51) and use reversibility to obtain that

$$\frac{d}{dt} \mathbb{E}_\rho \left[(\mathsf{A}_{s,\varepsilon} u_t)^2 \right] = 2 \mathbb{E}_\rho \left[\widetilde{\mathsf{A}_{s,\varepsilon}} u_t (\mathcal{L} u_t) \right].$$

We conclude that

$$\frac{d}{dt} S_{k,K,\beta,\varepsilon}(u_t) = 2\alpha_K \mathbb{E}_\rho \left[u_t (\mathcal{L} u_t) \right] + \int_k^K 2\alpha_s' \mathbb{E}_\rho \left[\widetilde{\mathsf{A}_{s,\varepsilon}} u_t (-\mathcal{L} u_t) \right] ds. \tag{5.52}$$

Step 2: Estimate of a localized Dirichlet energy. In this step, we will give an estimate for the term $\mathbb{E}_\rho \left[\widetilde{\mathsf{A}_{s,\varepsilon}} u_t (-\mathcal{L} u_t) \right]$ appeared in eq. (5.52). We will establish the following lemma.

Lemma 5.4.3. *For any $f \in \mathscr{H}_0^1(\mathbb{R}^d)$, we define that*

$$I_s^f := \mathbb{E}_\rho \left[\int_{Q_s} \nabla f \cdot \mathbf{a} \nabla f \, \mathrm{d}\mu \right], \tag{5.53}$$

then for $\widetilde{\mathsf{A}_{s,\varepsilon}} f$ introduced in eq. (5.49), for any $s, \theta, \varepsilon \in (0,\infty)$, we have

$$\mathbb{E}_\rho \left[\widetilde{\mathsf{A}_{s,\varepsilon}} f (-\mathcal{L} f) \right] \leqslant I_{s-1}^f + \Lambda \left(I_s^f - I_{s-1}^f \right) + \Lambda \left(\frac{\theta}{\varepsilon} + 1 \right) \left(I_{s+\varepsilon}^f - I_s^f \right) + \frac{\Lambda}{2\theta} \frac{d}{ds} \mathbb{E}_\rho \left[(\mathsf{A}_{s,\varepsilon} f)^2 \right]. \tag{5.54}$$

Proof. From eq. (5.50), we can decompose the quantity $\mathbb{E}_\rho \left[\widetilde{\mathsf{A}_{s,\varepsilon}} f (-\mathcal{L} f) \right]$ into three terms

$$\begin{aligned} \mathbb{E}_\rho \left[\widetilde{\mathsf{A}_{s,\varepsilon}} f (-\mathcal{L} f) \right] = &\underbrace{\mathbb{E}_\rho \left[\int_{Q_{s-1}} \nabla (\widetilde{\mathsf{A}_{s,\varepsilon}} f) \cdot \mathbf{a} \nabla f \, \mathrm{d}\mu \right]}_{\text{eq. (5.55)-a}} + \underbrace{\mathbb{E}_\rho \left[\int_{Q_s \setminus Q_{s-1}} \nabla (\widetilde{\mathsf{A}_{s,\varepsilon}} f) \cdot \mathbf{a} \nabla f \, \mathrm{d}\mu \right]}_{\text{eq. (5.55)-b}} \\ &+ \underbrace{\mathbb{E}_\rho \left[\int_{Q_{s+\varepsilon} \setminus Q_s} \nabla (\widetilde{\mathsf{A}_{s,\varepsilon}} f) \cdot \mathbf{a} \nabla f \, \mathrm{d}\mu \right]}_{\text{eq. (5.55)-c}}. \end{aligned} \tag{5.55}$$

For the first term eq. (5.55)-a, since $x \in Q_{s-1}$, then the coefficient is $\mathcal{F}_{Q_s}$ measurable. We use the formula eq. (5.50), eq. (5.49) and apply Jensen's inequality for conditional expectation

$$\begin{aligned}
\text{eq. (5.55)-a} &= \frac{2}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s-1}}\int_0^\varepsilon (\varepsilon - r)\mathsf{A}_{s+r}(\nabla f)\cdot \mathbf{a}\nabla f \,\mathrm{d}r\,\mathrm{d}\mu\right] \\
&= \frac{2}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s-1}}\int_0^\varepsilon (\varepsilon - r)\mathbb{E}_\rho\left[\mathsf{A}_{s+r}(\nabla f)\cdot \mathbf{a}\mathsf{A}_{s+r}(\nabla f)\,|\mathcal{F}_{Q_{s+r}}\right]\mathrm{d}r\,\mathrm{d}\mu\right] \\
&\leqslant \frac{2}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s-1}}\int_0^\varepsilon (\varepsilon - r)\mathbb{E}_\rho\left[\nabla f\cdot \mathbf{a}\nabla f\,|\mathcal{F}_{Q_{s+r}}\right]\mathrm{d}r\,\mathrm{d}\mu\right] \\
&= \mathbb{E}_\rho\left[\int_{Q_{s-1}} \nabla f\cdot \mathbf{a}\nabla f\,\mathrm{d}\mu\right]
\end{aligned}$$

For the second term eq. (5.55)-b, it is similar but $\mathbf{a}$ is no longer $\mathcal{F}_{Q_s}$ measurable. We use at first Young's inequality

$$\begin{aligned}
\text{eq. (5.55)-b} &\leqslant \frac{2}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-1}}\int_0^\varepsilon (\varepsilon - r)\mathsf{A}_{s+r}(\nabla f)\cdot \mathbf{a}\nabla f \,\mathrm{d}r\,\mathrm{d}\mu\right] \\
&\leqslant \frac{\Lambda}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-1}}\int_0^\varepsilon (\varepsilon - r)\left(|\mathsf{A}_{s+r}(\nabla f)|^2 + |\nabla f|^2\right)\mathrm{d}r\,\mathrm{d}\mu\right].
\end{aligned}$$

Then for the part with conditional expectation, we use the uniform bound $1 \leqslant \mathbf{a} \leqslant \Lambda$ that

$$\begin{aligned}
\frac{\Lambda}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-1}}\int_0^\varepsilon (\varepsilon - r)|\mathsf{A}_{s+r}(\nabla f)|^2\,\mathrm{d}r\,\mathrm{d}\mu\right] &\leqslant \frac{\Lambda}{2}\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-1}}|\nabla f|^2\mathrm{d}\mu\right] \\
&\leqslant \frac{\Lambda}{2}\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-1}}\nabla f\cdot\mathbf{a}\nabla f\mathrm{d}\mu\right].
\end{aligned}$$

This concludes that eq. (5.55)-b $\leqslant \Lambda\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-1}}\nabla f\cdot\mathbf{a}\nabla f\mathrm{d}\mu\right]$.

For the third term eq. (5.55)-c, we use eq. (5.50) and obtain

$$\begin{aligned}
&\text{eq. (5.55)-c} \leqslant \text{eq. (5.55)-c1} + \text{eq. (5.55)-c2} \\
&\text{eq. (5.55)-c1} = \frac{2}{\varepsilon^2}\left|\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}\int_{\tau(x)-s}^\varepsilon (\varepsilon - r)\mathsf{A}_{s+r}(\nabla f)\cdot\mathbf{a}\nabla f\,\mathrm{d}r\,\mathrm{d}\mu\right]\right| \\
&\text{eq. (5.55)-c2} = \frac{2}{\varepsilon^2}\left|\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))\Delta\mathsf{A}_{\tau(x)}f\,\vec{\mathbf{n}}(x)\cdot\mathbf{a}\nabla f\,\mathrm{d}\mu\right]\right|.
\end{aligned}$$

The part of eq. (5.55)-c1 is similar as that of eq. (5.55)-b and we have that

$$\text{eq. (5.55)-c1} \leqslant \Lambda\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}\nabla f\cdot\mathbf{a}\nabla f\mathrm{d}\mu\right].$$

We study the part eq. (5.55)-c2 with Young's inequality

$$\begin{aligned}
&\frac{2}{\varepsilon^2}\left|\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))\Delta\mathsf{A}_{\tau(x)}f\,\vec{\mathbf{n}}(x)\cdot\mathbf{a}\nabla f\,\mathrm{d}\mu\right]\right| \\
\leqslant &\frac{\Lambda}{\theta\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))|\Delta\mathsf{A}_{\tau(x)}f|^2\,\mathrm{d}\mu\right] + \frac{\theta\Lambda}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))|\nabla f|^2\,\mathrm{d}\mu\right] \\
\leqslant &\frac{\Lambda}{\theta\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))|\Delta\mathsf{A}_{\tau(x)}f|^2\,\mathrm{d}\mu\right] + \frac{\theta\Lambda}{\varepsilon}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}\nabla f\cdot\mathbf{a}\nabla f\,\mathrm{d}\mu\right].
\end{aligned}$$

The first part is in fact the bracket process defined in Corollary 5.4.1

$$\frac{\Lambda}{\theta\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))|\Delta\mathsf{A}_{\tau(x)}f|^2\,\mathrm{d}\mu\right]=\frac{\Lambda}{\theta\varepsilon^2}\mathbb{E}_\rho\left[\sum_{s\leqslant\tau\leqslant s+\varepsilon}(s+\varepsilon-\tau)|\Delta\mathscr{M}^f_\tau|^2\right].$$

Then we develop it with Fubini theorem and apply the $\mathscr{L}^2$-isometry of martingale that $\mathbb{E}_\rho\left[\left[\mathscr{M}^f\right]_s\right]=\mathbb{E}_\rho\left[(\mathscr{M}^f_s)^2\right]=\mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right]$

$$\begin{aligned}\frac{\Lambda}{\theta\varepsilon^2}\mathbb{E}_\rho\left[\sum_{s\leqslant\tau\leqslant s+\varepsilon}(s+\varepsilon-\tau)|\Delta\mathscr{M}^f_\tau|^2\right]&=\frac{\Lambda}{\theta\varepsilon^2}\mathbb{E}_\rho\left[\sum_{s\leqslant\tau\leqslant s+\varepsilon}\int_s^{s+\varepsilon}\mathbf{1}_{\{\tau\leqslant r\leqslant s+\varepsilon\}}\,\mathrm{d}r|\Delta\mathscr{M}^f_\tau|^2\right]\\&=\frac{\Lambda}{\theta\varepsilon^2}\mathbb{E}_\rho\left[\int_s^{s+\varepsilon}\sum_{s\leqslant\tau\leqslant r}|\Delta\mathscr{M}^f_\tau|^2\,\mathrm{d}r\right]\\&=\frac{\Lambda}{\theta\varepsilon^2}\mathbb{E}_\rho\left[\int_s^{s+\varepsilon}\left[\mathscr{M}^f\right]_r-\left[\mathscr{M}^f\right]_s\,\mathrm{d}r\right]\\&=\frac{\Lambda}{\theta\varepsilon^2}\int_0^\varepsilon\mathbb{E}_\rho\left[(\mathsf{A}_{s+r}f)^2\right]-\mathbb{E}_\rho\left[(\mathsf{A}_s f)^2\right]\mathrm{d}r\\&=\frac{\Lambda}{2\theta}\frac{d}{ds}\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon}f)^2\right].\end{aligned}$$

In the last step, we use the identity eq. (5.43). This concludes that

$$\text{eq. (5.55)-c}\leqslant\left(\frac{\theta\Lambda}{\varepsilon}+\Lambda\right)\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}\nabla f\cdot\mathbf{a}\nabla f\,\mathrm{d}\mu\right]+\frac{\Lambda}{2\theta}\frac{d}{ds}\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon}f)^2\right],$$

and we combine all the estimate for the three terms eq. (5.55)-a, eq. (5.55)-b, eq. (5.55)-c to obtain the desired result in eq. (5.54). □

Step 3: End of the proof. We take $k=\sqrt{t}, K>k$ and and put the estimate eq. (5.54) into eq. (5.52) with $\theta,\varepsilon,\beta>0$ to be fixed,

$$\begin{aligned}&\frac{d}{dt}S_{k,K,\beta,\varepsilon}(u_t)\\=&2\alpha_K\mathbb{E}_\rho\left[u_t(\mathcal{L}u_t)\right]+\int_k^K 2\alpha'_s\mathbb{E}_\rho\left[\widetilde{\mathsf{A}_{s,\varepsilon}}u_t(-\mathcal{L}u_t)\right]ds\\\leqslant&-2\alpha_K I^{u_t}_\infty+\int_k^K 2\alpha'_s\left\{I^{u_t}_{s-1}+\Lambda\left(I^{u_t}_s-I^{u_t}_{s-1}\right)+\Lambda\left(\frac{\theta}{\varepsilon}+1\right)(I^{u_t}_{s+\varepsilon}-I^{u_t}_s)+\frac{\Lambda}{2\theta}\frac{d}{ds}\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon}u_t)^2\right]\right\}\mathrm{d}s.\end{aligned}$$

We recall that $\alpha'_s=\frac{\alpha_s}{\beta}$, then we do some calculus and obtain that

$$\begin{aligned}\frac{d}{dt}S_{k,K,\beta,\varepsilon}(u_t)\leqslant&\int_{k-1}^{K+\varepsilon}\left(-2\alpha_{K\wedge(s+1)}+2\Lambda(\alpha_{s+1}-\alpha_s)+2\Lambda\left(\frac{\theta}{\varepsilon}+1\right)(\alpha_s-\alpha_{s-\varepsilon})\right)dI^{u_t}_s\\&+\int_0^{k-1}-2\alpha_k\,dI^{u_t}_s+\int_{K+\varepsilon}^\infty-2\alpha_K\,dI^{u_t}_s+\frac{\Lambda}{\beta\theta}\int_k^K\alpha_s\left(\frac{d}{ds}\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon}u_t)^2\right]\right)\mathrm{d}s.\end{aligned}$$

We see that the term $2\Lambda(\alpha_{s+1}-\alpha_s)\simeq\frac{2\Lambda}{\beta}\alpha_s$ and $2\Lambda\left(\frac{\theta}{\varepsilon}+1\right)(\alpha_s-\alpha_{s-\varepsilon})\simeq 2\Lambda\left(\frac{\theta}{\beta}+\frac{\varepsilon}{\beta}\right)\alpha_s$. One can choose the parameters $\theta=\frac{\beta}{2\Lambda}$, $\varepsilon=\frac{1}{2}$, then for $\beta>4\Lambda$, the part of integration with respect to $I^{u_t}_s$ is negative. We use the definition eq. (5.47) and obtain that

$$\frac{d}{dt}S_{k,K,\beta,\varepsilon}(u_t)\leqslant\frac{\Lambda}{\beta\theta}\int_k^K\alpha_s\left(\frac{d}{ds}\mathbb{E}_\rho\left[(\mathsf{A}_{s,\varepsilon}u_t)^2\right]\right)\mathrm{d}s\leqslant\frac{2\Lambda^2}{\beta^2}S_{k,K,\beta,\varepsilon}(u_t),$$

which implies that for $k = \sqrt{t} \geqslant l_u$, ($l_u$ the diameter of support of u_0 in Theorem 5.4.1)

$$\alpha_K \mathbb{E}_\rho\left[(u_t)^2 - (\mathsf{A}_{K,\varepsilon} u_t)^2\right] \leqslant S_{k,K,\beta,\varepsilon}(u_t) \leqslant \exp\left(\frac{2\Lambda^2 t}{\beta^2}\right) S_{k,K,\beta,\varepsilon}(u_0) = \exp\left(\frac{2\Lambda^2 t}{\beta^2}\right) \alpha_k \mathbb{E}_\rho\left[(u_0)^2\right].$$

Finally we remark that

$$\mathbb{E}_\rho\left[(u_t - \mathsf{A}_{K+\varepsilon} u_t)^2\right] = \mathbb{E}_\rho\left[(u_t)^2 - (\mathsf{A}_{K+\varepsilon} u_t)^2\right] \leqslant \mathbb{E}_\rho\left[(u_t)^2 - (\mathsf{A}_{K,\varepsilon} u_t)^2\right],$$

and choose $\beta = \sqrt{t}$ and it gives us the desired result, after shrinking a little the value of K. □

5.5 Spectral inequality, perturbation and perturbation

In this section, we collect several estimates used in the proof of the main result. They can also be read for independent interests.

5.5.1 Spectral inequality

The spectral inequality is an important topic in probability theory and Markov process, and it has its counterpart in analysis known as Poincaré's inequality.

Let $L > l > 0$ and $L/l \in \mathbb{N}$, and denote by $q = (L/l)^d, \{Q_l^i\}_{1 \leqslant i \leqslant q}$ the partition of Q_L by the small cube by scale l. Let $\mathbf{M}_{L,l} = (\mathbf{M}_1, \mathbf{M}_2 \cdots \mathbf{M}_q)$, be a random vector that $\mathbf{M}_i = \mu\left(Q_l^i\right)$, and we define $\mathsf{B}_{L,l} f := \mathbb{E}_\rho\left[f | \mathbf{M}_{L,l}\right]$, then we have the following estimate.

Proposition 5.5.1 (Spectral inequality). *There exists a finite positive number $R_0(d)$, such that for any $0 < l < L < \infty$, $L/l \in \mathbb{N}$, we have an estimate for any $f \in \mathscr{H}_0^1(\mathbb{R}^d)$,*

$$\mathbb{E}_\rho\left[(\mathsf{A}_L f - \mathsf{B}_{L,l} f)^2\right] \leqslant R_0 l^2 \mathbb{E}_\rho\left[\int_{Q_L} |\nabla f|^2 \, \mathrm{d}\mu\right]. \tag{5.56}$$

Proof. We prove at first a simple corollary from Efron-Stein inequality [60, Chapter 3]: let $f_n \in C^1(\mathbb{R}^{d \times n})$ and $X = (X_1, X_2 \cdots X_n)$, where $(X_i)_{1 \leqslant i \leqslant n}$ a family independent $\mathbb{R}^d$-valued random variables following uniform law in Q_l, then Efron-Stein inequality states

$$\mathbb{V}\mathrm{ar}\left[f_n(X)\right] \leqslant \frac{1}{2} \sum_{i=1}^n \mathbb{E}\left[\left(f_n(X) - f_n(X^i)\right)^2\right], \tag{5.57}$$

where $f_n(X^i) := \mathbb{E}\left[f_n(X) | X_1 \cdots X_{i-1}, X_{i+1}, \cdots X_n\right]$. From this, we calculate the expectation with respect to X_i for $\left(f_n(X) - f_n(X^i)\right)^2$, and apply the standard Poincaré's inequality for X_i

$$\begin{aligned}
\mathbb{E}_{X_i}\left[\left(f_n(X) - f(X^i)\right)^2\right] &= \fint_{Q_l} \left(f_n(x_1, x_2, \cdots x_n) - \fint_{Q_l} f_n(x_1, x_2, \cdots x_n) \, \mathrm{d}x_i\right)^2 \mathrm{d}x_i \\
&\leqslant C(d) l^2 \fint_{Q_l} |\nabla_{x_i} f_n|^2 (x_1, x_2, \cdots x_n) \, \mathrm{d}x_i, \\
\Longrightarrow \mathbb{E}\left[\left(f_n(X) - f(X^i)\right)^2\right] \leqslant C(d) l^2 \mathbb{E}[|\nabla_{x_i} f_n(X)|^2].
\end{aligned}$$

We combine the sum of all the term and obtain

$$\mathbb{V}\mathrm{ar}\left[f_n(X)\right] \leqslant C(d) l^2 \sum_{i=1}^n \mathbb{E}[|\nabla_{x_i} f_n(X)|^2]. \tag{5.58}$$

We then apply eq. (5.58) in eq. (5.56).

$$\mathbb{E}_\rho\left[(\mathsf{A}_L f - \mathsf{B}_{L,l} f)^2\right] = \sum_{M \in \mathbb{N}^q} \mathbb{P}_\rho[\mathbf{M}_{L,l} = M] \mathbb{E}_\rho\left[(\mathsf{A}_L f - \mathsf{B}_{L,l} f)^2 | \mathbf{M}_{L,l} = M\right].$$

Conditioned $\{\mathbf{M}_{L,l} = M\}$, we know that the expectation of $\mathsf{A}_L f$ is $\mathsf{B}_{L,l} f(M)$ and all the particles are distributed uniformly in its small cubes of size l, thus we can apply eq. (5.58) that

$$\begin{aligned}\mathbb{E}_\rho\left[(\mathsf{A}_L f - \mathsf{B}_{L,l} f)^2 | \mathbf{M}_{L,l} = M\right] &= \mathbb{V}\mathrm{ar}\rho\left[\mathsf{A}_L f | \mathbf{M}_{L,l} = M\right] \\ &\leqslant C(d) l^2 \mathbb{E}_\rho\left[\int_{Q_L} |\nabla \mathsf{A}_L f|^2 \, \mathrm{d}\mu \,\middle|\, \mathbf{M}_{L,l} = M\right] \\ &\leqslant C(d) l^2 \mathbb{E}_\rho\left[\int_{Q_L} |\mathsf{A}_L \nabla f|^2 \, \mathrm{d}\mu \,\middle|\, \mathbf{M}_{L,l} = M\right].\end{aligned}$$

Then we do the sum and concludes eq. (5.56). □

5.5.2 Perturbation

A similar version of the following lemma appears in [142], where the authors give some sketch and here we prove it in our model with some more details. We define a localized Dirichlet form for Borel set $U \subseteq \mathbb{R}^d$ that

$$\mathcal{E}_U(f, g) = \mathbb{E}_\rho[g(-\Delta_U f)] := \mathbb{E}_\rho\left[\int_U \nabla g(\mu, d) \cdot \nabla f(\mu, x) \, \mathrm{d}\mu(x)\right], \tag{5.59}$$

and we use $\mathcal{E}_U(f) := \mathcal{E}_U(f, f)$ and $\mathcal{E}(f) := \mathcal{E}_{\mathbb{R}^d}(f)$ for short.

Proposition 5.5.2 (Perturbation). *Let $u \in \mathscr{C}_c^\infty(\mathbb{R}^d)$ and $l_k := l_u + 2k$ be the minimal scale such that for any $|h| \leqslant k, \mathrm{supp}(\tau_h u) \subseteq Q_{l_k}$, then for any g such that $\mathbb{E}_\rho[g] = 1, \sqrt{g} \in \mathscr{H}_0^1(\mathbb{R}^d)$, we have*

$$(\mathbb{E}_\rho[g(u - \tau_h u)])^2 \leqslant C(d)(l_k \|u\|_{\mathscr{L}^\infty})^2 \mathcal{E}_{Q_{l_k}}(\sqrt{g}). \tag{5.60}$$

Proof. The proof of this proposition relies on the following lemma:

Lemma 5.5.1 (Lemma 4.2 of [142]). *Let $(\Omega, \mathbb{P}, \mathcal{F})$ be a probability space and let $\langle f, g\rangle = \int_\Omega f g \, \mathrm{d}\mathbb{P}$ denote the standard inner product on $L^2(\Omega, \mathbb{P}, \mathcal{F})$. Let A be a non-negative definite symmetric operator on $L^2(\Omega, \mathbb{P}, \mathcal{F})$, which has 0 as a simple eigenvalue with corresponding eigenfunction the constant function 1, and second eigenvalue $\delta > 0$ (the spectral gap). Let V be a function of means zero, $\langle 1, V\rangle = 0$ and assume that V is essential bounded. Denote by λ_ε the principal eigenvalue of $-A + \varepsilon V$ given by the variational formula*

$$\lambda_\varepsilon = \sup_{\|f\|_{L^2} = 1} \langle f, (-A + \varepsilon V) f\rangle. \tag{5.61}$$

Then for $0 < \varepsilon < \delta(2\|V\|_{L^\infty})^{-1}$,

$$0 \leqslant \lambda_\varepsilon \leqslant \frac{\varepsilon^2 \langle V, A^{-1} V\rangle}{1 - 2\|V\|_{L^\infty} \varepsilon \delta^{-1}}. \tag{5.62}$$

In our context, we should look for a good frame for this lemma. Since for any $|h| \leqslant k$, $(u - \tau_h u) \in \mathcal{F}_{Q_{l_k}}$, we have

$$\begin{aligned}\mathbb{E}_\rho[g(u-\tau_h u)] &= \mathbb{E}_\rho[(\mathsf{A}_{Q_{l_k}} g)(u-\tau_h u)] \\ &= \sum_{n=0}^{\infty} \mathbb{P}_\rho[\mu(Q_{l_k}) = n]\mathbb{E}_\rho[(\mathsf{A}_{Q_{l_k}} g)(u-\tau_h u)|\mu(Q_{l_k}) = n].\end{aligned} \tag{5.63}$$

Then, we focus on the estimate of $\mathbb{E}_\rho[(\mathsf{A}_{Q_{l_k}} g)(u-\tau_h u)|\mu(Q_{l_k}) = n]$: to shorten the notation, we use $\mathbb{P}_{\rho,n}$ for the probability $\mathbb{P}_\rho[\cdot|\mu(Q_{l_k}) = n]$ and $\mathbb{E}_{\rho,n}$ for its associated expectation. Then we apply Lemma 5.5.1 on the probability space $(\Omega, \mathcal{F}_{Q_{l_k}}, \mathbb{P}_{\rho,n})$, where we set $V = u - \tau_h u$ and the symmetric non-negative operator A is $(-\Delta_{Q_{l_k}})$ defined for any $f \in \mathscr{H}^1(Q_{l_k})$

$$\mathbb{E}_{\rho,n}[f(-\Delta_{Q_{l_k}} f)] := \mathbb{E}_{\rho,n}\left[\int_{Q_{l_k}} |\nabla f|^2 \, \mathrm{d}\mu\right].$$

We should check that this setting satisfies the condition of Lemma 5.5.1:

- Spectral gap for $A = -\Delta_{Q_{l_k}}$: by eq. (5.57) we have the spectral gap $\delta = (l_k)^{-2}$ for any function $f \in \mathscr{H}^1(Q_{l_k})$ with $\mathbb{E}_{\rho,n}[f] = 0$

$$\mathbb{E}_{\rho,n}[f^2] \leqslant (l_k)^2 \mathbb{E}_{\rho,n}[f(-\Delta_{Q_{l_k}} f)].$$

- Mean zero for $V = u - \tau_h u$: under the probability $\mathbb{P}_\rho$ this is clear by the transport invariant property of Poisson point process, while under $\mathbb{P}_{\rho,n}$ this requires some calculus. By the definition of l_k, we know that $\mathrm{supp}(u) \subseteq Q_{l_u}$, thus we denote by the projection $u(\mu) = \widetilde{u}_m(x_1, x_2, \cdots x_m)$ under the case $\mu \llcorner Q_{l_u} = \sum_{i=1}^m \delta_{x_i}$. Then we have

$$\begin{aligned}\mathbb{E}_{\rho,n}[u] &= \sum_{m=0}^{n} \mathbb{P}_{\rho,n}[\mu(Q_{l_u}) = m]\mathbb{E}_{\rho,n}[u|\mu(Q_{l_u}) = m] \\ &= \sum_{m=0}^{n} \binom{n}{m} \left(\frac{|Q_{l_u}|}{|Q_{l_k}|}\right)^m \left(1 - \frac{|Q_{l_u}|}{|Q_{l_k}|}\right)^{n-m} \fint_{(Q_{l_u})^m} \widetilde{u}_m(x_1, \cdots x_m) \, \mathrm{d}x_1 \cdots \mathrm{d}x_m,\end{aligned}$$

  because under $\mathbb{P}_{\rho,n}$, the number of particles in Q_{l_u} follows the law $\mathrm{Bin}\left(n, \frac{|Q_{l_u}|}{|Q_{l_k}|}\right)$ and they are uniformly distributed conditioned the number. We use the similar argument for the expectation of $\tau_h u$, where we should study the case for particles in $\tau_{-h} Q_{l_u} \subseteq Q_{l_k}$

$$\begin{aligned}\mathbb{E}_{\rho,n}[\tau_h u] &= \sum_{m=0}^{n} \mathbb{P}_{\rho,n}[\mu(\tau_{-h} Q_{l_u}) = m]\mathbb{E}_{\rho,n}[\tau_h u|\mu(\tau_{-h} Q_{l_u}) = m] \\ &= \sum_{m=0}^{n} \binom{n}{m} \left(\frac{|\tau_{-h} Q_{l_u}|}{|Q_{l_k}|}\right)^m \left(1 - \frac{|\tau_{-h} Q_{l_u}|}{|Q_{l_k}|}\right)^{n-m} \\ &\qquad \times \fint_{(\tau_{-h} Q_{l_u})^m} \widetilde{u}_m(x_1 + h, \cdots x_m + h) \, \mathrm{d}x_1 \cdots \mathrm{d}x_m \\ &= \sum_{m=0}^{n} \binom{n}{m} \left(\frac{|Q_{l_u}|}{|Q_{l_k}|}\right)^m \left(1 - \frac{|Q_{l_u}|}{|Q_{l_k}|}\right)^{n-m} \fint_{(Q_{l_u})^m} \widetilde{u}_m(x_1, \cdots x_m) \, \mathrm{d}x_1 \cdots \mathrm{d}x_m.\end{aligned}$$

  Thus we establish $\mathbb{E}_{\rho,n}[\tau_h u] = \mathbb{E}_{\rho,n}[u]$ and V has mean zero.

Now we can apply the lemma: for any $0 < \varepsilon < \frac{1}{8}(\|u\|_{L^\infty}(l_k)^2)^{-1}$, we put $\sqrt{\mathsf{A}_{Q_{l_k}} g}/\mathbb{E}_{\rho,n}[\mathsf{A}_{Q_{l_k}} g]$ at the place of f in eq. (5.61) and combine with eq. (5.62) to obtain that

$$\begin{aligned}\mathbb{E}_{\rho,n}\left[\mathsf{A}_{Q_{l_k}} g(u-\tau_h u)\right] \leqslant 2\varepsilon \mathbb{E}_{\rho,n}[(u-\tau_h u)((-\Delta_{Q_{l_k}})^{-1}(u-\tau_h u))]\mathbb{E}_{\rho,n}[\mathsf{A}_{Q_{l_k}} g] \\ + \frac{1}{\varepsilon}\mathbb{E}_{\rho,n}\left[\sqrt{\mathsf{A}_{Q_{l_k}} g}\left((-\Delta_{Q_{l_k}})\sqrt{\mathsf{A}_{Q_{l_k}} g}\right)\right].\end{aligned}$$

Notice that $(-\Delta_{Q_{l_k}})^{-1} : L^2 \to H^1$ well-defined thanks to the Lax-Milgram theorem and the spectral bound, we get

$$\begin{aligned}&\mathbb{E}_{\rho,n}\left[\mathsf{A}_{Q_{l_k}} g(u-\tau_h u)\right] \\ &\qquad \leqslant 8\varepsilon(l_k)^2\|u\|_{L^\infty}^2 \mathbb{E}_{\rho,n}[\mathsf{A}_{Q_{l_k}} g] + \frac{1}{\varepsilon}\mathbb{E}_{\rho,n}\left[\sqrt{\mathsf{A}_{Q_{l_k}} g}\left((-\Delta_{Q_{l_k}})\sqrt{\mathsf{A}_{Q_{l_k}} g}\right)\right]. \quad (5.64)\end{aligned}$$

For the case $\varepsilon > \frac{1}{8}(\|u\|_{L^\infty}(l_k)^2)^{-1}$, we have $1 \leqslant 8\varepsilon\|u\|_{L^\infty}(l_k)^2$, thus we use a trivial bound

$$\mathbb{E}_{\rho,n}\left[\mathsf{A}_{Q_{l_k}} g(u-\tau_h u)\right] \leqslant 2\|u\|_{L^\infty}\mathbb{E}_{\rho,n}\left[\mathsf{A}_{Q_{l_k}} g\right] \leqslant 16\varepsilon(l_k)^2\|u\|_{L^\infty}^2\left[\mathsf{A}_{Q_{l_k}} g\right]. \quad (5.65)$$

We combine eq. (5.64), eq. (5.65) and do optimization with for ε to obtain that

$$\mathbb{E}_{\rho,n}\left[\mathsf{A}_{Q_{l_k}} g(u-\tau_h u)\right] \leqslant 4 l_k\|u\|_{L^\infty}\left(\mathbb{E}_{\rho,n}\left[\mathsf{A}_{Q_{l_k}} g\right]\mathbb{E}_{\rho,n}\left[\sqrt{\mathsf{A}_{Q_{l_k}} g}\left((-\Delta_{Q_{l_k}})\sqrt{\mathsf{A}_{Q_{l_k}} g}\right)\right]\right)^{\frac{1}{2}}.$$

Here the term $\mathbb{E}_{\rho,n}\left[\sqrt{\mathsf{A}_{Q_{l_k}} g}\left((-\Delta_{Q_{l_k}})\sqrt{\mathsf{A}_{Q_{l_k}} g}\right)\right]$ is not the desired term and we should remove the conditional expectation here. For any $x \in Q_{l_k}$, using Cauchy-Schwartz inequality we have

$$\mathsf{A}_{Q_{l_k}}\left(\frac{|\nabla g(\mu,x)|^2}{g(\mu)}\right)\mathsf{A}_{Q_{l_k}} g(\mu) \geqslant \left(\mathsf{A}_{Q_{l_k}}|\nabla g(\mu,x)|\right)^2 \geqslant \left|\mathsf{A}_{Q_{l_k}}\nabla g(\mu,x)\right|^2.$$

Thus, in the term $\mathbb{E}_{\rho,n}\left[\sqrt{\mathsf{A}_{Q_{l_k}} g}\left((-\Delta_{Q_{l_k}})\sqrt{\mathsf{A}_{Q_{l_k}} g}\right)\right]$ we have

$$\begin{aligned}\mathbb{E}_{\rho,n}\left[\sqrt{\mathsf{A}_{Q_{l_k}} g}\left((-\Delta_{Q_{l_k}})\sqrt{\mathsf{A}_{Q_{l_k}} g}\right)\right] &= \frac{1}{4}\mathbb{E}_{\rho,n}\left[\int_{Q_{l_k}}\frac{\left|\mathsf{A}_{Q_{l_k}}\nabla g(\mu,x)\right|^2}{\mathsf{A}_{Q_{l_k}} g(\mu)}\,\mathrm{d}\mu\right] \\ &\leqslant \frac{1}{4}\mathbb{E}_{\rho,n}\left[\int_{Q_{l_k}}\mathsf{A}_{Q_{l_k}}\left(\frac{|\nabla g(\mu,x)|^2}{g(\mu)}\right)\mathrm{d}\mu\right] \\ &= \mathbb{E}_{\rho,n}\left[\sqrt{g}\left((-\Delta_{Q_{l_k}})\sqrt{g}\right)\right].\end{aligned}$$

Using the transpose invariant property for μ, we obtain

$$\left|\mathbb{E}_{\rho,n}\left[\mathsf{A}_{Q_{l_k}} g(u-\tau_h u)\right]\right| \leqslant 4 l_k\|u\|_{L^\infty}\left(\mathbb{E}_{\rho,n}\left[\mathsf{A}_{Q_{l_k}} g\right]\mathbb{E}_{\rho,n}\left[\sqrt{g}\left((-\Delta_{Q_{l_k}})\sqrt{g}\right)\right]\right)^{\frac{1}{2}},$$

and put it back to eq. (5.63) and use Cauchy-Schwartz inequality

$$
\begin{aligned}
&(\mathbb{E}_\rho[g(u-\tau_h u)])^2 \\
=&(l_k\|u\|_{L^\infty})^2\left(\sum_{n=0}^{\infty}\mathbb{P}_\rho[\mu(Q_{l_k})=n]\left(\mathbb{E}_{\rho,n}\left[\mathsf{A}_{Q_{l_k}}g\right]\mathbb{E}_{\rho,n}\left[\sqrt{g}\left((-\Delta_{Q_{l_k}})\sqrt{g}\right)\right]\right)^{\frac{1}{2}}\right)^2 \\
\leqslant&(l_k\|u\|_{L^\infty})^2\underbrace{\left(\sum_{n=0}^{\infty}\mathbb{P}_\rho[\mu(Q_{l_k})=n]\mathbb{E}_{\rho,n}\left[\mathsf{A}_{Q_{l_k}}g\right]\right)}_{=\mathbb{E}_\rho[\mathsf{A}_{Q_{l_k}}g]=\mathbb{E}_\rho[g]=1}\left(\sum_{n=0}^{\infty}\mathbb{P}_\rho[\mu(Q_{l_k})=n]\mathbb{E}_{\rho,n}\left[\sqrt{g}\left((-\Delta_{Q_{l_k}})\sqrt{g}\right)\right]\right) \\
=&(l_k\|u\|_{L^\infty})^2\mathbb{E}_\rho[\sqrt{g}(-\Delta_{Q_{l_k}}\sqrt{g})].
\end{aligned}
$$

□

5.5.3 Entropy

We recall the definition of δ-good configuration for $\frac{L}{l}\in\mathbb{N}$

$$\mathcal{C}_{L,l,\rho,\delta}=\left\{M\in\mathbb{N}^{\left(\frac{L}{l}\right)^d}\,\middle|\,\forall 1\leqslant i\leqslant\left(\frac{L}{l}\right)^d,\left|\frac{M_i}{\rho|Q_l|}-1\right|\leqslant\delta\right\}.$$

Lemma 5.5.2 (Bound for entropy)**.** *Given $l\geqslant 1,\frac{L}{l}\in\mathbb{N}$, $0<\delta<\frac{\rho}{2}$ for any $M\in\mathcal{C}_{L,l,\rho,\delta}$, we have a bound for the entropy of g_M defined in eq. (5.25) that*

$$H(g_M)\leqslant C(d,\rho)\left(\frac{L}{l}\right)^d\left(\log(l)+l^d\delta^2\right). \tag{5.66}$$

Proof.

$$H(g_M)=\mathbb{E}_\rho[g_M\log(g_M)]=-\mathbb{E}_\rho[g_M\log(\mathbb{P}_\rho[\mathbf{M}_{L,l}=M])].$$

It suffices to prove a upper bound for $-\log(\mathbb{P}_\rho[\mathbf{M}_{L,l}=M])$, which is

$$-\log(\mathbb{P}_\rho[\mathbf{M}_{L,l}=M])=-\log\left(\prod_{i=1}^{\left(\frac{L}{l}\right)^d}e^{-\rho|Q_l|}\frac{(\rho|Q_l|)^{M_i}}{M_i!}\right)=\sum_{i=1}^{\left(\frac{L}{l}\right)^d}-\log\left(e^{-\rho|Q_l|}\frac{(\rho|Q_l|)^{M_i}}{M_i!}\right). \tag{5.67}$$

For every term M_i, we set $\delta_i:=\frac{M_i}{\rho|Q_l|}-1$, and use Stirling's formula upper bound $n!\leqslant e\sqrt{n}\left(\frac{n}{e}\right)^n$ for any $n\in\mathbb{N}$

$$
\begin{aligned}
-\log\left(e^{-\rho|Q_l|}\frac{(\rho|Q_l|)^{M_i}}{M_i!}\right)&=\rho|Q_l|-M_i\log(\rho|Q_l|)+\log(M_i!) \\
&\leqslant\rho|Q_l|-M_i\log(\rho|Q_l|)+\log\left(e\sqrt{M_i}\left(\frac{M_i}{e}\right)^{M_i}\right) \\
&\leqslant\rho|Q_l|\left(\frac{M_i}{\rho|Q_l|}\log\left(\frac{M_i}{\rho|Q_l|}\right)+1-\frac{M_i}{\rho|Q_l|}\right)+\frac{1}{2}\log(M_i) \\
&=\rho|Q_l|\left((1+\delta_i)\underbrace{\log(1+\delta_i)}_{\leqslant\delta_i}-\delta_i\right)+\frac{1}{2}\underbrace{\log(M_i)}_{\leqslant C\log(l)} \\
&\leqslant\rho|Q_l|(\delta_i)^2+C\log(l).
\end{aligned}
$$

We use $|\delta_i|\leqslant\delta$ and put it back to eq. (5.67) and obtain the desired result. □

Chapter 6

Quantitative homogenization of interacting particle systems

For a class of interacting particle systems in continuous space, we show that finite-volume approximations of the bulk diffusion matrix converge at an algebraic rate. The models we consider are reversible with respect to the Poisson measures with constant density, and are of non-gradient type. Our approach is inspired by recent progress in the quantitative homogenization of elliptic equations. Along the way, we develop suitable modifications of the Caccioppoli and multiscale Poincaré inequalities, which are of independent interest.

This chapter corresponds to the article [115] written in collaboration with Arianna Giunti and Jean-Christophe Mourrat.

6.1 Introduction

The goal of this paper is to make progress on the quantitative analysis of interacting particle systems. We consider a class of models in which each particle follows a random evolution on $\mathbb{R}^d$ which is influenced by the configuration of neighboring particles. The models we consider are reversible with respect to the Poisson measures with constant density, uniformly elliptic, and of non-gradient type. For similar models in this class, the hydrodynamic limit and the equilibrium fluctuations have been identified rigorously. In both these results, the limit object is described in terms of the *bulk diffusion matrix*. The main result of this paper is a proof that finite-volume approximations of this diffusion matrix converge at an algebraic rate.

Our strategy is inspired by recent developments in the quantitative analysis of elliptic equations with random coefficients, and in particular on the renormalization approach developed in [31, 30, 23, 24, 25, 20, 21]; see also [185] for a gentle introduction, and [188, 182, 123, 124, 121, 125, 122] for another approach based on concentration inequalities. This renormalization approach has shown its versatility in a number of other settings, covering now the homogenization of parabolic equations [18], finite-difference equations on percolation clusters [19, 83, 85], differential forms [84], the "$\nabla\phi$" interface model [82, 29], and the Villain model [86].

Here as in the other settings mentioned above, we start from a representation of the finite-volume approximation of the bulk diffusion matrix as a family of variational problems, denoted by $\nu(U,p)$, where $U \subseteq \mathbb{R}^d$ and $p \in \mathbb{R}^d$ encodes a slope parameter. This quantity is subadditive as a function of the domain U. We then identify another subadditive quantity, denoted by $\nu^*(U,q)$, with $U \subseteq \mathbb{R}^d$ and $q \in \mathbb{R}^d$, such that $\nu^*(U,\cdot)$ is approximately convex dual to $\nu(U,\cdot)$. These quantities ν and ν^* provide with finite-volume lower and upper approximations of the limit diffusion matrix. Roughly speaking, the algebraic rate of convergence is obtained by showing that the defect in the convex duality between ν and ν^* can be controlled by the variation of ν and ν^* between two scales; we refer to [185, Section 3] for some intuition as to why a control of this sort is plausible.

Besides the identification of the most appropriate subadditive quantities ν and ν^*, one of the main difficulties we encounter relates to the development of certain functional inequalities. As is to be expected, we will make use of Poincaré inequalities, which allow to control the L^2 oscillation of a function by the L^2 norm of its gradient. However, we will need to be more precise than this. Indeed, we want to be able to assert that if the gradient of a function is small in some weaker norm, then we can control the L^2 oscillation of the function more tightly. In other words, we need some analogue of the inequality $\|u\|_{L^2} \leqslant C\|\nabla u\|_{H^{-1}}$. Recall that in the current paper, the functions of interest are defined over the space of all possible particle configurations. The precise statement of our "multiscale Poincaré inequality" is in Proposition 6.3.5.

Another crucial ingredient we need is a version of the Caccioppoli inequality. In the standard setting of elliptic equations, this inequality states that the L^2 norm of the gradient of a harmonic function can be controlled by the L^2 norm of its oscillation on a larger domain; one can think of this inequality as a "reverse Poincaré inequality" for harmonic functions. If u denotes the harmonic function, then a standard proof of this inequality consists in testing the equation for u with $u\phi$, where ϕ is a smooth cutoff function which is equal to 1 in the inner domain, and is equal to 0 outside of the larger domain.

In our context, we need to "turn off" the influence of *any* particle that would come too close to the boundary of the larger domain. In this case, a naive modification of the standard elliptic argument is inapplicable. This comes from the fact that, as the domains become large, there will essentially always be many particles that come dangerously close to the boundary of the larger domain; so the cutoff function ϕ would essentially always have to vanish, except on an event of very small probability. We therefore need to identify a different approach. In fact, we settle for a modified form of the Caccioppoli inequality, in which we control the L^2 norm of the gradient of a solution by the L^2 norm of the solution on a larger domain, plus *a fraction* of the L^2 norm of its gradient on the larger domain; see Proposition 6.3.6 for the precise statement.

At present, we think that the results presented here should allow to derive a quantitative version of the hydrodynamic limit, as well as to derive "near-equilibrium" fluctuation results. To be precise, for a domain of side length R and an initial density profile varying macroscopically, it should be possible to control the convergence to the hydrodynamic limit at a precision of $R^{-\alpha}$, for some $\alpha > 0$. Conversely, starting from a density profile that has variations of size bounded by $R^{-\frac{d}{2}+\alpha}$, it should be possible to identify the asymptotic fluctuations of the density field. These would represent first steps towards bridging the gap between these two results.

By analogy with the results obtained for elliptic equations and other contexts, see in particular [185, Section 3] and [25, Chapter 2 and following], we hope that the results obtained here will provide the seed for more refined, and hopefully sharp, quantitative results. This will hopefully allow to improve the exponent $\alpha > 0$ appearing in the previous paragraph to some explicit exponent (ideally $\alpha = \frac{d}{2}$), and thereby to bring us closer to a full understanding of non-equilibrium fluctuations.

We now turn to a brief overview of related works on interacting particle systems. The result in the literature that is possibly closest to ours is that of [166]. In this work, the authors consider the diffusion matrix associated with the long-time behavior of a tagged particle in the symmetric simple exclusion process, which is called the *self-diffusion matrix*. The main result of [166] is a proof that finite-volume approximations of the self-diffusion matrix converge to the correct limit. However, no rate of convergence could be obtained there. The qualitative result of [166] was extended to the mean-zero simple exclusion process, and to the asymmetric simple exclusion process in dimension $d \geqslant 3$, in [143].

An easy consequence of the results of the present paper is that the bulk diffusion matrix is Hölder continuous as a function of the density of particles. However, for related models, it was shown in [222, 165, 51, 167, 217, 189, 190, 191] that the diffusion matrix depends smoothly on the density of particles. The situation seems comparable to that encountered when considering Bernoulli perturbations of the law of the coefficient field for elliptic equations, see [183, 96]. Possibly more difficult situations for obtaining regularity results on the homogenized parameters, with less independence built into the nature of the perturbation, include the $\nabla\phi$ model [29], and nonlinear elliptic equations [20, 21].

Two classical approaches to the identification of the hydrodynamic limit have been developed. The first, called the entropy method, was introduced in [136], and extended to certain non-gradient models in [221, 204]. The second, called the relative entropy method, was introduced in [224], and was extended to a non-gradient model in [111].

The asymptotic description of the fluctuations of interacting particle systems at equilibrium has been obtained in [66, 214, 91, 69, 71]. The extension of this result to non-gradient models was obtained in [174, 70, 110].

We are not aware of any results concerning the non-equilibrium fluctuations of a non-gradient model. For gradient models (or small perturbations thereof), we refer in particular to [202, 90, 106, 71, 144]. We also refer to the books [215, 152, 157] for much more thorough expositions on these topics, and reviews of the literature.

In relation to the purposes of the present paper, several works considered the problem of obtaining a rate of convergence to equilibrium for a system of interacting particles [173, 93, 52, 142, 168, 68, 135]. Heat kernel bounds for the tagged particle in a simple exclusion process were obtained in [116].

In all likelihood, the results presented here can be extended to other reversible models of non-gradient type, provided that the invariant measures satisfy some mixing condition (an algebraic decay of correlations would suffice, see [30]). More challenging directions include dynamics that are not uniformly elliptic, such as hard spheres. Extensions to situations in which the noise only acts on the velocity variable are likely to also be very challenging. Even further away are purely deterministic dynamics of hard spheres, as considered for instance in [59]. For any of these models, it would of course also be desirable to make progress on the quantitative analysis of the large-scale behavior of a tagged particle.

The rest of the paper is organized as follows. In Section 6.2, we introduce some notation and state the main result precisely, see Theorem 6.2.1. We then prove several functional inequalities in Section 6.3, including the multiscale Poincaré inequality and the modified Caccioppoli inequality. In Section 6.4, we define the subadditive quantities, and establish their elementary properties. Finally, in Section 6.5 we prove Theorem 6.2.1.

6.2 Notation and main result

In this section, we introduce some notation and state our main result.

Let $\mathcal{M}_\delta(\mathbb{R}^d)$ be the set of σ-finite measures that are sums of Dirac masses on $\mathbb{R}^d$, which we think of as the space of configurations of particles. We denote by $\mathbb{P}_\rho$ the law on $\mathcal{M}_\delta(\mathbb{R}^d)$ of the Poisson point process of density $\rho \in (0,\infty)$, with $\mathbb{E}_\rho$ the associated expectation. We denote by $\mathcal{F}_U$ the σ-algebra generated by the mappings $V \mapsto \mu(V)$, for all Borel sets $V \subseteq U$, completed with all the $\mathbb{P}_\rho$-null sets, and we set $\mathcal{F} := \mathcal{F}_{\mathbb{R}^d}$. We give ourselves a function $\mathbf{a}_\circ : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}^{d\times d}_{\mathrm{sym}}$, where $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ is the set of d-by-d symmetric matrices. We assume that this mapping satisfies the following properties:

- *uniform ellipticity*: there exists $\Lambda < \infty$ such that for every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$,

$$\forall \xi \in \mathbb{R}^d, \quad |\xi|^2 \leqslant \xi \cdot \mathbf{a}_\circ(\mu)\xi \leqslant \Lambda |\xi|^2 \, ; \tag{6.1}$$

- *finite range of dependence*: denoting by B_1 the Euclidean ball of radius 1 centered at the origin, we assume that $\mathbf{a}_\circ$ is $\mathcal{F}_{B_1}$-measurable.

We denote by $\tau_{-x}\mu$ the translation of the measure μ by the vector $-x \in \mathbb{R}^d$; explicitly, for every Borel set U, we have $(\tau_{-x}\mu)(U) = \mu(x+U)$. We extend $\mathbf{a}_\circ$ by stationarity by setting, for every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$ and $x \in \mathbb{R}^d$,

$$\mathbf{a}(\mu, x) := \mathbf{a}_\circ(\tau_{-x}\mu).$$

While it would be possible to provide with a direct definition of the asymptotic bulk diffusion matrix, see for instance [152, Chapter 7], our purposes require that we identify suitable finite-volume versions of this quantity. Accordingly, for every bounded open set $U \subseteq \mathbb{R}^d$, we define the matrix $\bar{\mathbf{a}}(U) \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ to be such that, for every $p \in \mathbb{R}^d$,

$$\frac{1}{2} p \cdot \bar{\mathbf{a}}(U) p = \inf_{\phi \in \mathscr{H}^1_0(U)} \mathbb{E}_\rho \left[\frac{1}{\rho |U|} \int_U \frac{1}{2} (p + \nabla \phi(\mu, x)) \cdot \mathbf{a}(\mu, x) (p + \nabla \phi(\mu, x)) \, \mathrm{d}\mu(x) \right]. \tag{6.2}$$

In this expression, the gradient $\nabla\phi(\mu,x)$ is such that, for any sufficiently smooth function ϕ, $x \in \operatorname{supp}\mu$, and $k \in \{1,\dots,d\}$,

$$e_k \cdot \nabla\phi(\mu,x) = \lim_{h\to 0} \frac{\phi(\mu - \delta_x + \delta_{x+he_k}) - \phi(\mu)}{h}, \tag{6.3}$$

with $(e_1,\dots,e_d)$ being the canonical basis of $\mathbb{R}^d$. As will be explained in more details below, the space $\mathscr{H}_0^1(U)$ is a completion of a space of functions that are $\mathcal{F}_K$-measurable for some compact set $K \subseteq U$. The expectation $\mathbb{E}_\rho$ is taken with respect to the variable μ, a notation we will always use to denote the canonical random variable on $(\mathcal{M}_\delta(\mathbb{R}^d), \mathcal{F}, \mathbb{P}_\rho)$ (an explicit writing of $\int_U \cdots \mathrm{d}\mu(x)$ would actually involve a summation over every point in the intersection of U and the support of μ). For every $m \in \mathbb{N}$, we let $\square_m = Q_{3^m}$ denote the cube of side length 3^m. We define the *bulk diffusion matrix* $\bar{\mathbf{a}}$ as

$$\bar{\mathbf{a}} := \lim_{m\to\infty} \bar{\mathbf{a}}(\square_m). \tag{6.4}$$

Although we keep this implicit in the notation, we point out that the matrices $\bar{\mathbf{a}}(U)$ and $\bar{\mathbf{a}}$ depend on the density ρ of particles, which we keep fixed throughout the paper. Our main result is to obtain an algebraic rate for the convergence in (6.4).

Theorem 6.2.1. *The limit in (6.4) is well-defined. Moreover, there exist an exponent $\alpha(d,\Lambda,\rho) > 0$ and a constant $C(d,\Lambda,\rho) < \infty$ such that for every $m \in \mathbb{N}$,*

$$|\bar{\mathbf{a}}(\square_m) - \bar{\mathbf{a}}| \leqslant C 3^{-\alpha m}. \tag{6.5}$$

In the remainder of this section, we clarify some of the definitions appearing earlier, and introduce some more useful notation.

6.2.1 Continuum configuration space

For the purposes of the present paper, we will not need to construct the stochastic process of interacting particles whose large-scale behavior is captured by the bulk diffusion matrix $\bar{\mathbf{a}}$, so we contend ourselves with brief remarks here. Intuitively, the dynamics is a cloud of particles, which we can denote by

$$\mu(t) = \sum_{i=1}^{\infty} \delta_{X_i(t)} \in \mathcal{M}_\delta(\mathbb{R}^d), \qquad t \geqslant 0,$$

and each coordinate $(X_i(t))_{t\geqslant 0}$ performs a diffusion with local diffusivity matrix given by $\mathbf{a}(\mu(t), X_i(t))$. General properties of interacting diffusions on the space $\mathcal{M}_\delta(\mathbb{R}^d)$ have been studied using Dirichlet forms in [2, 3, 4, 5]; see also the survey [206]. In our current setup, for a finite N number of particles, the diffusion process can be defined in the standard way (say, using De Giorgi-Nash regularity results on the heat kernel, and Kolmogorov's theorems) as a diffusion on $(\mathbb{R}^d)^N$. For $\mathbb{P}_\rho$-almost every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$, one can then define the dynamics of the entire cloud of particles using finite-volume approximations.

Although we have defined $\mathbf{a}(\mu,x)$ for every $x \in \mathbb{R}^d$, we will in fact only need to appeal to this quantity in the case when x is in the support of μ. One possible example of local diffusivity function is $\mathbf{a}_\circ(\mu) := (1 + \mathbf{1}_{\{\{\mu(B_1)=1\}\}})\mathsf{Id}$. For this example, a particle at position $x \in \mathbb{R}^d$ follows a Brownian motion with variance 2 whenever there are no other particles in the unit ball centered at x, while it follows a Brownian motion with unit variance whenever there is at least one additional particle in this ball (there are also reflection effects at the transition between these two situations).

For every Borel set $U \subseteq \mathbb{R}^d$, we denote by $\mathcal{B}_U$ the set of Borel subsets of U. For every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$, we denote by $\operatorname{supp}\mu$ the support of μ, and by $\mu \llcorner U \in \mathcal{M}_\delta(\mathbb{R}^d)$ the measure such that, for every Borel set $V \subseteq \mathbb{R}^d$,

$$(\mu \llcorner U)(V) = \mu(U \cap V).$$

We will often use the following "disintegration" lemma for functions defined on $\mathcal{M}_\delta(\mathbb{R}^d)$. For definiteness, we state it for functions taking values in $\mathbb{R}$, but this plays no particular role. Its proof is deferred to Appendix 6.A. Whenever $U \subseteq \mathbb{R}^d$, we write U^c to denote the complement of U in $\mathbb{R}^d$.

Lemma 6.2.1 (Canonical projection). *Let $f : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}$ be a function, and for every Borel set U, measure $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$, and $n \in \mathbb{N}$, let $f_n(\cdot, \mu \llcorner U^c)$ denote the (permutation-invariant) function*

$$f_n(\cdot, \mu \llcorner U^c) : \begin{cases} U^n & \to & \mathbb{R} \\ (x_1, \ldots, x_n) & \mapsto & f\left(\sum_{i=1}^n \delta_{x_i} + \mu \llcorner U^c\right). \end{cases}$$

The following statements are equivalent.

(1) The function f is $\mathcal{F}$-measurable.

(2) For every $n \in \mathbb{N}$, the function f_n is $\mathcal{B}_U^{\otimes n} \otimes \mathcal{F}_{U^c}$-measurable.

Abusing notation, on the event that $\mu \llcorner U = \sum_{i=1}^n \delta_{x_i}$, we may sometimes write $f_n(\mu) = f_n(\sum_{i=1}^n \delta_{x_i} + \mu \llcorner U^c)$, so that $f(\mu) = \sum_{n=0}^\infty f_n(\mu)\, \mathbf{1}_{\{\{\mu(U)=n\}\}}$.

6.2.2 Lebesgue and Sobolev function spaces

We define $\mathscr{L}^2$ to be the space of $\mathcal{F}$-measurable functions f such that $\mathbb{E}_\rho[f^2]$ is finite.

Recall that for sufficiently smooth $f : \mathcal{M}_\delta(\mathbb{R}^d) \to \mathbb{R}$, $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$ and $x \in \operatorname{supp}\mu$, we define $\nabla f(\mu, x)$ according to the formula in (6.3). We write $\nabla f = (\partial_1 f, \cdots, \partial_d f)$.

For every open set $U \subseteq \mathbb{R}^d$, we define the sets of smooth functions $\mathscr{C}^\infty(U)$ and $\mathscr{C}_c^\infty(U)$ in the following way. We have that $f \in \mathscr{C}^\infty(U)$ if and only if f is an $\mathcal{F}$-measurable function, and for every bounded open set $V \subseteq U$, $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$ and $n \in \mathbb{N}$, the function $f_n(\cdot, \mu \llcorner V^c)$ appearing in Lemma 6.2.1 is infinitely differentiable on V^n. The space $\mathscr{C}_c^\infty(U)$ is the subspace of $\mathscr{C}^\infty(U)$ of functions that are $\mathcal{F}_K$-measurable for some compact set $K \subseteq U$.

We now define $\mathscr{H}^1(U)$, an infinite dimensional analogue of the classical Sobolev space H^1. For every $f \in \mathscr{C}^\infty(U)$, we set

$$\|f\|_{\mathscr{H}^1(U)} = \left(\mathbb{E}_\rho[f^2(\mu)] + \mathbb{E}_\rho\left[\int_U |\nabla f(\mu, x)|^2 \, d\mu(x)\right]\right)^{\frac{1}{2}}.$$

The space $\mathscr{H}^1(U)$ is the completion, with respect to this norm, of the space of functions $f \in \mathscr{C}^\infty(U)$ such that $\|f\|_{\mathscr{H}^1(U)}$ is finite (elements in this function space that coincide $\mathbb{P}_\rho$-almost surely are identified). As in classical Sobolev spaces, for every $f \in \mathscr{H}^1(U)$, we can interpret $\nabla f(\mu, x)$, with $x \in U$, in some weak sense. We stress that *functions in $\mathscr{H}^1(U)$ need not be $\mathcal{F}_U$-measurable.* Indeed, the function f can depend on $\mu \llcorner U^c$ in a relatively arbitrary (measurable) way, as long as $f \in \mathscr{L}^2$. If $V \subseteq U$ is another open set, then $\mathscr{H}^1(U) \subseteq \mathscr{H}^1(V)$.

We also define the space $\mathscr{H}_0^1(U)$ as the closure in $\mathscr{H}^1(U)$ of the space of functions $f \in \mathscr{C}_c^\infty(U)$ such that $\|f\|_{\mathscr{H}^1(U)}$ is finite. Notice in particular that, in stark contrast with functions in $\mathscr{H}^1(U)$, a function in $\mathscr{H}^1_0(U)$ does not depend on $\mu \llcorner U^c$. In the notation of Lemma 6.2.1, when $f \in \mathscr{H}^1_0(U)$, certain compatibility conditions between the functions $(f_n)_{n\in\mathbb{N}}$ also have to be satisfied. If $V \subseteq U$ is another open set, we have that $\mathscr{H}^1_0(V) \subseteq \mathscr{H}^1_0(U)$ (notice that the inclusion is in the opposite direction to that for $\mathscr{H}^1$ spaces). We also have the following result.

Lemma 6.2.2. *For every bounded open set $U \subseteq \mathbb{R}^d$ with Lipschitz boundary and $f \in \mathscr{H}^1_0(U)$, we have*

$$\mathbb{E}_\rho\left[\int_U \nabla f(\mu, x)\,\mathrm{d}\mu(x)\right] = 0. \tag{6.6}$$

Proof. By density, we can assume that $f \in \mathscr{C}^\infty_c(U)$. We use the functions $(f_n)_{n\in\mathbb{N}}$ appearing in Lemma 6.2.1; moreover, since $f(\mu)$ does not depend on $\mu \llcorner U^c$, we simply write $f_n(x_1, \ldots, x_n)$ in place of $f_n(x_1, \ldots, x_n, \mu \llcorner U^c)$. For every $k \in \{1, \ldots, d\}$, we have

$$\mathbb{E}_\rho\left[\int_U \partial_k f(\mu, x)\,\mathrm{d}\mu(x)\right] = \sum_{n=1}^{\infty} \mathbb{P}_\rho[\mu(U) = n] \sum_{i=1}^{n} \fint_{U^n} e_k \cdot \nabla_{x_i} f_n(x_1, \cdots, x_n)\,\mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

We use Green's formula for the integral $e_k \cdot \nabla_{x_i} f(x_1, \cdots, x_n)$ with respect to x_i

$$\int_U e_k \cdot \nabla_{x_i} f_n(x_1, \cdots, x_n)\,\mathrm{d}x_i = \int_{\partial U} f_n(x_1, \cdots, x_n) e_k \cdot \mathbf{n}(x_i)\,\mathrm{d}x_i,$$

where $\mathbf{n}(x_i)$ is the unit outer normal. Since $f \in \mathscr{C}^\infty_c(U)$, the quantity $f_n(x_1, \ldots, x_n)$ remains constant when x_i moves along the boundary ∂U. Denoting this constant (which depends on $(x_j)_{j\neq i}$) by c, we apply once again Green's formula to get

$$\int_U e_k \cdot \nabla_{x_i} f_n(x_1, \cdots, x_n)\,\mathrm{d}x_i = \int_{\partial U} c e_k \cdot \mathbf{n}(x_i)\,\mathrm{d}x_i = \int_U e_k \cdot \nabla_{x_i} c\,\mathrm{d}x_i = 0.$$

This proves the desired result. □

6.2.3 Localization operators

We now introduce families of operators that allow to localize a function defined on $\mathcal{M}_\delta(\mathbb{R}^d)$. We state some properties of these operators without proof, and refer to [135, Section 4.1](see Section 5.4.1) for more details.

Recall that for every $s > 0$, we write by $Q_s := \left(-\frac{s}{2}, \frac{s}{2}\right)^d$. We denote the closure of the cube Q_s by $\overline{Q}_s$, and define $\mathsf{A}_s f := \mathbb{E}_\rho[f \,|\, \mathcal{F}_{\overline{Q}_s}]$. For any $f \in \mathscr{L}^2$, the process $(\mathsf{A}_s f)_{s\geqslant 0}$ is a càdlàg $\mathscr{L}^2$-martingale with respect to $(\mathcal{M}_\delta(\mathbb{R}^d), (\mathcal{F}_{\overline{Q}_s})_{s\geqslant 0}, \mathbb{P}_\rho)$. We denote the jump at time s by

$$\Delta_s(\mathsf{A}f) := \mathsf{A}_s f - \mathsf{A}_{s-} f = \mathsf{A}_s f - \lim_{t<s, t\to s} \mathsf{A}_t f.$$

We can have $\Delta_s(\mathsf{A}f) \neq 0$ only on the event where the support of the measure μ intersects the boundary ∂Q_s. The bracket process $([\mathsf{A}f]_s)_{s\geqslant 0}$ is defined by

$$[\mathsf{A}f]_s := \sum_{0\leqslant \tau \leqslant s} \Delta_\tau(\mathsf{A}f). \tag{6.7}$$

We have that $((\mathsf{A}_s f)^2 - [\mathsf{A}f]_s)_{s\geqslant 0}$ is a martingale with respect to $(\mathcal{F}_{\overline{Q}_s})_{s\geqslant 0}$.

Notice that the operator A_s can be interpreted as an averaging of the variable $\mu \llcorner \overline{Q}_s^c$, keeping $\mu \llcorner \overline{Q}_s$ fixed. As a consequence, for every open set $Q_s \subseteq U$, if $f \in \mathscr{H}^1(U)$ and $x \in Q_s \cap \mathrm{supp}(\mu)$, there is no ambiguity in considering the quantity $\mathsf{A}_s(\nabla f)(\mu, x)$. Moreover,

$$\nabla \mathsf{A}_s f(\mu, x) = \mathsf{A}_s(\nabla f)(\mu, x), \tag{6.8}$$

and $\mathsf{A}_s f$ belongs to $\mathscr{H}^1(Q_s)$, by Jensen's inequality; see Proposition 6.A.1 for details. However, in general, this function does not belong to $\mathscr{H}_0^1(\mathbb{R}^d)$, or any other $\mathscr{H}_0^1$ space. This comes from the fact that the function $\mathsf{A}_s f$ may be discontinuous as a particle enters or leave $\overline{Q}_s$. To solve this problem, we regularize this conditional expectation in the following way. For any $s, \varepsilon > 0$, we define

$$\mathsf{A}_{s,\varepsilon} f := \frac{1}{\varepsilon} \int_0^\varepsilon \mathsf{A}_{s+t} f \, dt. \tag{6.9}$$

As above, for every open set U containing $Q_{s+\varepsilon}$, $f \in \mathscr{H}^1(U)$, and $x \in Q_{s+\varepsilon} \cap \mathrm{supp}(\mu)$, the quantity $\mathsf{A}_{s,\varepsilon}(\nabla f)(\mu, x)$ is well-defined. Irrespectively of the position of the point $x \in \mathrm{supp}(\mu)$, the gradient of $\mathsf{A}_{s,\varepsilon} f$ can be calculated explicitly. Indeed, writing $\tau(x) := \inf\{r \in \mathbb{R} : x \in Q_r\}$, and $\vec{\mathbf{n}}(x)$ for the outer unit normal to $Q_{\tau(x)}$ at the point x, we have

$$\nabla \mathsf{A}_{s,\varepsilon} f(\mu, x) = \begin{cases} \mathsf{A}_{s,\varepsilon} \left(\nabla f\right) \left(\mu, x\right) & \text{if } x \in Q_s; \\ \frac{1}{\varepsilon} \int_{\tau(x)-s}^{\varepsilon} \mathsf{A}_{s+t} \left(\nabla f(\mu, x)\right) \mathrm{d}t - \frac{\vec{\mathbf{n}}(x)}{\varepsilon} \Delta_{\tau(x)}(\mathsf{A} f) & \text{if } x \in \left(Q_{s+\varepsilon} \backslash \overline{Q}_s\right); \\ 0 & \text{if } x \in \overline{Q}_{s+\varepsilon}^c. \end{cases} \tag{6.10}$$

Recalling that $Q_{s+\varepsilon} \subseteq U$, one can check that $\mathsf{A}_{s,\varepsilon} f \in \mathscr{H}_0^1(U)$. Similarly, one can define another regularized localization operator $\widetilde{\mathsf{A}}_{s,\varepsilon}$

$$\widetilde{\mathsf{A}}_{s,\varepsilon} f := \frac{2}{\varepsilon^2} \int_0^\varepsilon (\varepsilon - t) \mathsf{A}_{s+t} f \, \mathrm{d}t, \tag{6.11}$$

which can be obtained by applying $\mathsf{A}_{s,\varepsilon}$ twice: $\widetilde{\mathsf{A}}_{s,\varepsilon} = \mathsf{A}_{s,\varepsilon} \circ \mathsf{A}_{s,\varepsilon}$. We have the identity

$$\mathbb{E}_\rho[(\mathsf{A}_{s,\varepsilon} f)^2] = \mathbb{E}_\rho[f(\widetilde{\mathsf{A}}_{s,\varepsilon} f)] = \mathbb{E}_\rho\left[\frac{2}{\varepsilon^2} \int_0^\varepsilon (\varepsilon - t)(\mathsf{A}_{s+t} f)^2 \, \mathrm{d}t\right]. \tag{6.12}$$

The operator $\widetilde{\mathsf{A}}_{s,\varepsilon}$ satisfies properties similar to those of $\mathsf{A}_{s,\varepsilon}$, and we have

$$\nabla \widetilde{\mathsf{A}}_{s,\varepsilon} f(\mu, x) = \\ \begin{cases} \widetilde{\mathsf{A}}_{s,\varepsilon} \left(\nabla f(\mu, x)\right) & x \in Q_s; \\ \frac{2}{\varepsilon^2} \left(\int_{\tau(x)-s}^{\varepsilon} (\varepsilon - t) \mathsf{A}_{s+t} \left(\nabla f(\mu, x)\right) \mathrm{d}t - (s + \varepsilon - \tau(x)) \Delta_{\tau(x)}(\mathsf{A} f) \vec{\mathbf{n}}(x) \right) & x \in \left(Q_{s+\varepsilon} \backslash \overline{Q}_s\right); \\ 0 & x \in \overline{Q}_{s+\varepsilon}^c. \end{cases} \tag{6.13}$$

6.3 Functional inequalities

The goal of this section is to derive functional inequalities that will be fundamental to the proof of our main result. The first crucial estimate is a multiscale Poincaré inequality, see Proposition 6.3.5. This inequality is an improvement over the standard Poincaré inequality that substitutes the L^2 norm of the gradient of the function of interest by a weighted sum of spatial averages of this gradient. It has a structure comparable to that of $\|u\|_{L^2} \lesssim \|\nabla u\|_{H^{-1}}$, where we moreover decompose the H^{-1} norm into a series a scales, in analogy with the standard definition of Besov spaces, or the equivalent definition of H^{-1} norm in terms of spatial averages, see for instance [25, Appendix D]. The proof of this estimate is based on an H^2 estimate for solutions of "$-\Delta u = f$", with "Δ" being the relevant Laplacian adapted to our setting; see Proposition 6.3.4.

The second crucial functional inequality derived here is a Caccioppoli inequality, see Proposition 6.3.6. In the standard elliptic setting, the Caccioppoli inequality allows to control the L^2 norm of the gradient of a solution by the L^2 norm of the function itself, on a larger domain; it can thus be thought of as a reverse Poincaré inequality for solutions. In our context, we are not able to prove such a strong estimate, but prove instead a weaker version of this inequality that allows to control the $\mathscr{L}^2$ norm of the gradient of a solution by the $\mathscr{L}^2$ norm of the function itself, plus *a fraction* of the $\mathscr{L}^2$ norm of the gradient on a larger domain.

For every $k \leqslant n \in \mathbb{N}$, we define $\mathcal{Z}_{n,k} := 3^k\mathbb{Z}^d \cap \square_n$. Up to a set of null measure, the family $(z + \square_k)_{z\in\mathcal{Z}_{n,k}}$ forms a partition of $\square_n$. For any $y \in \mathbb{R}^d$, we write $\square_n(y)$ to denote the unique cube containing y that can be written in the form $z + \square_n$ for some $z \in 3^n\mathbb{Z}^d$. This is well-defined except for some y's in a set of null measure; we can decide on an arbitrary convention for these remaining cases. We also write $\mathcal{Z}_{n,k}(y) := 3^k\mathbb{Z}^d \cap \square_n(y)$.

The following "multiscale spatial filtration" will be useful in the rest of the paper: for every $n, k \in \mathbb{N}$ with $k \leqslant n$, and $y \in \mathbb{R}^d$, we define the σ-algebra $\mathcal{G}^y_{n,k}$ by

$$\mathcal{G}^y_{n,k} := \sigma\left(\{\mu(z + \square_k)\}_{z\in\mathcal{Z}_{n,k}(y)}, \mu \llcorner (\mathbb{R}^d \backslash \square_n(y))\right). \tag{6.14}$$

We use the shorthand $\mathcal{G}_{n,k} := \mathcal{G}^0_{n,k}$ and $\mathcal{G}_n := \mathcal{G}_{n,n}$. One can verify that, for every $n, n', k, k' \in \mathbb{N}$ and $y, y' \in \mathbb{R}^d$,

$$n \leqslant n', k \leqslant k' \text{ and } \square_n(y) \subseteq \square_{n'}(y') \quad \Longrightarrow \quad \mathcal{G}^{y'}_{n',k'} \subseteq \mathcal{G}^y_{n,k}. \tag{6.15}$$

We also define the analogue of $\mathcal{G}_n$ for a general Borel set $U \subseteq \mathbb{R}^d$ as

$$\mathcal{G}_U := \sigma\left(\mu(U), \mu \llcorner (\mathbb{R}^d \backslash U)\right). \tag{6.16}$$

The condition $\mathbb{E}_\rho[f|\mathcal{G}_U] = 0$ will appear many times in this paper, usually in the context of centering a function in $\mathscr{H}^1(U)$. Using the functions (f_n) defined defined in Lemma 6.2.1, we can rewrite the condition $\mathbb{E}_\rho[f|\mathcal{G}_U] = 0$ as: for every $n \in \mathbb{N}$ and $\mathbb{P}_\rho$-almost every $\mu \in \mathcal{M}_\delta(\mathbb{R}^d)$,

$$\int_{U^n} f_n(x_1, \cdots, x_n, \mu \llcorner U^c)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n = 0. \tag{6.17}$$

6.3.1 Poincaré inequality

We present two types of Poincaré inequalities: one for the space $\mathscr{H}^1_0(U)$, and one for the space $\mathscr{H}^1(U)$. We first state an elementary version for product spaces and functions in the standard Sobolev H^1 space. The proof is classical and can be found for instance in [171, Theorem 13.36 and Proposition 13.34]. For any bounded Borel set $U \subseteq \mathbb{R}^d$, we write $\mathrm{diam}(U)$ to denote the diameter of U, and for every $f \in L^1(U)$, we denote the Lebesgue integral of f, normalized by the Lebesgue measure of U, by

$$\fint_U f := |U|^{-1} \int_U f.$$

Proposition 6.3.1 (Poincaré inequality in classical Sobolev spaces)**.** *There exists a constant $C(d) < \infty$ such that for every bounded convex open set $U \subseteq \mathbb{R}^d$, $n \in \mathbb{N}$, and $f \in H^1(U^n)$, we have*

$$\fint_{U^n} \left(f - \left(\fint_{U^n} f \right) \right)^2 \leqslant C \operatorname{diam}(U)^2 \sum_{i=1}^{n} \fint_{U^n} |\nabla_{x_i} f|^2. \tag{6.18}$$

A direct application of Proposition 6.3.1 gives the following proposition.

Proposition 6.3.2 (Poincaré inequality in $\mathscr{H}^1(U)$)**.** *There exists a constant $C(d) < \infty$ such that for every bounded convex open set and $f \in \mathscr{H}^1(U)$, we have*

$$\mathbb{E}_\rho \left[(f - \mathbb{E}_\rho[f \,|\, \mathcal{G}_U])^2 \right] \leqslant C \operatorname{diam}(U)^2 \mathbb{E}_\rho \left[\int_U |\nabla f|^2 \, \mathrm{d}\mu \right]. \tag{6.19}$$

Proof. Without loss of generality, we may assume that $\mathbb{E}_\rho[f|\mathcal{G}_U] = 0$; subtracting $\mathbb{E}_\rho[f|\mathcal{G}_U]$ from f does not change the right side of eq. (6.19). We use the functions (f_n) from Lemma 6.2.1, so that $f = \sum_{n=0}^\infty f_n \mathbf{1}_{\{\mu(U)=n\}}$, and recall that since $\mathbb{E}_\rho[f|\mathcal{G}_U] = 0$, we have that every function f_n is centered; see eq. (6.17). We can apply Proposition 6.3.1 to every f_n: for a constant $C < \infty$ independent of n, we have

$$\fint_{U^n} |f_n(x_1, \cdots, x_n, \mu \llcorner U^c)|^2 \, \mathrm{d}x_1 \cdots \mathrm{d}x_n \leqslant C \operatorname{diam}(U)^2 \sum_{i=1}^{n} \fint_{U^n} |\nabla_{x_i} f_n(x_1, \cdots, x_n, \mu \llcorner U^c)|^2 \, \mathrm{d}x_1 \cdots \mathrm{d}x_n.$$

We then sum over n and take the expectation to obtain the result. □

Functions in the space $\mathscr{H}_0^1(U)$ enjoy certain continuity properties as particles enter and leave the domain U. For this reason, it suffices to center the function by its mean value to have a Poincaré inequality.

Proposition 6.3.3 (Poincaré inequality in $\mathscr{H}_0^1(U)$)**.** *There exists a constant $C(d) < \infty$ such that for every bounded open set $U \subseteq \mathbb{R}^d$, and every $f \in \mathscr{H}_0^1(U)$,*

$$\mathbb{E}_\rho \left[(f - \mathbb{E}_\rho[f])^2 \right] \leqslant C \operatorname{diam}(U)^2 \, \mathbb{E}_\rho \left[\int_U |\nabla f|^2 \, \mathrm{d}\mu \right]. \tag{6.20}$$

Proof. Without loss of generality, we assume that $\mathbb{E}_\rho[f] = 0$. By density, we may restrict to $f \in \mathscr{C}_c^\infty(U)$. Applying [169, Theorem 18.7] to f, we have that

$$\mathbb{E}_\rho \left[f^2 \right] \leqslant \rho \int_{\mathbb{R}^d} \mathbb{E}_\rho \left[(f(\mu + \delta_x) - f(\mu))^2 \right] \mathrm{d}x.$$

By the Fubini-Tonelli theorem, and since f is $\mathcal{F}_U$-measurable, this reduces to

$$\mathbb{E}_\rho \left[f^2 \right] \leqslant \rho \mathbb{E}_\rho \left[\int_U (f(\mu + \delta_x) - f(\mu))^2 \mathrm{d}x \right].$$

To establish Proposition 6.3.3, it thus only remains to show that

$$\mathbb{E}_\rho \left[\int_U (f(\mu + \delta_x) - f(\mu))^2 \mathrm{d}x \right] \leqslant \frac{C(d)}{\rho} \, \mathbb{E}_\rho \left[\int_U |\nabla f|^2 \mathrm{d}\mu \right]. \tag{6.21}$$

We recall that

$$\int_U \mathbb{E}_\rho\left[(f(\cdot+\delta_x)-f(\cdot))^2\right]\mathrm{d}x$$
$$= \sum_{n\in\mathbb{N}} \mathbb{P}(\mu(U)=n) \fint_{U^n}\left(\int_U |f_{n+1}(x_1,\cdots,x_n,x)-f_n(x_1,\cdots,x_n)|^2\mathrm{d}x\right)\mathrm{d}x_1\cdots\mathrm{d}x_n, \quad (6.22)$$

where we used the notation (similar but simpler than in Lemma 6.2.1)

$$f_n(x_1,\ldots,x_n) := f\left(\sum_{k=1}^n \delta_{x_k}\right), \qquad x_1,\cdots,x_n\in U. \quad (6.23)$$

Let $n\in\mathbb{N}$ be fixed. Since $f\in\mathscr{C}_c^\infty(U)$, for every $\bar{x}\in\partial U$ we have that

$$f_n(x_1,\cdots,x_n) = f_{n+1}(x_1,\cdots,x_n,\bar{x}).$$

That is, for every $x_1,\cdots,x_n\in U^n$, the (smooth) function

$$G: U\to\mathbb{R}, \qquad G(\cdot) := f_{n+1}(x_1,\cdots,x_n,\cdot)-f_n(x_1,\cdots,x_n),$$

belongs to the (standard) Sobolev space $H_0^1(U)$. We may thus apply the standard Poincaré inequality for functions in $H_0^1(U)$ to infer that

$$\int_U |f_{n+1}(x_1,\cdots,x_n,x)-f_n(x_1,\cdots,x_n)|^2\mathrm{d}x$$
$$\leqslant C(d)\,\mathrm{diam}(U)^2\int_U |\nabla_x f_{n+1}(x_1,\cdots,x_n,x)|^2\mathrm{d}x.$$

Inserting this into (6.22), using that $\mathbb{P}(\mu(U)=n) = e^{-\rho|U|}\frac{(\rho|U|)^n}{n!}$ and relabelling $n+1$ as n, yields that

$$\int_U \mathbb{E}_\rho\left[(f(\cdot+\delta_x)-f(\cdot))^2\right]\mathrm{d}x$$
$$\leqslant \frac{C(d)}{\rho}\mathrm{diam}(U)^2\sum_{n\in\mathbb{N}}\mathbb{P}(\mu(U)=n)n\fint_{U^n}|\nabla_{x_n}f_n(x_1,\cdots,x_n)|^2\mathrm{d}x_1\cdots\mathrm{d}x_n.$$

To establish (6.21) from this, it only remains to observe that, since by definition (6.23) each function f_n is invariant under permutations, we have

$$\fint_{U^n}|\nabla_{x_1}f_n|^2 = \fint_{U^n}|\nabla_{x_i}f_n|^2 \quad \text{for all } i=1,\cdots,n.$$

This concludes the proof of (6.21) and establishes Proposition 6.3.3. □

6.3.2 $\mathscr{H}^2$ estimate for the homogeneous equation

When the diffusion matrix **a** is a constant, the solutions to the corresponding equation have a better regularity than otherwise, and in particular, the following $\mathscr{H}^2$ estimate holds. One can define the function with higher derivative iteratively: for $x,y\in\mathrm{supp}(\mu), x\neq y$

$$\partial_j\partial_k f(\mu,x,y) := \lim_{h\to 0}\frac{\partial_k f(\mu-\delta_y+\delta_{y+he_j},x)-\partial_k f(\mu,x)}{h},$$

and for the case $x = y$, it makes sense as

$$\partial_j \partial_k f(\mu, x, x) := \lim_{h \to 0} \frac{\partial_k f(\mu - \delta_x + \delta_{x+he_j}, x + he_j) - \partial_k f(\mu, x)}{h}.$$

We also denote by $\nabla^2 f(\mu, x, y)$ the matrix $\{\partial_j \partial_k f(\mu, x, y)\}_{1 \leqslant j,k \leqslant d}$, and its norm is defined as

$$|\nabla^2 f(\mu, x, y)|^2 := \sum_{1 \leqslant j,k \leqslant d} |\partial_j \partial_k f(\mu, x, y)|^2.$$

Proposition 6.3.4 ($\mathscr{H}^2$ estimate)**.** *Let $f \in \mathscr{L}^2$, and let $u \in \mathscr{H}^1(Q_r)$ solve "$-\Delta u = f$" in the sense that for any $v \in \mathscr{H}^1(Q_r)$,*

$$\mathbb{E}_\rho \left[\int_{Q_r} \nabla u(\mu, x) \cdot \nabla v(\mu, x) \, d\mu \right] = \mathbb{E}_\rho[fv]. \tag{6.24}$$

We have the $\mathscr{H}^2(Q_r)$ estimate

$$\mathbb{E}_\rho \left[\int_{(Q_r)^2} |\nabla^2 u(\mu, x, y)|^2 \, \mathrm{d}\mu(x) \mathrm{d}\mu(y) \right] \leqslant \mathbb{E}_\rho[f^2]. \tag{6.25}$$

Remark. By testing eq. (6.24) with $v = \mathbf{1}_{\{\mu(Q_r)=n, \mu \llcorner Q_r^c(V)=m\}}$, we see that f has to satisfy $\mathbb{E}_\rho[f \,|\, \mathcal{G}_{Q_r}] = 0$ as a condition of compatibility.

Proof of Proposition 6.3.4. Although this is not really part of the statement, we start by showing that for every $f \in \mathscr{L}^2$ satisfying the compatibility condition $\mathbb{E}_\rho[f \,|\, \mathcal{G}_{Q_r}] = 0$, there exists a solution u to eq. (6.24), and we will show its link with the classical elliptic equation. At first, we notice that the problem can be studied on the space of functions

$$W = \{g \in \mathscr{H}^1(Q_r) : \mathbb{E}_\rho[g \,|\, \mathcal{G}_{Q_r}] = 0\}.$$

Because for a general function $v \in \mathscr{H}^1(Q_r)$, $\mathbb{E}_\rho[v|\mathcal{G}_{Q_r}]$ can be seen as a constant in eq. (6.24): its derivative is 0 so the left-hand side of eq. (6.24) is 0. For the right-hand side, we have

$$\mathbb{E}_\rho[f \mathbb{E}_\rho[v|\mathcal{G}_{Q_r}]] = \mathbb{E}_\rho \left[\mathbb{E}_\rho[f|\mathcal{G}_{Q_r}] \mathbb{E}_\rho[v|\mathcal{G}_{Q_r}] \right] = 0.$$

Thus when applying the operation $v \mapsto v - \mathbb{E}_\rho[v|\mathcal{G}_{Q_r}]$, we do not change eq. (6.24) and we can restrict the Laplace equation on W. Moreover, with the notation in Lemma 6.2.1 that $v = \sum_{n=0}^\infty v_n \mathbf{1}_{\{\mu(Q_r)=n\}}$, $\mathbb{E}_\rho[v|\mathcal{G}_{Q_r}] = 0$ implies every v_n is centered; see eq. (6.17).

Secondly, we test eq. (6.24) by conditioning the environment outside Q_r and the number of particles $\mu(Q_r)$, i.e. let

$$v = v_n(x_1, \cdots, x_n, \mu \llcorner Q_r^c) \mathbf{1}_{\{\mu(Q_r)=n\}} \mathbf{1}_{\{\mu \llcorner Q_r^c(V)=m\}},$$

for some Borel set V using the canonical projection Lemma 6.2.1. Then for arbitrary choices of n, m, V, in fact we have a classical elliptic equation

$$\begin{aligned} \int_{(Q_r)^n} \sum_{k=1}^n \nabla_{x_k} u_n(x_1, \cdots, x_n, \mu \llcorner Q_r^c) \cdot \nabla_{x_k} v_n(x_1, \cdots, x_n, \mu \llcorner Q_r^c) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n \\ = \int_{(Q_r)^n} f_n(x_1, \cdots, x_n, \mu \llcorner Q_r^c) v_n(x_1, \cdots, x_n, \mu \llcorner Q_r^c) \, \mathrm{d}x_1 \cdots \mathrm{d}x_n, \end{aligned} \tag{6.26}$$

with the notation of the canonical projection. Thus the solution u can be described as follows: we sample the environment outside Q_r and fix the number of particle $\mu(Q_r) = n$ at first, then solve the classical elliptic equation in $H^1(\mathbb{R}^{nd})$ with mean zero. Finally we combine all the u_n and this gives the solution of eq. (6.24). In other words, the statement of eq. (6.24) can be reinforced as

$$\forall v \in W, \qquad \mathbb{E}_\rho\left[\int_{Q_r} \nabla u(\mu, x) \cdot \nabla v(\mu, x)\, d\mu \;\middle|\; \mathcal{G}_{Q_r} \right] = \mathbb{E}_\rho\left[f v \,\middle|\, \mathcal{G}_{Q_r} \right].$$

We now turn to study the $\mathscr{H}^2$ estimate. We apply the classical $H^2(\mathbb{R}^{nd})$ estimate for eq. (6.26) (see for instance [25, Lemma B.19] and its proof)

$$\int_{(Q_r)^n} \sum_{1 \leqslant i,j \leqslant n} |\nabla_{x_i} \nabla_{x_j} u_n|^2 (x_1, \cdots, x_n, \mu \llcorner Q_r^c)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n \\ \leqslant \int_{(Q_r)^n} |f_n|^2 (x_1, \cdots, x_n, \mu \llcorner Q_r^c)\, \mathrm{d}x_1 \cdots \mathrm{d}x_n, \quad (6.27)$$

Taking the expectation of eq. (6.27) then gives the result. □

6.3.3 Multiscale Poincaré inequality

For cubes of size 3^n, the Poincaré inequalities derived in the previous subsection (say with $k = n$ in Proposition 6.3.2) have a right-hand side that scales like 3^{2n}. In this subsection, we derive a multiscale version of the Poincaré inequality, that aims to improve upon this scaling, provided that some local average of the gradient of the function is not too large. We recall that the multiscale spatial filtration $\mathcal{G}^y_{n,k}$ is defined in eq. (6.14). For every $k \leqslant n \in \mathbb{N}$, $x, y \in \mathbb{R}^d$ such that $x \in \square_n(y)$, open set U containing $\square_k(x)$, and $f \in \mathscr{H}^1(U)$, the following quantity is well defined

$$(\mathsf{S}^y_{n,k} \nabla f)(\mu, x) := \mathbb{E}_\rho\left[\fint_{\square_k(x)} \nabla f \, \mathrm{d}\mu \;\middle|\; \mathcal{G}^y_{n,k} \right], \quad (6.28)$$

where we use the notation, for every Borel set V such that $\mu(V) \in (0, \infty)$ and function g defined on $\operatorname{supp}(\mu) \cap V$,

$$\fint_V g \, \mathrm{d}\mu := \frac{1}{\mu(V)} \int_V g \, \mathrm{d}\mu, \quad (6.29)$$

and for definiteness, we also set $\fint_V g \, \mathrm{d}\mu = 0$ if $\mu(V) = 0$. We use the shorthand notation $\mathsf{S}_{n,k} := \mathsf{S}^0_{n,k}$ and $\mathsf{S}_n := \mathsf{S}_{n,n}$. This operator has a convenient spatial martingale structure, as displayed in the following lemma.

Lemma 6.3.1 (Martingale structure for $\mathsf{S}_{n,k}$)**.** *For every $n, n', k, k' \in \mathbb{N}$, $y, y' \in \mathbb{R}^d$ satisfying*

$$n \leqslant n', \quad k \leqslant k', \quad \square_n(y) \subseteq \square_{n'}(y'),$$

every $x \in \square_{k'}(y')$, and $f \in \mathscr{H}^1(\square_{n'}(y'))$, we have

$$\mathsf{S}^{y'}_{n',k'} \nabla f(\mu, x) = \mathbb{E}_\rho\left[\fint_{\square_{k'}(x)} (\mathsf{S}^y_{n,k} \nabla f) \, \mathrm{d}\mu \;\middle|\; \mathcal{G}^{y'}_{n',k'} \right]. \quad (6.30)$$

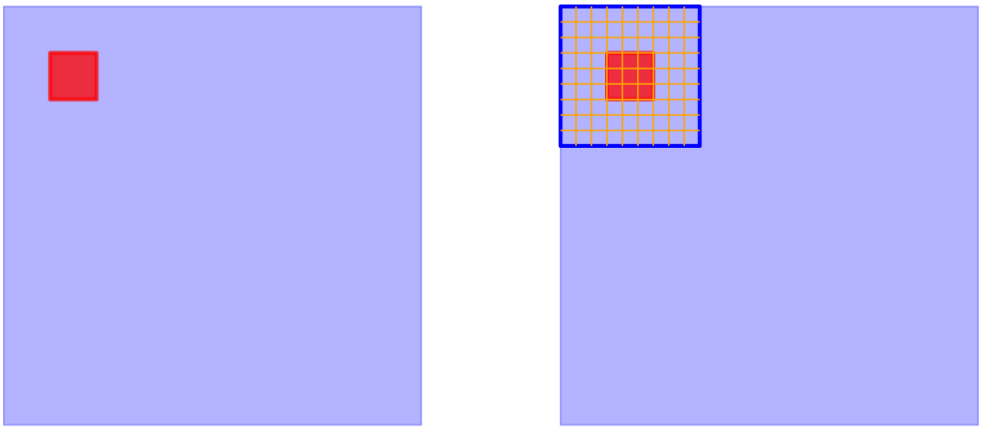

Figure 6.1: The largest cube on this figure is $\square_{n'}(y')$. The operator $\mathsf{S}^{y'}_{n',k'}$ computes the spatial average in every subcubeof size $3^{k'}$, for example the cube in red in the image. We can apply at first the operator $\mathsf{S}^{y}_{n,k}$, which works on the finer scales 3^k and 3^n, represented by the cubes with orange and blue boundaries respectively.

Proof. The key observation is eq. (6.15), stating that $\mathcal{G}^y_{n,k}$ is a finer σ-algebra than $\mathcal{G}^{y'}_{n',k'}$, so that

$$\begin{aligned}&\mathsf{S}^{y'}_{n',k'}\nabla f(\mu,x)\\&\quad=\mathbb{E}_\rho\left[\frac{1}{\mu(\square_{k'}(x))}\int_{\square_{k'}(x)}\nabla f\,\mathrm{d}\mu\;\middle|\;\mathcal{G}^{y'}_{n',k'}\right]\\&\quad=\mathbb{E}_\rho\left[\sum_{z\in\mathcal{Z}_{n,k}\cap\square_{k'}(x)}\frac{\mu(z+\square_k)}{\mu(\square_{k'}(x))}\mathbb{E}_\rho\left[\frac{1}{\mu(z+\square_k)}\int_{\square_k(z)}\nabla f\,\mathrm{d}\mu\;\middle|\;\mathcal{G}^y_{n,k}\right]\;\middle|\;\mathcal{G}^{y'}_{n',k'}\right].\end{aligned}$$

By the definition of $\mathsf{S}^y_{n,k}\nabla f(\mu,z)$, we obtain

$$\begin{aligned}\mathsf{S}^{y'}_{n',k'}\nabla f(\mu,x)&=\mathbb{E}_\rho\left[\sum_{z\in\mathcal{Z}_{n,k}\cap\square_{k'}(x)}\frac{\mu(z+\square_k)}{\mu(\square_{k'}(x))}(\mathsf{S}^y_{n,k}\nabla f)(\mu,z)\;\middle|\;\mathcal{G}^{y'}_{n',k'}\right]\\&=\mathbb{E}_\rho\left[\frac{1}{\mu(\square_{k'}(x))}\int_{\square_{k'}(x)}\mathsf{S}^y_{n,k}\nabla f\,\mathrm{d}\mu\;\middle|\;\mathcal{G}^{y'}_{n',k'}\right].\end{aligned}$$

This is eq. (6.30). □

To prepare further for the multiscale Poincaré inequality, we also give the following explicit expression for $\mathsf{S}^y_{n,k}\nabla f$. We use the notation

$$\fint_{(z_i+\square_k)_{1\leqslant i\leqslant N}}:=\fint_{z_i+\square_k}\cdots\fint_{z_N+\square_k}.$$

Lemma 6.3.2. *Using the notation of Lemma 6.2.1 with $\mu\llcorner\square_n(y)=\sum_{i=1}^N\delta_{x_i}$, for any $x\in\square_n(y)$ and any $f\in\mathscr{H}^1(\square_n(y))$, we have*

$$\begin{aligned}&(\mathsf{S}^y_{n,k}\nabla f)(\mu,x)\prod_{z\in\mathcal{Z}_{n,k}(y)}\mathbf{1}_{\{\mu(z+\square_k)=N_z\}}\\&\quad=\frac{1}{\mu(\square_k(x))}\sum_{j:x_j\in\square_k(x)}\fint_{(z_i+\square_k)_{1\leqslant i\leqslant N}}\nabla_{x_j}f_N(\cdot,\mu\llcorner\square_n^c)\prod_{z\in\mathcal{Z}_{n,k}(y)}\mathbf{1}_{\{\mu(z+\square_k)=N_z\}},\quad(6.31)\end{aligned}$$

with $N = \sum_{z\in\mathcal{Z}_{n,k}(y)} N_z$ *and* $\{z_i\}_{1\leqslant i\leqslant N}$ *any fixed sequence such that*

$$\forall z \in \mathcal{Z}_{n,k}(y), \qquad |\{i \in \{1,\dots,N\} \ : \ z_i = z\}| = N_z. \tag{6.32}$$

Moreover, for every j, j' *such that* $x_j, x_{j'} \in \square_k(x)$*, we have*

$$\fint_{(z_i+\square_k)_{1\leqslant i\leqslant N}} \nabla_{x_j} f_N(\cdot, \mu \llcorner \square_n^c) = \fint_{(z_i+\square_k)_{1\leqslant i\leqslant N}} \nabla_{x_{j'}} f_N(\cdot, \mu \llcorner \square_n^c). \tag{6.33}$$

Proof. Without loss of generality, we set $y = 0$. Then let $N = \sum_{z\in\mathcal{Z}_{n,k}} N_z$ and we use the canonical projection

$$\begin{aligned}&(\mathsf{S}_{n,k}\nabla f)(\mu,x) \prod_{z\in\mathcal{Z}_{n,k}(y)} \mathbf{1}_{\{\mu(z+\square_k)=N_z\}}\\ =&\frac{1}{\mu(\square_k(x))}\mathbb{E}_\rho\left[\sum_{x_j\in\square_k(x)} \nabla_{x_j} f_N(\cdot,\mu\llcorner\square_n^c) \prod_{z\in\mathcal{Z}_{n,k}} \mathbf{1}_{\{\mu(z+\square_k)=N_z\}} \middle| \mathcal{G}_{n,k}\right].\end{aligned}$$

The key point is to write $\prod_{z\in\mathcal{Z}_{n,k}} \mathbf{1}_{\{\mu(z+\square_k)=N_z\}}$ with respect to $\{x_i\}_{1\leqslant i\leqslant N}$ such that $\mu \llcorner \square_n = \sum_{i=1}^N \delta_{x_i}$. Let $\{z_i\}_{1\leqslant x_i\leqslant N}$ be any fixed sequence so that every z in $\mathcal{Z}_{n,k}$ appears exactly N_z times, as displayed in eq. (6.32). We have

$$\prod_{z\in\mathcal{Z}_{n,k}} \mathbf{1}_{\{\mu(z+\square_k)=N_z\}} = \sum_{\sigma\in S_N}\prod_{i=1}^N \mathbf{1}_{\{x_{\sigma(i)}\in z_i+\square_k\}},$$

where S_N is the symmetric group. Moreover, under $\mathcal{G}_{n,k}$ every permutation has equal probability, and then each x_i is uniformly distributed in the associated cube $z_{\sigma(i)} + \square_k$. Thus, we have

$$\begin{aligned}&(\mathsf{S}_{n,k}\nabla f)(\mu,x) \prod_{z\in\mathcal{Z}_{n,k}} \mathbf{1}_{\{\mu(z+\square_k)=N_z\}}\\ =&\frac{1}{\mu(\square_k(x))}\frac{1}{|S_N|}\sum_{\sigma\in S_N}\fint_{(z_{\sigma(i)}+\square_k)_{1\leqslant i\leqslant N}}\sum_{x_j\in\square_k(x)} \nabla_{x_j} f_N(\cdot,\mu\llcorner\square_n^c)\prod_{i=1}^N \mathbf{1}_{\{x_i\in z_{\sigma(i)}+\square_k\}}.\end{aligned}$$

Notice that for every $1 \leqslant i \leqslant N$, $x_i \in z_{\sigma(i)}+\square_k$ means $x_{\sigma^{-1}(i)} \in z_i+\square_k$, and $\sum_{x_j\in\square_k(x)} \nabla_{x_j} f_N(\cdot,\mu\llcorner\square_n^c)$ is permutation-invariant. So we have

$$\begin{aligned}&\sum_{x_j\in\square_k(x)} \nabla_{x_j} f_N(x_1,\cdots,x_N,\mu\llcorner\square_n^c)\prod_{i=1}^N \mathbf{1}_{\{x_i\in z_{\sigma(i)}+\square_k\}}\\ =&\sum_{x_j\in\square_k(x)} \nabla_{x_j} f_N(x_1,\cdots,x_N,\mu\llcorner\square_n^c)\prod_{i=1}^N \mathbf{1}_{\left\{x_{\sigma^{-1}(i)}\in z_i+\square_k\right\}}\\ =&\sum_{x_{\sigma^{-1}(j)}\in\square_k(x)} \nabla_{x_{\sigma^{-1}(j)}} f_N(x_{\sigma^{-1}(1)},\cdots,x_{\sigma^{-1}(N)},\mu\llcorner\square_n^c)\prod_{i=1}^N \mathbf{1}_{\left\{x_{\sigma^{-1}(i)}\in z_i+\square_k\right\}}\\ =&\sum_{x_j\in\square_k(x)} \nabla_{x_j} f_N(x_1,\cdots,x_N,\mu\llcorner\square_n^c)\prod_{i=1}^N \mathbf{1}_{\{x_i\in z_i+\square_k\}}.\end{aligned}$$

Therefore, the term for each permutation has the same contribution, and we thus obtain eq. (6.31).

Then we prove eq. (6.33). To avoid possible confusion in the notation, we let $y_j, y_{j'}$ be the j-th and j'-th coordinates, then we exchange them and use the invariance under permutation of f_N,

$$\begin{aligned} e_k \cdot \nabla_{x_j} f_N(\cdots, y_j, \cdots y_{j'}, \cdots) &= \lim_{h\to 0} \frac{f_N(\cdots, y_j + he_k, \cdots y_{j'}, \cdots) - f_N(\cdots, y_j, \cdots y_{j'}, \cdots)}{h} \\ &= \lim_{h\to 0} \frac{f_N(\cdots, y_{j'}, \cdots y_j + he_k, \cdots) - f_N(\cdots, y_{j'}, \cdots y_j, \cdots)}{h} \\ &= e_k \cdot \nabla_{x_{j'}} f_N(\cdots, y_{j'}, \cdots y_j, \cdots). \end{aligned} \tag{6.34}$$

Moreover, the condition $x_j, x_{j'} \in \square_k(x)$ implies that $z_j = z_{j'}$ and

$$\mathbf{1}_{\{y_j \in z_j + \square_k\}} \mathbf{1}_{\{y_{j'} \in z_{j'} + \square_k\}} = \mathbf{1}_{\{y_{j'} \in z_j + \square_k\}} \mathbf{1}_{\{y_j \in z_{j'} + \square_k\}}. \tag{6.35}$$

We combine eq. (6.34) and eq. (6.35) to conclude eq. (6.33). □

We now use the operators $\mathsf{S}^y_{n,k}$ as our locally averaged gradient to obtain the following multiscale Poincaré inequality. Notice in particular the factor of 3^k inside the sum on the right side of eq. (6.36), which we aim to leverage upon later by combining this with information on the smallness of $\mathsf{S}_{n,k}\nabla u$ for k close to n.

Proposition 6.3.5 (Multiscale Poincaré inequality)**.** *There exists a constant $C(d) < \infty$ such that for every function $u \in \mathscr{H}^1(\square_n)$ satisfying $\mathbb{E}_\rho[u \,|\, \mathcal{G}_n] = 0$, we have*

$$\|u\|_{\mathscr{L}^2} \leqslant C \left(\mathbb{E}_\rho \left[\int_{\square_n} |\nabla u|^2 \,\mathrm{d}\mu \right] \right)^{\frac{1}{2}} + C \sum_{k=0}^{n} 3^k \left(\mathbb{E}_\rho \left[\int_{\square_n} |\mathsf{S}_{n,k} \nabla u|^2 \,\mathrm{d}\mu \right] \right)^{\frac{1}{2}}. \tag{6.36}$$

Proof. Let $w \in \mathscr{H}^1(\square_n)$ be such that $\mathbb{E}_\rho[w \,|\, \mathcal{G}_n] = 0$ and that solves "$-\Delta w = u$", in the sense that

$$\forall v \in \mathscr{H}^1(\square_n), \qquad \mathbb{E}_\rho \left[\int_{\square_n} \nabla w \cdot \nabla v \,\mathrm{d}\mu \right] = \mathbb{E}_\rho[uv], \tag{6.37}$$

and this relation also holds conditionally on $\mathcal{G}_n$:

$$\forall v \in \mathscr{H}^1(\square_n), \qquad \mathbb{E}_\rho \left[\int_{\square_n} \nabla w \cdot \nabla v \,\mathrm{d}\mu \,\middle|\, \mathcal{G}_n \right] = \mathbb{E}_\rho[uv \,|\, \mathcal{G}_n]. \tag{6.38}$$

Thanks to the condition $\mathbb{E}_\rho[u \,|\, \mathcal{G}_n] = 0$, these equations are well-defined; see the proof of Proposition 6.3.4 for a detailed discussion. This proposition asserts that

$$\mathbb{E}_\rho \left[\int_{(\square_n)^2} |\nabla^2 w(\mu, x, y)|^2 \,\mathrm{d}\mu(x) \mathrm{d}\mu(y) \right] \leqslant \mathbb{E}_\rho[u^2]. \tag{6.39}$$

We test eq. (6.37) with u and write a telescopic sum with $(\mathsf{S}_{n,k}\nabla w)_{0 \leqslant k \leqslant n}$ to get

$$\begin{aligned} \mathbb{E}_\rho[u^2] = \mathbb{E}_\rho \left[\int_{\square_n} \nabla w \cdot \nabla u \,\mathrm{d}\mu \right] &= \text{eq. (6.40)-a} + \text{eq. (6.40)-b} + \text{eq. (6.40)-c}, \\ \text{eq. (6.40)-a} &= \mathbb{E}_\rho \left[\int_{\square_n} (\nabla w - \mathsf{S}_{n,0} \nabla w) \cdot \nabla u \,\mathrm{d}\mu \right], \\ \text{eq. (6.40)-b} &= \sum_{k=0}^{n-1} \mathbb{E}_\rho \left[\int_{\square_n} (\mathsf{S}_{n,k} \nabla w - \mathsf{S}_{n,k+1} \nabla w) \cdot \nabla u \,\mathrm{d}\mu \right], \\ \text{eq. (6.40)-c} &= \mathbb{E}_\rho \left[\int_{\square_n} (\mathsf{S}_{n,n} \nabla w) \cdot \nabla u \,\mathrm{d}\mu \right]. \end{aligned} \tag{6.40}$$

We treat each of these three terms in turn. For eq. (6.40)-a, we use the Cauchy-Schwarz inequality to write

$$\text{eq. (6.40)-a} \leqslant \left(\mathbb{E}_\rho\left[\int_{\square_n} |\nabla w - \mathsf{S}_{n,0}\nabla w|^2 \,\mathrm{d}\mu\right]\right)^{\frac{1}{2}} \left(\mathbb{E}_\rho\left[\int_{\square_n} |\nabla u|^2 \,\mathrm{d}\mu\right]\right)^{\frac{1}{2}}.$$

The first term on the right side above can be rewritten as

$$\mathbb{E}_\rho\left[\int_{\square_n} |\nabla w - \mathsf{S}_{n,0}\nabla w|^2 \,\mathrm{d}\mu\right] = \mathbb{E}_\rho\left[\sum_{z\in\mathcal{Z}_{n,0}} \mathbb{E}_\rho\left[\int_{z+\square_0} |\nabla w - \mathsf{S}_{n,0}\nabla w|^2 \,\mathrm{d}\mu \,\middle|\, \mathcal{G}_{n,0}\right]\right]. \tag{6.41}$$

We use the canonical projection Lemma 6.2.1 for w with $\mu \llcorner \square_n = \sum_{i=1}^N \delta_{x_i}$, and the do the decomposition conditioned on $\mathcal{G}_{n,0}$ that

$$w(\mu) = \sum_{N=0}^{\infty} \sum_{\sum_{z\in\mathcal{Z}_{n,0}} N_z = N} w_N(x_1,\cdots,x_N,\mu\llcorner\square_n^c) \prod_{z\in\mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}}.$$

It suffices to study one term $w_N(x_1,\cdots,x_N,\mu\llcorner\square_n^c)\prod_{z\in\mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}}$. We can apply eq. (6.31): let $\{z_i\}_{1\leqslant i\leqslant N}$ be a fixed sequence such that eq. (6.32) holds (with $y=0$ there). For any $x\in\square_n$ we have

$$\begin{aligned} &\mathsf{S}_{n,0}\nabla w(\mu,x)\prod_{z\in\mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}} \\ &\qquad = \frac{1}{\mu(\square_0(x))}\sum_{x_j\in\square_0(x)} \fint_{(z_i+\square_0)_{1\leqslant i\leqslant N}} \nabla_{x_j} w_N(\cdot,\mu\llcorner\square_n^c)\prod_{z\in\mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}}. \end{aligned} \tag{6.42}$$

We apply eq. (6.42) in eq. (6.41) and just study the sum over one z' in $\mathcal{Z}_{n,0}$

$$\begin{aligned} &\mathbb{E}_\rho\left[\int_{z'+\square_0} |\nabla w - \mathsf{S}_{n,0}\nabla w|^2\,\mathrm{d}\mu \prod_{z\in\mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}} \,\middle|\, \mathcal{G}_{n,0}\right] \\ &= \sum_{x_j\in z'+\square_0} \fint_{(z_i+\square_0)_{1\leqslant i\leqslant N}} \left|\nabla_{x_j} w_N(\cdot,\mu\llcorner\square_n^c) - \frac{1}{\mu(z'+\square_0)}\sum_{x_{j'}\in z'+\square_0}\fint_{(z_i+\square_0)_{1\leqslant i\leqslant N}} \nabla_{x_{j'}} w_N(\cdot,\mu\llcorner\square_n^c)\right|^2 \\ &\qquad \times \prod_{z\in\mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}}. \end{aligned}$$

Then we use the symmetry proved in eq. (6.33), that in fact every $\nabla_{x_j} w_N$ has the same contribution for all $x_j \in z'+\square_0$,

$$\begin{aligned} &\mathbb{E}_\rho\left[\int_{z'+\square_0} |\nabla w - \mathsf{S}_{n,0}\nabla w|^2\,\mathrm{d}\mu \prod_{z\in\mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}} \,\middle|\, \mathcal{G}_{n,0}\right] \\ &= \sum_{x_j\in z'+\square_0} \fint_{(z_i+\square_0)_{1\leqslant i\leqslant N}} \left|\nabla_{x_j} w_N(\cdot,\mu\llcorner\square_n^c) - \fint_{(z_i+\square_0)_{1\leqslant i\leqslant N}} \nabla_{x_j} w_N(\cdot,\mu\llcorner\square_n^c)\right|^2 \prod_{z\in\mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}}. \end{aligned}$$

For the equation above, we can use the Poincaré inequality Proposition 6.3.1 because it is centered and every x_i lives uniformly in its associated small cube $z_i+\square_0$. We remark that the constant C here is independent of N.

$$\begin{aligned} &\mathbb{E}_\rho\left[\int_{z'+\square_0} |\nabla w - \mathsf{S}_{n,0}\nabla w|^2\,\mathrm{d}\mu \prod_{z\in\mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}} \,\middle|\, \mathcal{G}_{n,0}\right] \\ &\qquad \leqslant C\sum_{1\leqslant i\leqslant N}\sum_{x_j\in z'+\square_0} \fint_{(z_i+\square_0)_{1\leqslant i\leqslant N}} \left|\nabla_{x_i}\nabla_{x_j} w_N(\cdot,\mu\llcorner\square_n^c)\right|^2 \prod_{z\in\mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}}. \end{aligned}$$

We put this estimate back to eq. (6.41), do the sum over all $z' \in \mathcal{Z}_{n,0}$

$$\begin{aligned}
&\sum_{z' \in \mathcal{Z}_{n,0}} \mathbb{E}_\rho \left[\int_{z'+\square_0} |\nabla w - \mathsf{S}_{n,0} \nabla w|^2 \, \mathrm{d}\mu \prod_{z \in \mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}} \,\middle|\, \mathcal{G}_{n,0} \right] \\
&\quad \leqslant C \sum_{1 \leqslant i,j \leqslant N} \fint_{(z_i+\square_0)_{1\leqslant i \leqslant N}} \left| \nabla_{x_i} \nabla_{x_j} w_N(\cdot, \mu \llcorner \square_n^c) \right|^2 \prod_{z \in \mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}} \\
&\quad = C \mathbb{E}_\rho \left[\int_{(\square_n)^2} \left| \nabla^2 w(\mu, x, y) \right|^2 \, \mathrm{d}\mu(x) \mathrm{d}\mu(y) \prod_{z \in \mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}} \,\middle|\, \mathcal{G}_{n,0} \right].
\end{aligned}$$

Finally, we do the expectation and the sum over all $\prod_{z \in \mathcal{Z}_{n,0}} \mathbf{1}_{\{\mu(z+\square_0)=N_z\}}$, and use the $\mathscr{H}^2$-estimate eq. (6.39) to obtain that

$$\begin{aligned}
&\mathbb{E}_\rho \left[\int_{\square_n} |\nabla w - \mathsf{S}_{n,0} \nabla w|^2 \, \mathrm{d}\mu \right] \\
&\quad \leqslant C \mathbb{E}_\rho \left[\int_{(\square_n)^2} \left| \nabla^2 w(\mu, x, y) \right|^2 \, \mathrm{d}\mu(x) \mathrm{d}\mu(y) \right] \\
&\quad \leqslant C \mathbb{E}_\rho [u^2],
\end{aligned} \tag{6.43}$$

and this concludes that

$$\text{eq. (6.40)-a} \leqslant C (\mathbb{E}_\rho [u^2])^{\frac{1}{2}} \left(\mathbb{E}_\rho \left[\int_{\square_n} |\nabla u|^2 \, \mathrm{d}\mu \right] \right)^{\frac{1}{2}}. \tag{6.44}$$

The term eq. (6.40)-b can be treated similarly. For every k, we apply at first the conditional expectation with respect to $\mathcal{G}_{n,k}$

$$\begin{aligned}
&\mathbb{E}_\rho \left[\int_{\square_n} (\mathsf{S}_{n,k} \nabla w - \mathsf{S}_{n,k+1} \nabla w) \cdot \nabla u \, \mathrm{d}\mu \right] \\
&\quad = \sum_{z \in \mathcal{Z}_{n,k}} \mathbb{E}_\rho \left[\mathbb{E}_\rho \left[\int_{z+\square_k} (\mathsf{S}_{n,k} \nabla w - \mathsf{S}_{n,k+1} \nabla w) \cdot \nabla u \, \mathrm{d}\mu \,\middle|\, \mathcal{G}_{n,k} \right] \right] \\
&\quad = \mathbb{E}_\rho \left[\sum_{z \in \mathcal{Z}_{n,k}} \int_{z+\square_k} (\mathsf{S}_{n,k} \nabla w - \mathsf{S}_{n,k+1} \nabla w) \cdot (\mathsf{S}_{n,k} \nabla u) \, \mathrm{d}\mu \right].
\end{aligned}$$

Then we use the Cauchy-Schwarz inequality to obtain that

$$\begin{aligned}
&\mathbb{E}_\rho \left[\int_{\square_n} (\mathsf{S}_{n,k} \nabla w - \mathsf{S}_{n,k+1} \nabla w) \cdot \nabla u \, \mathrm{d}\mu \right] \\
&\quad \leqslant \left(\mathbb{E}_\rho \left[\sum_{z \in \mathcal{Z}_{n,k}} \int_{z+\square_k} |\mathsf{S}_{n,k} \nabla w - \mathsf{S}_{n,k+1} \nabla w|^2 \, \mathrm{d}\mu \right] \right)^{\frac{1}{2}} \left(\mathbb{E}_\rho \left[\sum_{z \in \mathcal{Z}_{n,k}} \int_{z+\square_k} |\mathsf{S}_{n,k} \nabla u|^2 \, \mathrm{d}\mu \right] \right)^{\frac{1}{2}}.
\end{aligned}$$

We use the definition in eq. (6.28) and Jensen's inequality for $|\mathsf{S}_{n,k} \nabla w - \mathsf{S}_{n,k+1} \nabla w|^2$. For every $z \in \mathcal{Z}_{n,k}$, since $(\mathsf{S}_{n,k+1} \nabla w)(\mu, z)$ is $\mathcal{G}_{n,k}$-measurable,

$$\begin{aligned}
|\mathsf{S}_{n,k} \nabla w - \mathsf{S}_{n,k+1} \nabla w|^2(\mu, z) &= \left(\mathbb{E}_\rho \left[\fint_{z+\square_k} (\nabla w - \mathsf{S}_{n,k+1} \nabla w) \, \mathrm{d}\mu \,\middle|\, \mathcal{G}_{n,k} \right] \right)^2 \\
&\leqslant \mathbb{E}_\rho \left[\fint_{z+\square_k} |\nabla w - \mathsf{S}_{n,k+1} \nabla w|^2 \, \mathrm{d}\mu \,\middle|\, \mathcal{G}_{n,k} \right].
\end{aligned} \tag{6.45}$$

Then we sum over all $z \in \mathcal{Z}_{n,k}$, and we can treat it like eq. (6.40)-a and eq. (6.43) with the Poincaré inequality in the scale 3^k and the $\mathscr{H}^2$-estimate eq. (6.39), yielding

$$\mathbb{E}_\rho\left[\sum_{z\in\mathcal{Z}_{n,k}} \int_{z+\square_k} |\mathsf{S}_{n,k}\nabla w - \mathsf{S}_{n,k+1}\nabla w|^2 \, d\mu\right] \leqslant \mathbb{E}_\rho\left[\int_{\square_n} |\nabla w - \mathsf{S}_{n,k+1}\nabla w|^2 \, d\mu\right] \leqslant C3^{2k}\mathbb{E}_\rho[u^2].$$

We have thus shown that

$$\text{eq. (6.40)-b} \leqslant C(\mathbb{E}_\rho[u^2])^{\frac{1}{2}} \left(\sum_{k=0}^{n-1} 3^k \left(\mathbb{E}_\rho\left[\int_{\square_n} |\mathsf{S}_{n,k}\nabla u|^2 \, d\mu\right]\right)^{\frac{1}{2}}\right). \tag{6.46}$$

For eq. (6.40)-c, we use eq. (6.28) and the Cauchy-Schwarz inequality to get that

$$\begin{aligned}\text{eq. (6.40)-c} &= \mathbb{E}_\rho\left[\int_{\square_n} (\mathsf{S}_{n,n}\nabla w)\cdot(\mathsf{S}_{n,n}\nabla u) \, d\mu\right] \\ &\leqslant \left(\mathbb{E}_\rho\left[\int_{\square_n} |\mathsf{S}_{n,n}\nabla w|^2 \, d\mu\right]\right)^{\frac{1}{2}} \left(\mathbb{E}_\rho\left[\int_{\square_n} |\mathsf{S}_{n,n}\nabla u|^2 \, d\mu\right]\right)^{\frac{1}{2}}.\end{aligned}$$

To treat the term $\mathbb{E}_\rho\left[\int_{\square_n} |\mathsf{S}_{n,n}\nabla w|^2 \, d\mu\right]$, we define the random affine function

$$\mathsf{p} := \frac{(\mathsf{S}_{n,n}\nabla w)(\mu,0)}{|(\mathsf{S}_{n,n}\nabla w)(\mu,0)|}, \qquad \ell_{\mathsf{p},\square_n} := \int_{\square_n} \mathsf{p}\cdot x \, d\mu(x). \tag{6.47}$$

Notice that here p is random, but when the particles in $\square_n$ move within $\square_n$, it does mot change the value; more precisely, the slope p is $\mathcal{G}_{n,n}$-measureable. We test $\ell_{\mathsf{p},\square_n}$ with eq. (6.38),

$$\begin{aligned}\mathbb{E}_\rho\left[u\ell_{\mathsf{p},\square_n} \,\middle|\, \mathcal{G}_{n,n}\right] &= \mathbb{E}_\rho\left[\int_{\square_n} \nabla w\cdot\mathsf{p} \, d\mu \,\middle|\, \mathcal{G}_{n,n}\right] \\ &= \mathbb{E}_\rho\left[\int_{\square_n} \nabla w \, d\mu \,\middle|\, \mathcal{G}_{n,n}\right]\cdot\mathsf{p} \\ &= \int_{\square_n} (\mathsf{S}_{n,n}\nabla w)\cdot\mathsf{p} \, d\mu.\end{aligned}$$

Recalling the definition in eq. (6.47), we obtain that

$$\begin{aligned}\int_{\square_n} |\mathsf{S}_{n,n}\nabla w| \, d\mu &= \mathbb{E}_\rho\left[u\ell_{\mathsf{p},\square_n} \,\middle|\, \mathcal{G}_{n,n}\right] \\ &\leqslant \left(\mathbb{E}_\rho\left[u^2 \,\middle|\, \mathcal{G}_{n,n}\right]\right)^{\frac{1}{2}} \left(\mathbb{E}_\rho\left[\ell^2_{\mathsf{p},\square_n} \,\middle|\, \mathcal{G}_{n,n}\right]\right)^{\frac{1}{2}} \\ &\leqslant C\sqrt{\mu(\square_n)}3^n \left(\mathbb{E}_\rho\left[u^2 \,\middle|\, \mathcal{G}_{n,n}\right]\right)^{\frac{1}{2}},\end{aligned}$$

where in the last step, we use a direct calculation of $\left(\mathbb{E}_\rho\left[\ell^2_{\mathsf{p},\square_n} \,\middle|\, \mathcal{G}_{n,n}\right]\right)^{\frac{1}{2}}$, and where the constant C may depend on d. Since $\mathsf{S}_{n,n}\nabla w$ is constant for every point in $\square_n$, we have shown that

$$\sqrt{\mu(\square_n)}|\mathsf{S}_{n,n}\nabla w|(\mu,0) \leqslant C3^n \left(\mathbb{E}_\rho\left[u^2 \,\middle|\, \mathcal{G}_{n,n}\right]\right)^{\frac{1}{2}}.$$

We thus obtain that

$$\begin{aligned}\mathbb{E}_\rho\left[\int_{\square_n} |\mathsf{S}_{n,n}\nabla w|^2 \, d\mu\right] &= \mathbb{E}_\rho\left[\mu(\square_n)|\mathsf{S}_{n,n}\nabla w|^2(\mu,0)\right] \\ &\leqslant C3^{2n}\mathbb{E}_\rho\left[\mathbb{E}_\rho\left[u^2 \,\middle|\, \mathcal{G}_{n,n}\right]\right] \\ &= C3^{2n}\mathbb{E}_\rho[u^2],\end{aligned}$$

and therefore

$$\text{eq. (6.40)-c} \leqslant C3^n (\mathbb{E}_\rho[u^2])^{\frac{1}{2}} \left(\mathbb{E}_\rho \left[\int_{\square_n} |\mathsf{S}_{n,n} \nabla u|^2 \, \mathrm{d}\mu \right] \right)^{\frac{1}{2}}. \tag{6.48}$$

We now combine eq. (6.40), (6.44), (6.46), and (6.48), to obtain eq. (6.36). □

6.3.4 Caccioppoli inequality

For every bounded open set $U \subseteq \mathbb{R}^d$, we define the space of $\mathbf{a}$-harmonic functions on $\mathcal{M}_\delta(\mathbb{R}^d)$ by

$$\mathcal{A}(U) := \left\{ u \in \mathscr{H}^1(U) : \ \forall \varphi \in \mathscr{H}^1_0(U), \ \mathbb{E}_\rho \left[\int_U \nabla u \cdot \mathbf{a} \nabla \varphi \, \mathrm{d}\mu \right] = 0 \right\}. \tag{6.49}$$

Recalling that, for any two bounded open sets $V \subseteq U$, we have $\mathscr{H}^1(U) \subseteq \mathscr{H}^1(V)$ and $\mathscr{H}^1_0(V) \subseteq \mathscr{H}^1_0(U)$, we see that $\mathcal{A}(U) \subseteq \mathcal{A}(V)$. For the classical Caccioppoli inequality, a standard proof is as follows: we multiply the harmonic function by a cutoff function, and then use this as a test function against the harmonic function itself. Adapting this argument to our space of particle configurations is not immediate. A naive approach would be to introduce a cutoff that brings the value of the function to zero whenever a particle approaches the boundary of the domain. But proceeding in this way is a very bad idea, since as we increase the size of the domain, there will essentially always be some particles near the boundary. We will instead rely on a suitable averaging procedure for particles that fall outside of a given region, using the localization operators defined in Subsection 6.2.3. Notice that our goal thus is not to bring the function to zero as a particle approaches the boundary of the box. Rather, it is only to produce a function that stops depending on the position of a particle that progressively approaches the boundary of the domain, in agreement with our definition of the space $\mathscr{H}^1_0(U)$ (and departing from the traditional definition of the Sobolev H^1_0 spaces).

Proposition 6.3.6 (Modified Caccioppoli inequality)**.** *There exist $\theta(d,\Lambda) \in (0,1)$, $C(d,\Lambda) < \infty$, and $R_0(d,\Lambda) < \infty$ such that for every $r \geqslant R_0$ and $u \in \mathcal{A}(Q_{3r})$, we have*

$$\mathbb{E}_\rho \left[\frac{1}{\rho |Q_r|} \int_{Q_r} \nabla (\mathsf{A}_{r+2} u) \cdot \mathbf{a} \nabla (\mathsf{A}_{r+2} u) \, \mathrm{d}\mu \right] \leqslant \frac{C}{r^2 \rho |Q_{3r}|} \mathbb{E}_\rho[u^2] + \theta \mathbb{E}_\rho \left[\frac{1}{\rho |Q_{3r}|} \int_{Q_{3r}} \nabla u \cdot \mathbf{a} \nabla u \, \mathrm{d}\mu \right]. \tag{6.50}$$

Remark. Inequality eq. (6.50) controls the norm of the gradient of a harmonic function in the small cube Q_r by a sum of terms involving the norm of the gradient in the larger cube Q_{3r}. This does not seem to be useful at first glance. However, the key point is that the multiplicative factor θ is smaller than one.

The proof of Proposition 6.3.6 will be divided into two steps. In the first step, provided by the lemma below, we prove a weaker Caccioppoli inequality, without the normalization of the volume. In the second step we use an iterative argument to improve the result and obtain Proposition 6.3.6.

Recall that $\mathsf{A}_{s,\varepsilon}$ is the regularized localization operator defined in eq. (6.9).

Lemma 6.3.3 (Weak Caccioppoli inequality). *Fix* $\theta'(\Lambda) := \frac{2\Lambda}{2\Lambda+1} \in (0,1)$. *For every* $r > 0$, $s \geqslant r+2, \varepsilon > 0$ *and* $u \in \mathcal{A}(Q_{s+\varepsilon})$, *we have*

$$\frac{\theta'}{2\varepsilon^2}\mathbb{E}_\rho[(\mathsf{A}_s u)^2] + \mathbb{E}_\rho\left[\int_{Q_r} \nabla(\mathsf{A}_{s,\varepsilon}u)\cdot \mathbf{a}\nabla(\mathsf{A}_{s,\varepsilon}u)\,\mathrm{d}\mu\right] \leqslant \theta'\left(\frac{1}{2\varepsilon^2}\mathbb{E}_\rho[(\mathsf{A}_{s+\varepsilon}u)^2] + \mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}} \nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]\right). \tag{6.51}$$

Proof. The proof of this lemma borrows some elements from [135, Lemma 4.8]; in both settings, the main point is to construct and analyze an appropriate "cut-off" version of the function u. We use the function $\widetilde{\mathsf{A}}_{s,\varepsilon}u \in \mathscr{H}_0^1(Q_{s+\varepsilon})$ defined in eq. (6.11) as a cut-off of the function u and test it against $u \in \mathcal{A}(Q_{s+\varepsilon})$ to get

$$\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}} \nabla(\widetilde{\mathsf{A}}_{s,\varepsilon}u)\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right] = 0. \tag{6.52}$$

Combining this with the decomposition

$$\begin{aligned}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}} \nabla(\widetilde{\mathsf{A}}_{s,\varepsilon}u)\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right] &= \underbrace{\mathbb{E}_\rho\left[\int_{Q_{s-2}} \nabla(\widetilde{\mathsf{A}}_{s,\varepsilon}u)\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]}_{\text{eq. (6.53)-a}} \\ &\quad + \underbrace{\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-2}} \nabla(\widetilde{\mathsf{A}}_{s,\varepsilon}u)\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]}_{\text{eq. (6.53)-b}} \\ &\quad + \underbrace{\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s} \nabla(\widetilde{\mathsf{A}}_{s,\varepsilon}u)\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]}_{\text{eq. (6.53)-c}},\end{aligned} \tag{6.53}$$

we obtain that

$$\text{eq. (6.53)-a} \leqslant |\text{eq. (6.53)-b}| + |\text{eq. (6.53)-c}|. \tag{6.54}$$

We now study each of these three terms. For the first term eq. (6.53)-a, since $x \in Q_{s-2}$, the coefficient $\mathbf{a}$ is $\mathcal{F}_{Q_s}$-measurable. We can thus use eq. (6.8), eq. (6.11), (6.13) and (6.12) to get

$$\begin{aligned}\text{eq. (6.53)-a} &= \frac{2}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s-2}}\int_0^\varepsilon (\varepsilon-t)\mathsf{A}_{s+t}(\nabla u)\cdot\mathbf{a}\nabla u\,\mathrm{d}t\,\mathrm{d}\mu\right] \\ &= \frac{2}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s-2}}\int_0^\varepsilon (\varepsilon-t)\mathbb{E}_\rho\left[\mathsf{A}_{s+t}(\nabla u)\cdot\mathbf{a}\mathsf{A}_{s+t}(\nabla u)\,|\mathcal{F}_{\bar{Q}_{s+t}}\right]\mathrm{d}t\,\mathrm{d}\mu\right] \\ &= \mathbb{E}_\rho\left[\int_{Q_{s-2}} \nabla(\mathsf{A}_{s,\varepsilon}u)\cdot\mathbf{a}\nabla(\mathsf{A}_{s,\varepsilon}u)\,\mathrm{d}\mu\right].\end{aligned}$$

We then apply eq. (6.13) for the second term eq. (6.53)-b. We notice that $\mathbf{a}$ is no longer $\mathcal{F}_{Q_s}$-measurable and use Young's inequality and the bound $\mathbf{a} \leqslant \Lambda \mathsf{Id}$

$$\begin{aligned}|\text{eq. (6.53)-b}| &= \frac{2}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-2}}\int_0^\varepsilon (\varepsilon-t)\mathsf{A}_{s+t}(\nabla u)\cdot\mathbf{a}\nabla u\,\mathrm{d}t\,\mathrm{d}\mu\right] \\ &\leqslant \frac{\Lambda}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-2}}\int_0^\varepsilon (\varepsilon-t)\left(|\mathsf{A}_{s+t}(\nabla u)|^2 + |\nabla u|^2\right)\mathrm{d}t\,\mathrm{d}\mu\right].\end{aligned}$$

For the part with conditional expectation, we use Jensen's inequality and the uniform bound $\mathsf{Id} \leqslant \mathbf{a} \leqslant \Lambda \mathsf{Id}$

$$\frac{\Lambda}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-2}}\int_0^\varepsilon (\varepsilon - t)\,|\mathsf{A}_{s+t}(\nabla u)|^2 \,\mathrm{d}t\,\mathrm{d}\mu\right] \leqslant \frac{\Lambda}{2}\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-2}} |\nabla u|^2 \,\mathrm{d}\mu\right]$$
$$\leqslant \frac{\Lambda}{2}\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-2}} \nabla u\cdot \mathbf{a}\nabla u\,\mathrm{d}\mu\right].$$

This concludes that $|\text{eq. (6.53)-b}| \leqslant \Lambda\mathbb{E}_\rho\left[\int_{Q_s\setminus Q_{s-2}} \nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]$.

For the third term eq. (6.53)-c, we use eq. (6.13) and obtain

$$|\text{eq. (6.53)-c}| \leqslant \text{eq. (6.53)-c1} + \text{eq. (6.53)-c2}$$
$$\text{eq. (6.53)-c1} = \frac{2}{\varepsilon^2}\left|\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}\int_{\tau(x)-s}^{\varepsilon}(\varepsilon - t)\mathsf{A}_{s+t}(\nabla u)\cdot\mathbf{a}\nabla u\,\mathrm{d}t\,\mathrm{d}\mu\right]\right|$$
$$\text{eq. (6.53)-c2} = \frac{2}{\varepsilon^2}\left|\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))\Delta_{\tau(x)}(\mathsf{A}u)\vec{\mathbf{n}}(x)\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]\right|.$$

The part of eq. (6.53)-c1 can be treated as that of eq. (6.53)-b, so that

$$\text{eq. (6.53)-c1} \leqslant \Lambda\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}\nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right].$$

We study the part eq. (6.53)-c2 using Young's inequality with a parameter $\beta > 0$ to be fixed later:

$$\begin{aligned}
&\frac{2}{\varepsilon^2}\left|\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))\Delta_{\tau(x)}(\mathsf{A}u)\vec{\mathbf{n}}(x)\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]\right| \\
\leqslant & \frac{\Lambda}{\beta\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))|\Delta_{\tau(x)}(\mathsf{A}u)|^2\,\mathrm{d}\mu\right] + \frac{\beta\Lambda}{\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))|\nabla u|^2\,\mathrm{d}\mu\right] \\
\leqslant & \frac{\Lambda}{\beta\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))|\Delta_{\tau(x)}(\mathsf{A}u)|^2\,\mathrm{d}\mu\right] + \frac{\beta\Lambda}{\varepsilon}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}\nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right].
\end{aligned} \tag{6.55}$$

The first term above will be responsible for producing the $\mathscr{L}^2$ term on the right side of eq. (6.51). We start by writing

$$\frac{\Lambda}{\beta\varepsilon^2}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}(s+\varepsilon-\tau(x))|\Delta_{\tau(x)}(\mathsf{A}u)|^2\,\mathrm{d}\mu\right]$$
$$= \frac{\Lambda}{\beta\varepsilon^2}\mathbb{E}_\rho\left[\sum_{s\leqslant\tau\leqslant s+\varepsilon}(s+\varepsilon-\tau)|\Delta_\tau(\mathsf{A}u)|^2\right],$$

where on the right side, the sum is over all τ's that are jump discontinuites for $(\mathsf{A}_s u)_{s\geqslant 0}$. Recalling the definition of the bracket process $([\mathsf{A}u]_s)_{s\geqslant 0}$ defined in eq. (6.7), we use Fubini's lemma and the $\mathscr{L}^2$ isometry $\mathbb{E}_\rho\left[[\mathsf{A}u]_s\right] = \mathbb{E}_\rho\left[(\mathsf{A}_s u)^2\right]$:

$$\begin{aligned}\frac{\Lambda}{\beta\varepsilon^2}\mathbb{E}_\rho\left[\sum_{s\leqslant\tau\leqslant s+\varepsilon}(s+\varepsilon-\tau)|\Delta_\tau(\mathsf{A}u)|^2\right] &= \frac{\Lambda}{\beta\varepsilon^2}\mathbb{E}_\rho\left[\sum_{s\leqslant\tau\leqslant s+\varepsilon}\int_s^{s+\varepsilon}\mathbf{1}_{\{\tau\leqslant t\leqslant s+\varepsilon\}}\,\mathrm{d}t|\Delta_\tau(\mathsf{A}u)|^2\right]\\ &= \frac{\Lambda}{\beta\varepsilon^2}\mathbb{E}_\rho\left[\int_s^{s+\varepsilon}\sum_{s\leqslant\tau\leqslant t}|\Delta_\tau(\mathsf{A}u)|^2\,\mathrm{d}t\right]\\ &= \frac{\Lambda}{\beta\varepsilon^2}\mathbb{E}_\rho\left[\int_s^{s+\varepsilon}\left([\mathsf{A}u]_t-[\mathsf{A}u]_s\right)\mathrm{d}t\right]\\ &= \frac{\Lambda}{\beta\varepsilon^2}\int_s^{s+\varepsilon}\left(\mathbb{E}_\rho\left[(\mathsf{A}_t u)^2\right]-\mathbb{E}_\rho\left[(\mathsf{A}_s u)^2\right]\right)\mathrm{d}t\\ &\leqslant \frac{\Lambda}{\beta\varepsilon}\left(\mathbb{E}_\rho\left[(\mathsf{A}_{s+\varepsilon}u)^2\right]-\mathbb{E}_\rho\left[(\mathsf{A}_s u)^2\right]\right).\end{aligned}$$

Putting this estimate back into eq. (6.55), we conclude the estimating of the term eq. (6.53)-c2, obtaining

$$\text{eq. (6.53)-c2} \leqslant \frac{\Lambda}{\beta\varepsilon}\left(\mathbb{E}_\rho\left[(\mathsf{A}_{s+\varepsilon}u)^2\right]-\mathbb{E}_\rho\left[(\mathsf{A}_s u)^2\right]\right)+\frac{\beta\Lambda}{\varepsilon}\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_s}\nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right].$$

By choosing $\beta=\varepsilon$, recalling eq. (6.54), and that $r\leqslant s-2$, we can combine this estimate with those of eq. (6.53)-a, eq. (6.53)-b, and eq. (6.53)-c1 to get

$$\begin{aligned}&\frac{\Lambda}{\varepsilon^2}\mathbb{E}_\rho\left[(\mathsf{A}_s u)^2\right]+\mathbb{E}_\rho\left[\int_{Q_r}\nabla(\mathsf{A}_{s,\varepsilon}u)\cdot\mathbf{a}\nabla(\mathsf{A}_{s,\varepsilon}u)\,\mathrm{d}\mu\right]\\ &\qquad\leqslant\frac{\Lambda}{\varepsilon^2}\mathbb{E}_\rho\left[(\mathsf{A}_{s+\varepsilon}u)^2\right]+2\Lambda\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}\setminus Q_r}\nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right].\end{aligned}$$

We now proceed with a hole-filling argument: adding $2\Lambda\mathbb{E}_\rho\left[\int_{Q_r}\nabla\mathsf{A}_{s,\varepsilon}u\cdot\mathbf{a}\nabla\mathsf{A}_{s,\varepsilon}u\,\mathrm{d}\mu\right]$ to both sides of the equation above, and using Jensen's inequality, we obtain

$$\begin{aligned}&\frac{\Lambda}{\varepsilon^2}\mathbb{E}_\rho\left[(\mathsf{A}_s u)^2\right]+(2\Lambda+1)\mathbb{E}_\rho\left[\int_{Q_r}\nabla(\mathsf{A}_{s,\varepsilon}u)\cdot\mathbf{a}\nabla(\mathsf{A}_{s,\varepsilon}u)\,\mathrm{d}\mu\right]\\ &\qquad\leqslant\frac{\Lambda}{\varepsilon^2}\mathbb{E}_\rho\left[(\mathsf{A}_{s+\varepsilon}u)^2\right]+2\Lambda\mathbb{E}_\rho\left[\int_{Q_{s+\varepsilon}}\nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right].\end{aligned}$$

Dividing both sides by $(2\Lambda+1)$, and setting $\theta' := \frac{2\Lambda}{2\Lambda+1}$, we obtain the desired inequality eq. (6.51). □

We remark that eq. (6.51) does not imply directly eq. (6.50). For example, let $r>2$ and we choose $s=2r$ and $\varepsilon=r$ in eq. (6.51), then with a normalization of volume we get

$$\begin{aligned}&\frac{\theta'}{2r^2\rho|Q_r|}\mathbb{E}_\rho[(\mathsf{A}_{2r}u)^2]+\mathbb{E}_\rho\left[\frac{1}{\rho|Q_r|}\int_{Q_r}\nabla(\mathsf{A}_{2r,r}u)\cdot\mathbf{a}\nabla(\mathsf{A}_{2r,r}u)\,\mathrm{d}\mu\right]\\ &\qquad\leqslant 3^d\theta'\left(\frac{1}{2r^2\rho|Q_{3r}|}\mathbb{E}_\rho[(\mathsf{A}_{3r}u)^2]+\mathbb{E}_\rho\left[\frac{1}{\rho|Q_{3r}|}\int_{Q_{3r}}\nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]\right).\end{aligned}$$

Then another factor 3^d will be added, and we typically do not have $3^d\theta'\in(0,1)$, since we recall that $\theta'=\frac{2\Lambda}{2\Lambda+1}$.

Proof of Proposition 6.3.6. We apply Lemma 6.3.3 iteratively, with very small increments of the volume. Let $\delta > 0$ to be fixed later, and choose $s = (1+\delta)r, \varepsilon = \delta r$. For convenience, we assume that r is sufficiently large that

$$s = (1+\delta)r \geqslant r + 2, \quad \text{that is} \quad r \geqslant 2\delta^{-1}. \tag{6.56}$$

Equation (6.51) and Jensen's inequality give us that, provided $(1+2\delta)r \leqslant 3r$,

$$\begin{aligned}\mathbb{E}_\rho\left[\frac{1}{\rho|Q_r|}\int_{Q_r}\nabla(\mathsf{A}_{(1+\delta)r,\delta r}u)\cdot\mathbf{a}\nabla(\mathsf{A}_{(1+\delta)r,\delta r}u)\,\mathrm{d}\mu\right]\\ \leqslant \widetilde{\theta}\left(\frac{1}{2(\delta r)^2\rho|Q_{(1+2\delta)r}|}\mathbb{E}_\rho[u^2] + \mathbb{E}_\rho\left[\frac{1}{\rho|Q_{(1+2\delta)r}|}\int_{Q_{(1+2\delta)r}}\nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right]\right),\end{aligned} \tag{6.57}$$

with $\widetilde{\theta} = (1+2\delta)^d\theta'$. We choose the constant $\delta > 0$ sufficiently small that $\widetilde{\theta} < 1$. In order to obtain eq. (6.50), we will now apply eq. (6.57) iteratively, from the cube Q_r to the larger cube Q_{3r}.

We give the details for this argument—see also Figure 6.2 for an illustration. We plan to use eq. (6.57) $(N+1)$ times, and and let $\delta \in (0,1)$, $N \in \mathbb{N}$ satisfy

$$\widetilde{\theta} = (1+2\delta)^d\theta' < 1, \qquad (1+2\delta)^{N+1} = 3. \tag{6.58}$$

Then we set the scale and the $\mathbf{a}$-harmonic functions in every scale

$$\begin{cases} r_n = (1+2\delta)^n r & 0 \leqslant n \leqslant N+1, \\ u_{N+1} = u & , \\ u_n = \mathsf{A}_{(1+\delta)r_n,\delta r_n}u_{n+1} & 0 \leqslant n \leqslant N. \end{cases} \tag{6.59}$$

We can prove by induction that $u_n \in \mathcal{A}(Q_{r_n})$ under the condition eq. (6.56). Then for every $0 \leqslant n \leqslant N$, we apply eq. (6.57) from u_n on Q_{r_n} to u_{n+1} on $Q_{r_{n+1}}$

$$\begin{aligned}\mathbb{E}_\rho\left[\frac{1}{\rho|Q_{r_n}|}\int_{Q_{r_n}}\nabla u_n\cdot\mathbf{a}\nabla u_n\,\mathrm{d}\mu\right]\\ \leqslant \widetilde{\theta}\left(\frac{1}{2(\delta r_n)^2\rho|Q_{(1+2\delta)r_n}|}\mathbb{E}_\rho[(u_{n+1})^2] + \mathbb{E}_\rho\left[\frac{1}{\rho|Q_{r_{n+1}}|}\int_{Q_{r_{n+1}}}\nabla u_{n+1}\cdot\mathbf{a}\nabla u_{n+1}\,\mathrm{d}\mu\right]\right).\end{aligned} \tag{6.60}$$

Iterating on eq. (6.60) until $u_{N+1} = u$ on Q_{3r}, we get

$$\begin{aligned}\mathbb{E}_\rho\left[\frac{1}{\rho|Q_r|}\int_{Q_r}\nabla u_0\cdot\mathbf{a}\nabla u_0\,\mathrm{d}\mu\right]\\ \leqslant \left(\frac{3^d}{2}\sum_{n=0}^{N}(1+2\delta)^{-2n}\right)\frac{1}{(\delta r)^2\rho|Q_{3r}|}\mathbb{E}_\rho[u^2] + (\widetilde{\theta})^{N+1}\mathbb{E}_\rho\left[\frac{1}{\rho|Q_{3r}|}\int_{Q_{3r}}\nabla u\cdot\mathbf{a}\nabla u\,\mathrm{d}\mu\right].\end{aligned}$$

We notice that u_0 can be seen as as a weighted sum of $\mathsf{A}_{s'}u$, for scales s' satisfying $s' \geqslant (1+\delta)r \geqslant r+2$, by eq. (6.56). So we apply once Jensen's inequality for u_0 and obtain eq. (6.50) by setting

$$C(d,\Lambda) \coloneqq \frac{3^d}{2\delta^2}\sum_{n=0}^{N}(1+2\delta)^{-2n}, \qquad \theta \coloneqq (\widetilde{\theta})^{N+1}.$$ □

Although we will not use this later, we now give more explicit estimates for the choice of the parameters in the proof above, resulting from the conditions listed in eq. (6.56) and eq. (6.58). It suffices to pick an integer N larger than $\left\lfloor \frac{d\log 3}{\log(1+\frac{1}{2\Lambda})} \right\rfloor$, and then in eq. (6.58) use $\delta := \frac{1}{2}(3^{\frac{1}{N+1}} - 1)$ to fix δ, and in eq. (6.56) we require $r \geqslant 2\delta^{-1}$, which gives the condition for the minimal scale R_0. A possible choice is the following

$$N := 2\left\lfloor \frac{d\log 3}{\log\left(1+\frac{1}{2\Lambda}\right)} \right\rfloor + 1, \quad \delta := \frac{1}{2}(3^{\frac{1}{N+1}} - 1) \simeq \frac{1}{8d\Lambda}, \quad R_0 := 2\delta^{-1} \simeq 16 d\Lambda,$$
$$\widetilde{\theta} := \theta'(1+2\delta)^d \simeq \left(1 + \frac{1}{2\Lambda}\right)^{-\frac{1}{2}}, \quad \theta := (\widetilde{\theta})^{N+1} \simeq 3^{-d}, \quad C := \frac{3^d}{2\delta^2}\sum_{n=0}^{N}(1+2\delta)^{-2n} \simeq 2^8 3^d d^3 \Lambda^3.$$

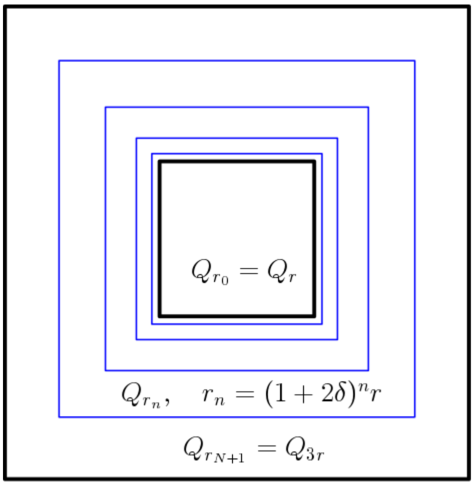

Figure 6.2: An illustration of the iterative argument for the proof of Proposition 6.3.6. Since Lemma 6.3.3 can imply Proposition 6.3.6 only for a comparison from scale r to $(1+2\delta)r$ with δ very small, we add many intermediate scales $r_n = (1+2\delta)^n r$ between r and $3r$.

6.4 Subadditive quantities

We aim to adapt the strategy in [25, Chapter 2] for our model in continuum configuration space. In this section, we define several subadditive quantities, denoted by ν, ν^*, J, and develop their elementary properties. We then we use them and a renormalization argument to obtain a quantitative rate of convergence for $\bar{\mathbf{a}}$ in Section 6.5.

6.4.1 Subadditive quantities ν and ν^*

For every bounded domain $U \subseteq \mathbb{R}^d$ and $p, q \in \mathbb{R}^d$, we define the affine function in U with slope p by

$$\ell_{p,U}(\mu) := \int_U p \cdot x \, \mathrm{d}\mu(x), \tag{6.61}$$

and introduce the subadditive quantities

$$\begin{aligned} \nu(U,p) &:= \inf_{v \in \ell_{p,U} + \mathscr{H}_0^1(U)} \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla v \cdot \mathbf{a}\nabla v \, \mathrm{d}\mu\right], \\ \nu^*(U,q) &:= \sup_{u \in \mathscr{H}^1(U)} \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(-\frac{1}{2}\nabla u \cdot \mathbf{a}\nabla u + q \cdot \nabla u\right) \mathrm{d}\mu\right]. \end{aligned} \tag{6.62}$$

The quantity ν can be thought of as the average energy per unit volume of the solution which matches with the behavior of the affine function $\ell_{p,U}$ when a particle leaves the domain U. The quantity ν^* is analogous to a Neumann problem with prescribed average flux of q. As will be seen below, the quantities ν and ν^* are approximately dual to one another; the quality of this approximation as the domain U grows to $\mathbb{R}^d$ will be central to the proof of Theorem 6.2.1. If the matrix $\mathbf{a}$ were constant, then by eq. (6.6) the minimizer for $\nu(U,p)$ would be $\ell_{p,U}$, and we would have $\nu(U,p) = \frac{1}{2}p\cdot\mathbf{a}p$; and similarly, were $\mathbf{a}$ constant, we would have $\nu^*(U,q) = \frac{1}{2}q\cdot\mathbf{a}^{-1}q$.

We start by recording elementary properties satisfied by ν and ν^*. We recall that $\mathcal{G}_U = \sigma(\mu(U), \mu \llcorner (\mathbb{R}^d \backslash U))$. For every $r > 0$, we denote by $B_r(U)$ the r-enlargement of U, that is, $B_r(U) := \{x \in \mathbb{R}^d : \operatorname{dist}(x,U) < r\}$.

Proposition 6.4.1 (Elementary properties of ν and ν^*). *The following properties hold for every bounded domain $U \subseteq \mathbb{R}^d$ with Lipschitz boundary and $p,q,p',q' \in \mathbb{R}^d$.*

(1) There exists a unique solution for the optimization problem in the definition of $\nu(U,p)$ that satisfies $\mathbb{E}_\rho[v - \ell_{p,U}] = 0$; we denote it by $v(\cdot,U,p)$. For the optimization problem in the definition of $\nu^*(U,q)$, there exists a maximizer $u(\cdot,U,q)$ that is $\mathcal{F}_{B_1(U)}$-measurable and such that $\mathbb{E}_\rho[u \,|\, \mathcal{G}_U] = 0$. They are $\mathbf{a}$-harmonic functions on U, i.e. $v(\cdot,U,p), u(\cdot,U,q) \in \mathcal{A}(U)$.

(2) There exist two $d \times d$ symmetric matrices $\bar{\mathbf{a}}(U)$ and $\bar{\mathbf{a}}_*(U)$ such that

$$\nu(U,p) = \frac{1}{2}p\cdot\bar{\mathbf{a}}(U)p, \qquad \nu^*(U,q) = \frac{1}{2}q\cdot\bar{\mathbf{a}}_*^{-1}(U)q, \tag{6.63}$$

and these matrices satisfy $\mathsf{Id} \leqslant \bar{\mathbf{a}}(U) \leqslant \Lambda\mathsf{Id}$ and $\mathsf{Id} \leqslant \bar{\mathbf{a}}_*(U) \leqslant \Lambda\mathsf{Id}$. Moreover,

$$p'\cdot\bar{\mathbf{a}}(U)p = \mathbb{E}_U\left[\frac{1}{\rho|U|}\int_U p'\cdot\mathbf{a}(\mu,x)\nabla v(\mu,x,U,p)\,\mathrm{d}\mu(x)\right], \tag{6.64}$$

$$q'\cdot\bar{\mathbf{a}}_*^{-1}(U)q = \mathbb{E}_U\left[\frac{1}{\rho|U|}\int_U q'\cdot\nabla u(\mu,x,U,q)\,\mathrm{d}\mu(x)\right]. \tag{6.65}$$

(3) Slope: $v(\mu,U,p)$ satisfies

$$\mathbb{E}_\rho\left[\fint_U \nabla v(\mu,x,U,p)\,\mathrm{d}\mu(x)\;\middle|\;\mathcal{G}_U\right] = \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla v(\mu,x,U,p)\,\mathrm{d}\mu(x)\right] = p. \tag{6.66}$$

For the function $u(\cdot,U,q)$, there exists a $d\times d$ symmetric matrix $\mathsf{Id} \leqslant \mathbf{a}_*(U;\mathcal{G}_U) \leqslant \Lambda\mathsf{Id}$ such that

$$\mathbb{E}_\rho\left[\fint_U \nabla u(\mu,x,U,q)\,\mathrm{d}\mu(x)\;\middle|\;\mathcal{G}_U\right] = \mathbf{a}_*^{-1}(U;\mathcal{G}_U)q, \tag{6.67}$$

and $\bar{\mathbf{a}}_*^{-1}(U) = \frac{1}{\rho|U|}\mathbb{E}_\rho[\mathbf{a}_*^{-1}(U;\mathcal{G}_U)\mu(U)]$, so that

$$\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla u(\mu,x,U,q)\,\mathrm{d}\mu(x)\right] = \bar{\mathbf{a}}_*^{-1}(U)q. \tag{6.68}$$

(4) Quadratic response: for every $v' \in \ell_{p,U} + \mathscr{H}_0^1(U)$, we have

$$\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla(v' - v(\mu,U,p))\cdot\mathbf{a}\nabla(v' - v(\mu,U,p))\,\mathrm{d}\mu\right] = \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla v'\cdot\mathbf{a}\nabla v'\,\mathrm{d}\mu\right] - \nu(U,p). \tag{6.69}$$

Similarly, for every $u' \in \mathscr{H}^1(U)$, we have

$$\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla(u'-u(\mu,U,q))\cdot\mathbf{a}\nabla(u'-u(\mu,U,q))\,\mathrm{d}\mu\right] \\ = \nu^*(U,q) - \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\left(-\frac{1}{2}\nabla u'\cdot\mathbf{a}\nabla u' + q\cdot\nabla u'\right)\mathrm{d}\mu\right]. \quad (6.70)$$

(5) The quantities ν and ν^* are subadditive: for every $n\in\mathbb{N}$,

$$\nu(\square_{n+1},p)\leqslant\nu(\square_n,p), \qquad \nu^*(\square_{n+1},q)\leqslant\nu^*(\square_n,q). \quad (6.71)$$

Proof. We prove each of these points in turn.

(1)We study at first the maximizer for the problem $\nu^*(U,q)$. A first observation is that the maximizer can be found in $\mathcal{F}_{B_1(U)}$-measurable functions. Because for any $u\in\mathscr{H}^1(U)$, its conditional expectation $\mathbb{E}_\rho[u\,|\,\mathcal{F}_{B_1(U)}]$ reaches a larger value for the functional in $\nu^*(U,q)$. We use Jensen's inequality that

$$\begin{aligned}&\mathbb{E}_\rho\left[\int_U\left(-\frac{1}{2}\nabla\mathbb{E}_\rho[u\,|\,\mathcal{F}_{B_1(U)}]\cdot\mathbf{a}\nabla\mathbb{E}_\rho[u\,|\,\mathcal{F}_{B_1(U)}] + q\cdot\nabla\mathbb{E}_\rho[u\,|\,\mathcal{F}_{B_1(U)}]\right)\mathrm{d}\mu\right]\\ =&\mathbb{E}_\rho\left[\mathbb{E}_\rho\left[\int_U\left(-\frac{1}{2}\mathbb{E}_\rho[\nabla u\,|\,\mathcal{F}_{B_1(U)}]\cdot\mathbf{a}\mathbb{E}_\rho[\nabla u\,|\,\mathcal{F}_{B_1(U)}] + q\cdot\mathbb{E}_\rho[\nabla u\,|\,\mathcal{F}_{B_1(U)}]\right)\mathrm{d}\mu\,\middle|\,\mathcal{F}_{B_1(U)}\right]\right]\\ \geqslant&\mathbb{E}_\rho\left[\int_U\left(-\frac{1}{2}\nabla u\cdot\mathbf{a}\nabla u + q\cdot\nabla u\right)\mathrm{d}\mu\right].\end{aligned}$$

By a variational calculus, we know the characterization of a maximizer with elliptic equation that for any $\phi\in\mathscr{H}^1(U)$

$$\mathbb{E}_\rho\left[\int_U \nabla u\cdot\mathbf{a}\nabla\phi\,\mathrm{d}\mu\right] = \mathbb{E}_\rho\left[\int_U q\cdot\nabla\phi\,\mathrm{d}\mu\right]. \quad (6.72)$$

Similarly to the discussion in the proof of Proposition 6.3.4, we know that a solution for this problem also satisfies the more precise equation

$$\mathbb{E}_\rho\left[\int_U \nabla u\cdot\mathbf{a}\nabla\phi\,\mathrm{d}\mu\,\middle|\,\mathcal{G}_U\right] = \mathbb{E}_\rho\left[\int_U q\cdot\nabla\phi\,\mathrm{d}\mu\,\middle|\,\mathcal{G}_U\right], \quad (6.73)$$

and we can define its solution in the space

$$W = \{f\in\mathscr{H}^1(U) : \mathbb{E}_\rho[f\,|\,\mathcal{G}_U] = 0\}.$$

In this space, we have

$$\mathbb{E}_\rho[f^2\,|\,\mathcal{G}_U] \leqslant C\,\mathrm{diam}(U)^2\mathbb{E}_\rho\left[\int_U |\nabla f|^2\,\mathrm{d}\mu\,\middle|\,\mathcal{G}_U\right],$$

by the Poincaré inequality Proposition 6.3.2. Then the coercivity on left hand side in eq. (6.73) is ensured and we can apply the Lax-Milgram theorem. We call this maximizer $u(\mu,U,q)$. Testing eq. (6.72) with $\phi\in\mathscr{H}^1_0(U)$, eq. (6.6) implies that its right hand side is 0, so we have $u(\mu,U,q)\in\mathcal{A}(U)$.

Then we turn to $\nu(U,p)$. By a first order variation calculus, we know that a minimizer v for $\nu(U,p)$ is characterized by an elliptic equation that for any $\phi\in\mathscr{H}^1_0(U)$

$$\mathbb{E}_\rho\left[\int_U \nabla(v-\ell_{p,U})\cdot\mathbf{a}\nabla\phi\,\mathrm{d}\mu\right] = \mathbb{E}_\rho\left[\int_U -p\cdot\mathbf{a}\nabla\phi\,\mathrm{d}\mu\right]. \quad (6.74)$$

We remark that one cannot treat this equation as eq. (6.72), because $\mathbb{E}_\rho[v|\mathcal{G}_U]$ is not an element in $\mathscr{H}_0^1(U)$ and we cannot subtract it. On the other hand, we can apply the Lax-Milgram theorem on the space

$$V = \{f \in \mathscr{H}_0^1(U) : \mathbb{E}_\rho[f] = 0\},$$

to define the unique solution $v - \ell_{p,U} \in V$. We notice that the right hand side of eq. (6.74) is clearly a bounded linear functional, and the coercivity of the left hand side of eq. (6.74) is ensured by the Poincaré inequality Proposition 6.3.3 on V. We denote this minimizer by $v(\mu, U, p)$, and eq. (6.74) implies that $v(\mu, U, p) \in \mathcal{A}(U)$.

(2) We test at first eq. (6.74) with $v(\mu, U, p') - \ell_{p',U} \in \mathscr{H}_0^1(U)$ and obtain that

$$\mathbb{E}_\rho\left[\int_U \nabla v(\mu, x, U, p) \cdot \mathbf{a}(\mu, x) \nabla v(\mu, x, U, p') \,\mathrm{d}\mu(x)\right] = \mathbb{E}_\rho\left[\int_U \nabla v(\mu, x, U, p) \cdot \mathbf{a}(\mu, x) p' \,\mathrm{d}\mu(x)\right], \quad (6.75)$$

and this implies $(p, p') \mapsto \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla v(\mu, x, U, p) \cdot \mathbf{a}(\mu, x) \nabla v(\mu, x, U, p') \,\mathrm{d}\mu(x)\right]$ is a bilinear map $p \cdot \bar{\mathbf{a}}(U) p'$. This definition with eq. (6.75) proves eq. (6.64). We let $p = p'$ and obtain that $\nu(U, p) = \frac{1}{2} p \cdot \bar{\mathbf{a}}(U) p$. To obtain the bound of $\bar{\mathbf{a}}(U)$, we use the bound of $\mathbf{a}$ and the definition of eq. (6.62)

$$\inf_{v \in \ell_{p,U} + \mathscr{H}_0^1(U)} \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}|\nabla v|^2 \,\mathrm{d}\mu\right] \leqslant \nu(U, p) = \frac{1}{2} p \cdot \bar{\mathbf{a}}(U) p \leqslant \inf_{v \in \ell_{p,U} + \mathscr{H}_0^1(U)} \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{\Lambda}{2}|\nabla v|^2 \,\mathrm{d}\mu\right].$$

We can check that $\ell_{p,U}$ is the minimizer for $\inf_{v \in \ell_{p,U} + \mathscr{H}_0^1(U)} \mathbb{E}_\rho\left[\int_{\mathbb{R}^d} \frac{\Lambda}{2}|\nabla v|^2 \,\mathrm{d}\mu\right]$, then it concludes the proof of the bound $\mathsf{Id} \leqslant \bar{\mathbf{a}}(U) \leqslant \Lambda \mathsf{Id}$.

The same argument works for $\nu^*(U, q)$. We test eq. (6.72) with $u(\mu, U, q')$ and obtain that

$$\mathbb{E}_\rho\left[\int_U \nabla u(\mu, x, U, q) \cdot \mathbf{a}(\mu, x) \nabla u(\mu, x, U, q') \,\mathrm{d}\mu(x)\right] = \mathbb{E}_\rho\left[\int_U q \cdot \nabla u(\mu, x, U, q') \,\mathrm{d}\mu(x)\right]. \quad (6.76)$$

This proves that $(q, q') \mapsto \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla u(\mu, x, U, q) \cdot \mathbf{a}(\mu, x) \nabla u(\mu, x, U, q')\right]$ is also bilinear and we denote it by $q \cdot \bar{\mathbf{a}}_*^{-1}(U) q'$, and this also concludes eq. (6.65). Then we put $q' = q$ and eq. (6.76) in the definition of eq. (6.62) that

$$\begin{aligned}
&\nu^*(U, q) \\
&= \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(-\frac{1}{2}\nabla u(\mu, x, U, q) \cdot \mathbf{a}(\mu, x) \nabla u(\mu, x, U, q) + q \cdot \nabla u(\mu, x, U, q)\right) \mathrm{d}\mu(x)\right] \\
&= \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla u(\mu, x, U, q) \cdot \mathbf{a}(\mu, x) \nabla u(\mu, x, U, q) \,\mathrm{d}\mu(x)\right] \\
&= \frac{1}{2} q \cdot \bar{\mathbf{a}}_*^{-1}(U) q.
\end{aligned}$$

This proves the bilinear map expression for $\nu^*(U,q)$. Concerning the bound for the matrix $\bar{\mathbf{a}}_*^{-1}(U)$, we use the bound for $\mathbf{a}$ and the equations above to obtain that

$$\begin{aligned}\sup_{u\in\mathscr{H}^1(U)} \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\left(-\frac{\Lambda}{2}|\nabla u|^2+q\cdot\nabla u\right)\mathrm{d}\mu\right] &\leqslant \nu^*(U,q)\\ &\leqslant \sup_{u\in\mathscr{H}^1(U)} \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\left(-\frac{1}{2}|\nabla u|^2+q\cdot\nabla u\right)\mathrm{d}\mu\right]. \qquad (6.77)\end{aligned}$$

One can check for the lower bound, $\ell_{\frac{q}{\Lambda},U}$ attains the maximum and for the upper bound it is $\ell_{q,U}$ that attains the maximum. Then we put the expression $\nu^*(U,q)=\frac{1}{2}q\cdot\bar{\mathbf{a}}_*^{-1}(U)q$ and obtain that

$$\frac{\Lambda^{-1}}{2}|q|^2\leqslant\nu^*(U,q)=\frac{1}{2}q\cdot\bar{\mathbf{a}}_*^{-1}(U)q\leqslant\frac{1}{2}|q|^2,$$

which implies the bound for $\bar{\mathbf{a}}_*(U)$.

(3) The slope identity eq. (6.66) for $v(\mu,U,p)$ is directly the result from eq. (6.6) that

$$\mathbb{E}_\rho\left[\fint_U\nabla v(\mu,x,U,p)\,\mathrm{d}\mu(x)\;\middle|\;\mu(U)\right]=\mathbb{E}_\rho\left[\fint_U p\,\mathrm{d}\mu\;\middle|\;\mu(U)\right]=p.$$

For the function $u(\mu,U,q)$, the identity eq. (6.68) comes directly from eq. (6.65), but conditioned $\mathcal{G}_U$, the averaged slope is not $\bar{\mathbf{a}}_*^{-1}(U)q$. In fact, we recall that $u(\mu,U,q)$ is also the conditioned maximizer for eq. (6.73), so we can define the matrix $\mathbf{a}_*^{-1}(U;\mathcal{G}_U)$, the quenched slope eq. (6.67). The estimate for this matrix is then obtained by repeating the argument in eq. (6.63) for eq. (6.73).

(4) We test eq. (6.74) with $(v'-\ell_{p,U})$ and put it in the left hand side of eq. (6.69)

$$\begin{aligned}&\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\frac{1}{2}\nabla(v'-v(\cdot,U,p))\cdot\mathbf{a}\nabla(v'-v(\cdot,U,p))\,\mathrm{d}\mu\right]\\ =&\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\frac{1}{2}\nabla v'\cdot\mathbf{a}\nabla v'\,\mathrm{d}\mu\right]+\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\frac{1}{2}\nabla v(\cdot,U,p)\cdot\mathbf{a}\nabla v(\cdot,U,p)\,\mathrm{d}\mu\right]\\ &\quad-\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\nabla v'\cdot\mathbf{a}\nabla v(\cdot,U,p)\,\mathrm{d}\mu\right] \qquad (6.78)\\ =&\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\frac{1}{2}\nabla v'\cdot\mathbf{a}\nabla v'\,\mathrm{d}\mu\right]+\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\frac{1}{2}\nabla v(\cdot,U,p)\cdot\mathbf{a}\nabla v(\cdot,U,p)\,\mathrm{d}\mu\right]\\ &\quad-\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U p\cdot\mathbf{a}\nabla v(\cdot,U,p)\,\mathrm{d}\mu\right].\end{aligned}$$

The term $\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U p\cdot\mathbf{a}\nabla v(\cdot,U,p)\,\mathrm{d}\mu\right]$ also appears on the right side of eq. (6.64) with $p=p'$, thus we obtain that

$$\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U p\cdot\mathbf{a}\nabla v(\cdot,U,p)\,\mathrm{d}\mu\right]=p\cdot\bar{\mathbf{a}}(U)p=\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\nabla v(\cdot,U,p)\cdot\mathbf{a}\nabla v(\cdot,U,p)\,\mathrm{d}\mu\right],$$

and we put it back to eq. (6.78) to conclude for the validity of eq. (6.69).

Similarly, we develop the left hand side of eq. (6.70) as eq. (6.78), and use eq. (6.72) with $\phi = u'$ to treat the inner product term of u' and $u(\cdot, U, q)$.

$$\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla u' \cdot \mathbf{a}\nabla u(\cdot, U, q)\,\mathrm{d}\mu\right] = \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla u' \cdot q\,\mathrm{d}\mu\right].$$

We put this term in the left hand side of eq. (6.70) and use the bilinear map expression of $\nu^*(U, q)$ to obtain that

$$\begin{aligned}
&\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla(u' - u(\cdot, U, q)) \cdot \mathbf{a}\nabla(u' - u(\cdot, U, q))\,\mathrm{d}\mu\right]\\
=&\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla u' \cdot \mathbf{a}\nabla u'\,\mathrm{d}\mu\right] + \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla u(\cdot, U, q) \cdot \mathbf{a}\nabla u(\cdot, U, q)\,\mathrm{d}\mu\right]\\
&\quad - \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla u' \cdot \mathbf{a}\nabla u(\cdot, U, q)\,\mathrm{d}\mu\right]\\
=&\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla u' \cdot \mathbf{a}\nabla u'\,\mathrm{d}\mu\right] + \nu^*(U, q) - \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla u' \cdot q\,\mathrm{d}\mu\right].
\end{aligned}$$

This concludes the proof of eq. (6.70).

(5) For $\nu(\square_{n+1}, p)$, we test the associated variational problem with a sub-minimiser $v' = \sum_{z \in \mathcal{Z}_{n+1,n}} v(\cdot, z + \square_n, p)$, which is an element of $\ell_{p,\square_{n+1}} + \mathscr{H}_0^1(\square_{n+1})$, so that

$$\begin{aligned}
\nu(\square_{n+1}, p) &\leqslant \mathbb{E}_\rho\left[\frac{1}{\rho|\square_{n+1}|}\int_{\square_{n+1}} \nabla v' \cdot \mathbf{a}\nabla v'\,\mathrm{d}\mu\right]\\
&= 3^{-d}\sum_{z \in \mathcal{Z}_{n+1,n}} \mathbb{E}_\rho\left[\frac{1}{\rho|\square_n|}\int_{z+\square_n} \nabla v(\cdot, z + \square_n, p) \cdot \mathbf{a}\nabla v(\cdot, z + \square_n, p)\,\mathrm{d}\mu\right]\\
&= \nu(\square_n, p).
\end{aligned}$$

In the last step, we also use the stationarity of the coefficient field $\mathbf{a}$.

For $\nu^*(\square_{n+1}, p)$, we also use that, for every $z \in \mathcal{Z}_{n+1,n}$, we have the inclusion $\mathscr{H}^1(\square_{n+1}) \subseteq \mathscr{H}^1(z + \square_n)$, so its unit energy on every small cube $z + \square_n$ is less than the maximum $\nu^*(z + \square_n, p)$, thus

$$\begin{aligned}
&\nu^*(\square_{n+1}, q)\\
=&3^{-d}\sum_{z \in \mathcal{Z}_{n+1,n}} \mathbb{E}_\rho\left[\frac{1}{\rho|\square_n|}\int_{z+\square_n} -\frac{1}{2}\nabla u(\cdot, \square_{n+1}, q) \cdot \mathbf{a}\nabla u(\cdot, \square_{n+1}, q) + q \cdot \nabla u(\cdot, \square_{n+1}, q)\,\mathrm{d}\mu\right]\\
\leqslant&3^{-d}\sum_{z \in \mathcal{Z}_{n+1,n}} \nu^*(z + \square_n, q)\\
=&\nu^*(\square_n, q).
\end{aligned}$$

□

6.4.2 Subadditive quantity J

We now study the quantity J defined by

$$\begin{aligned}
J(U, p, q) :&= \nu(U, p) + \nu^*(U, q) - p \cdot q\\
&= \frac{1}{2}p \cdot \bar{\mathbf{a}}(U)p + \frac{1}{2}q \cdot \bar{\mathbf{a}}_*^{-1}(U)q - p \cdot q.
\end{aligned} \tag{6.79}$$

By the properties of ν and ν^*, the quantity J is also subadditive. We briefly explain why this quantity will be convenient for our purposes. If the functions $\nu(U, \cdot)$ and $\nu^*(U, \cdot)$ were exactly convex dual of one another, then we would have that $J \geqslant 0$ and that for every $p \in \mathbb{R}^d$, the infimum of $J(U,p,\cdot)$ is zero. This would correspond to the situation in which $\bar{\mathbf{a}}(U)$ and $\bar{\mathbf{a}}_*(U)$ are equal, and for every $p \in \mathbb{R}^d$, we would in fact have that $J(U,p,\bar{\mathbf{a}}(U)p) = 0$. Instead, we will show below that, for any symmetric matrix $\mathsf{Id} \leqslant \widetilde{\mathbf{a}} \leqslant \Lambda\mathsf{Id}$, we have

$$|\widetilde{\mathbf{a}} - \bar{\mathbf{a}}(U)| + |\widetilde{\mathbf{a}} - \mathbf{a}_*(U)| \leqslant \sup_{p\in B_1} C(J(U,p,\widetilde{\mathbf{a}}p))^{\frac{1}{2}}.$$

The right side of the inequality above can be thought of as a measure of the defect in the convex duality relationship between ν and ν^*. For $U = \square_m$ and using $\widetilde{\mathbf{a}} = \bar{\mathbf{a}}_*(\square_m)$, we obtain that

$$|\bar{\mathbf{a}}_*(\square_m) - \bar{\mathbf{a}}(\square_m)| \leqslant \sup_{p\in B_1} C(J(U,p,\bar{\mathbf{a}}_*(\square_m)p))^{\frac{1}{2}}.$$

Since we know that $\{\bar{\mathbf{a}}(\square_m)\}_{m\geqslant 0}$ is a decreasing sequence while $\{\bar{\mathbf{a}}_*(\square_m)\}_{m\geqslant 0}$ is a increasing sequence from eq. (6.63) and eq. (6.71), each sequence has a limit. Therefore, once we prove a rate of convergence to zero for $J(U,p,\bar{\mathbf{a}}_*(\square_m)p)$, we get that the two limits coincide, and also a rate for the convergence of $\{\bar{\mathbf{a}}(\square_m)\}_{m\geqslant 0}$.

The rest of this section will present this strategy in details. We establish at first a variational description for the quantity J and the properties mentioned above.

Lemma 6.4.1. *(1) For every $p,q \in \mathbb{R}^d$, we have the variational representation*

$$J(U,p,q) = \sup_{w\in\mathcal{A}(U)} \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(-\frac{1}{2}\nabla w\cdot \mathbf{a}\nabla w - p\cdot\mathbf{a}\nabla w + q\cdot\nabla w\right)\mathrm{d}\mu\right]. \tag{6.80}$$

(2) We have that $J(U,p,q) \geqslant 0$ and $\bar{\mathbf{a}}(U) \geqslant \bar{\mathbf{a}}_*(U)$.

(3) There exists a constant $C(d,\Lambda) < \infty$ such that and for every symmetric matrix $\widetilde{\mathbf{a}}$ satisfying $\mathsf{Id} \leqslant \widetilde{\mathbf{a}} \leqslant \Lambda\mathsf{Id}$, we have

$$|\widetilde{\mathbf{a}} - \bar{\mathbf{a}}(U)| + |\widetilde{\mathbf{a}} - \bar{\mathbf{a}}_*(U)| \leqslant C\sup_{p\in B_1}(J(U,p,\widetilde{\mathbf{a}}p))^{\frac{1}{2}}. \tag{6.81}$$

Proof. (1) We start by rewriting the expression of $J(U,p,q)$ using the definition of $\nu^*(U,q)$ and the quadratic expression of $\nu(U,p)$. Noting also that the maximizer of $\nu^*(U,q)$ belongs to $\mathcal{A}(U)$, we can write

$$\begin{aligned} J(U,p,q) = \mathbb{E}_\rho&\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla v(\cdot,U,p)\cdot\mathbf{a}\nabla v(\cdot,U,p)\,\mathrm{d}\mu\right] \\ &+ \sup_{u\in\mathcal{A}(U)}\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\left(-\frac{1}{2}\nabla u\cdot\mathbf{a}\nabla u + q\cdot\nabla u\right)\mathrm{d}\mu\right] - p\cdot q. \end{aligned} \tag{6.82}$$

We claim that for any $u \in \mathcal{A}(U)$, with $w := u - v(\cdot,U,p)$, we have

$$\begin{aligned} \mathbb{E}_\rho&\left[\frac{1}{\rho|U|}\int_U\left(\frac{1}{2}\nabla v(\cdot,U,p)\cdot\mathbf{a}\nabla v(\cdot,U,p) - \frac{1}{2}\nabla u\cdot\mathbf{a}\nabla u + q\cdot\nabla u\right)\mathrm{d}\mu\right] - p\cdot q \\ &= \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\left(-\frac{1}{2}\nabla w\cdot\mathbf{a}\nabla w - p\cdot\mathbf{a}\nabla w + q\cdot\nabla w\right)\mathrm{d}\mu\right]. \end{aligned} \tag{6.83}$$

To prove it, we can develop the right hand side of eq. (6.83)

$$\begin{aligned}
&\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\left(-\frac{1}{2}\nabla w\cdot\mathbf{a}\nabla w - p\cdot\mathbf{a}\nabla w + q\cdot\nabla w\right)\mathrm{d}\mu\right]\\
=&\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\left(\frac{1}{2}\nabla v(\cdot,U,p)\cdot\mathbf{a}\nabla v(\cdot,U,p) - \frac{1}{2}\nabla u\cdot\mathbf{a}\nabla u + q\cdot\nabla u\right)\mathrm{d}\mu\right] - p\cdot q \\
&\quad + \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla\left(v(\cdot,U,p)-\ell_{p,U}\right)\cdot\left(\mathbf{a}\nabla u - \mathbf{a}\nabla v(\cdot,U,p) - q\ \right)\mathrm{d}\mu\right].
\end{aligned} \tag{6.84}$$

Because $(v(\cdot,U,p)-\ell_{p,U})\in\mathscr{H}^1_0(U)$, we apply $u, v(\cdot,U,p)\in\mathcal{A}(U)$ and eq. (6.6), the last line of eq. (6.84) is 0 and we prove eq. (6.83). Then we take the maximum as eq. (6.82) and obtain the definition eq. (6.80).

(2) The properties that $J(U,p,q)\geqslant 0$ comes from the definition of $\nu^*(U,q)$: we test the functional in the definition of $\nu^*(U,q)$ with the minimizer $v(\cdot,U,p)$ of $\nu(U,p)$ and obtain that

$$\begin{aligned}
&\nu^*(U,q)\\
&\geqslant \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U\left(-\frac{1}{2}\nabla v(\mu,x,U,p)\cdot\mathbf{a}(\mu,x)\nabla v(\mu,x,U,p) + q\cdot\nabla v(\mu,x,U,p)\right)\mathrm{d}\mu\right]\\
&= p\cdot q - \nu(U,p),
\end{aligned}$$

so that

$$J(U,p,q) = \nu(U,p) + \nu^*(U,q) - p\cdot q \geqslant 0.$$

Then we test $J(U,p,q)\geqslant 0$ with that $q=\bar{\mathbf{a}}_*(U)p$ and obtain that

$$0\leqslant J(U,p,\bar{\mathbf{a}}_*(U)p) = \frac{1}{2}p\cdot\bar{\mathbf{a}}(U)p + \frac{1}{2}(\bar{\mathbf{a}}_*(U)p)\cdot\bar{\mathbf{a}}_*^{-1}(U)(\bar{\mathbf{a}}_*(U)p) - p\cdot\bar{\mathbf{a}}_*(U)p,$$

and therefore $\bar{\mathbf{a}}(U)\geqslant\bar{\mathbf{a}}_*(U)$.

(3) Using this property, we have

$$\begin{aligned}
J(U,p,q) &= \frac{1}{2}p\cdot\bar{\mathbf{a}}(U)p + \frac{1}{2}q\cdot\bar{\mathbf{a}}_*^{-1}(U)q - p\cdot q\\
&\geqslant \frac{1}{2}p\cdot\bar{\mathbf{a}}(U)p + \frac{1}{2}q\cdot\bar{\mathbf{a}}^{-1}(U)q - p\cdot q\\
&= \frac{1}{2}(\bar{\mathbf{a}}(U)p - q)\bar{\mathbf{a}}^{-1}(U)\cdot(\bar{\mathbf{a}}(U)p - q).
\end{aligned}$$

We put $q=\widetilde{\mathbf{a}}p$ and obtain $|\bar{\mathbf{a}}(U)-\widetilde{\mathbf{a}}|\leqslant C\sup_{p\in B_1}(J(U,p,\widetilde{\mathbf{a}}p))^{\frac{1}{2}}$. The proof of the statement concerning $|\bar{\mathbf{a}}_*(U)-\widetilde{\mathbf{a}}|$ is similar. □

In view of the definition of J, this functional enjoys properties similar to those described in Proposition 6.4.1 for ν and ν^*.

Proposition 6.4.2 (Elementary properties of J)**.** *For every bounded domain $U\subseteq\mathbb{R}^d$ with Lipschitz boundary and $p,q\in\mathbb{R}^d$, the quantity $J(U,p,q)$ defined in eq. (6.79) satisfies the following properties:*

(1) Characterization of optimizer: the optimization problem in eq. (6.80) admits a unique solution $v(\cdot,U,p,q)\in\mathscr{H}^1(U)$ such that $\mathbb{E}_\rho[v(\cdot,U,p,q)\,|\,\mathcal{G}_U]=0$. This solution is such that for every $w\in\mathcal{A}(U)$,

$$\mathbb{E}_\rho\left[\int_U \nabla v(\cdot,U,p,q)\cdot\mathbf{a}\nabla w\,\mathrm{d}\mu\right] = \mathbb{E}_\rho\left[\int_U(-p\cdot\mathbf{a}\nabla w + q\cdot\nabla w)\,\mathrm{d}\mu\right], \tag{6.85}$$

and $(p,q) \mapsto v(\cdot,U,p,q)$ a linear map. The function $v(\cdot,U,p,q)$ can be expressed in terms of the optimizers in eq. (6.62) as

$$v(\mu,U,p,q) = u(\mu,U,q) - v(\mu,U,p) - \mathbb{E}_\rho[u(\mu,U,q) - v(\mu,U,p) \,|\, \mathcal{G}_U]. \tag{6.86}$$

We have the quadratic expression

$$J(U,p,q) = \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \frac{1}{2}\nabla v(\cdot,U,p,q)\cdot \mathbf{a}\nabla v(\cdot,U,p,q)\,\mathrm{d}\mu\right]. \tag{6.87}$$

(2) Slope: $v(\cdot,U,p,q)$ satisfies

$$\begin{aligned} \mathbb{E}_\rho\left[\fint_U \nabla v(\cdot,U,p,q)\,\mathrm{d}\mu \,\middle|\, \mathcal{G}_U\right] &= \mathbf{a}_*^{-1}(U;\mathcal{G}_U)q - p, \\ \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla v(\cdot,U,p,q)\,\mathrm{d}\mu\right] &= \bar{\mathbf{a}}_*^{-1}(U)q - p, \end{aligned} \tag{6.88}$$

where the matrix $\mathbf{a}_*^{-1}(U;\mathcal{G}_U)$ is defined in eq. (6.67).

(3) Quadratic response: for every $w \in \mathcal{A}(U)$, we have

$$\begin{aligned} &\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(\frac{1}{2}\nabla(w - v(\cdot,U,p,q))\cdot\mathbf{a}\nabla(w - v(\cdot,U,p,q))\right)\mathrm{d}\mu\right] \\ &\qquad = J(U,p,q) - \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(-\frac{1}{2}\nabla w\cdot\mathbf{a}\nabla w - p\cdot\mathbf{a}\nabla w + q\cdot\nabla w\right)\mathrm{d}\mu\right]. \end{aligned} \tag{6.89}$$

(4) Subadditivity: for every $n \in \mathbb{N}$, we have

$$J(\square_{n+1},p,q) \leqslant J(\square_n,p,q). \tag{6.90}$$

Proof. (1) The equation eq. (6.85) comes directly from the first order variation calculus. The proof of the existence and uniqueness of the solution $v(\cdot,U,p,q)$ is similar as the one for $\nu^*(U,q)$. Equation (6.85) also implies that the map $(p,q) \mapsto v(\cdot,U,p,q)$ is linear because for any $p_1,p_2,q_1,q_2 \in \mathbb{R}^d$, and any $w \in \mathcal{A}(U)$ we have

$$\begin{aligned} &\mathbb{E}_\rho\left[\int_U \nabla v(\cdot,U,p_1+p_2,q_1+q_2)\cdot\mathbf{a}\nabla w\,\mathrm{d}\mu\right] \\ &\quad = \mathbb{E}_\rho\left[\int_U -(p_1+p_2)\cdot\mathbf{a}\nabla w + (q_1+q_2)\cdot\nabla w\,\mathrm{d}\mu\right] \\ &\quad = \mathbb{E}_\rho\left[\int_U \nabla(v(\cdot,U,p_1,q_1)+v(\cdot,U,p_2,q_2))\cdot\mathbf{a}\nabla w\,\mathrm{d}\mu\right]. \end{aligned}$$

Then $(v(\cdot,U,p_1,q_1)+v(\cdot,U,p_2,q_2))$ is also a solution for the problem eq. (6.85) with parameter (p_1+p_2,q_1+q_2). Notice that we have

$$\mathbb{E}_\rho[(v(\cdot,U,p_1,q_1)+v(\cdot,U,p_2,q_2))\,|\,\mathcal{G}_U] = 0,$$

it implies $v(\mu,U,p_1+p_2,q_1+q_2) = v(\mu,U,p_1,q_1)+v(\mu,U,p_2,q_2)$ and the linearity of the map.

The exact expression of $v(\mu,U,P,q)$ comes from the equivalent definition eq. (6.80) of $J(U,p,q)$ and its proof. We put $v(\mu,U,p,q)$ in the first order variation eq. (6.85)

$$\begin{aligned} &\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U (-p\cdot\mathbf{a}\nabla v(\cdot,U,p,q) + q\cdot\nabla v(\cdot,U,p,q))\,\mathrm{d}\mu\right] \\ &\qquad\qquad = \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \nabla v(\cdot,U,p,q)\cdot\mathbf{a}\nabla v(\cdot,U,p,q)\,\mathrm{d}\mu\right]. \end{aligned}$$

Then we put this equation into eq. (6.80) to get eq. (6.87).

(2) The slope identity eq. (6.88) comes from eq. (6.86), (6.66), (6.67), and (6.68).

(3) We use the expression in eq. (6.86) with $w := u' - v(\cdot, U, p)$, then we use the quadratic response for $\nu^*(U,q)$ eq. (6.70) that

$$\begin{aligned}
&\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(\frac{1}{2}\nabla(w - v(\cdot,U,p,q))\cdot \mathbf{a}\nabla(w - v(\cdot,U,p,q))\right) \mathrm{d}\mu\right] \\
&\quad = \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(\frac{1}{2}\nabla(u' - u(\cdot,U,q))\cdot \mathbf{a}\nabla(u' - u(\cdot,U,q))\right) \mathrm{d}\mu\right] \\
&\quad = \nu^*(U,q) - \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(-\frac{1}{2}\nabla u'\cdot \mathbf{a}\nabla u' + q\cdot\nabla u'\right) \mathrm{d}\mu\right].
\end{aligned}$$

Then we add back the term $\nu(U,p)$ and it gives the desired result

$$\begin{aligned}
&\mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(\frac{1}{2}\nabla(w - v(\cdot,U,p,q))\cdot \mathbf{a}\nabla(w - v(\cdot,U,p,q))\right) \mathrm{d}\mu\right] \\
&\quad = J(U,p,q) - \left(\nu(U,p) + \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(-\frac{1}{2}\nabla u'\cdot \mathbf{a}\nabla u' + q\cdot\nabla u'\right) \mathrm{d}\mu\right] - p\cdot q\right) \\
&\quad = J(U,p,q) - \mathbb{E}_\rho\left[\frac{1}{\rho|U|}\int_U \left(-\frac{1}{2}\nabla w\cdot \mathbf{a}\nabla w - p\cdot\mathbf{a}\nabla w + q\cdot\nabla w\right) \mathrm{d}\mu\right].
\end{aligned}$$

(4) Equation (6.90) is a consequence of eq. (6.71) and eq. (6.79). □

We conclude this section with the following lemma.

Lemma 6.4.2 (Comparison between two scales)**.** *For every $n, k \in \mathbb{N}$ with $k \leqslant n$, and $p, q \in \mathbb{R}^d$, writing $v(U)$ as shorthand for $v(\cdot, U, p, q)$, we have*

$$\begin{aligned}
\frac{1}{|\mathcal{Z}_{n,k}|}\sum_{z\in\mathcal{Z}_{n,k}} \mathbb{E}_\rho\left[\frac{1}{\rho|\square_k|}\int_{z+\square_k} \frac{1}{2}|\nabla v(\square_n) - \nabla v(z+\square_k)|^2 \,\mathrm{d}\mu\right]& \\
\leqslant J(\square_k,p,q) - J(\square_n,p,q).& \qquad (6.91)
\end{aligned}$$

Proof. For any $z \in \mathcal{Z}_{n,k}$, since $v(\square_n) \in \mathcal{A}(z+\square_k)$, we use the quadratic response eq. (6.89) for $J(z+\square_k,p,q)$ that

$$\begin{aligned}
&\mathbb{E}_\rho\left[\frac{1}{\rho|\square_k|}\int_{z+\square_k} \frac{1}{2}|\nabla v(\square_n) - \nabla v(z+\square_k)|^2 \,\mathrm{d}\mu\right] \\
&\quad \leqslant \mathbb{E}_\rho\left[\frac{1}{\rho|\square_k|}\int_{z+\square_k} \frac{1}{2}(\nabla v(\square_n) - \nabla v(z+\square_k))\cdot\mathbf{a}(\nabla v(\square_n) - \nabla v(z+\square_k))\,\mathrm{d}\mu\right] \\
&\quad = J(z+\square_k,p,q) \\
&\qquad - \mathbb{E}_\rho\left[\frac{1}{\rho|\square_k|}\int_{z+\square_k} \left(-\frac{1}{2}\nabla v(\square_n)\cdot\mathbf{a}\nabla v(\square_n) - p\cdot\mathbf{a}\nabla v(\square_n) + q\cdot\nabla v(\square_n)\right) \mathrm{d}\mu\right].
\end{aligned}$$

We sum this expression over all $z \in \mathcal{Z}_{n,k}$ to obtain that

$$\begin{aligned}
&\frac{1}{|\mathcal{Z}_{n,k}|} \sum_{z \in \mathcal{Z}_{n,k}} \mathbb{E}_\rho \left[\frac{1}{\rho|\square_k|} \int_{z+\square_k} \frac{1}{2} |\nabla v(\square_n) - \nabla v(z+\square_k)|^2 \, \mathrm{d}\mu \right] \\
&\leqslant \frac{1}{|\mathcal{Z}_{n,k}|} \sum_{z \in \mathcal{Z}_{n,k}} \Bigg(J(z+\square_k, p, q) \\
&\quad - \mathbb{E}_\rho \left[\frac{1}{\rho|\square_k|} \int_{z+\square_k} \left(-\frac{1}{2} \nabla v(\square_n) \cdot \mathbf{a} \nabla v(\square_n) - p \cdot \mathbf{a} \nabla v(\square_n) + q \cdot \nabla v(\square_n) \right) \mathrm{d}\mu \right] \Bigg) \\
&= J(\square_k, p, q) - J(\square_n, p, q).
\end{aligned}$$

In the last step, we use the stationarity of J and also eq. (6.80) for $v(\square_n)$. □

6.5 Quantitative rate of convergence

We are now ready to prove Theorem 6.2.1. We decompose the argument into a series of four steps.

6.5.1 Step 1: setup

We use the shorthand $\bar{\mathbf{a}}_n := \bar{\mathbf{a}}_*(\square_n)$, so that by eq. (6.88), the average slope of the function $v(\cdot, \square_n, p, q)$ is $\bar{\mathbf{a}}_n^{-1} q - p$, in the sense that

$$\mathbb{E}_\rho \left[\frac{1}{\rho|\square_n|} \int_U \nabla v(\cdot, \square_n, p, q) \, \mathrm{d}\mu \right] = \bar{\mathbf{a}}_n^{-1} q - p. \tag{6.92}$$

We let τ_n denote a measure of the defect in the subadditivity of J, precisely,

$$\begin{aligned}
\tau_n &:= \sup_{p,q \in B_1} \left(J(\square_n, p, q) - J(\square_{n+1}, p, q) \right) \\
&= \sup_{p \in B_1} \left(\nu(\square_n, p) - \nu(\square_{n+1}, p) \right) + \sup_{q \in B_1} \left(\nu^*(\square_n, q) - \nu^*(\square_{n+1}, q) \right).
\end{aligned} \tag{6.93}$$

A direct corollary from eq. (6.93) is that for any integers $n < m$,

$$|\bar{\mathbf{a}}_n^{-1} - \bar{\mathbf{a}}_m^{-1}| = \sup_{q \in B_1} q \cdot (\bar{\mathbf{a}}_n^{-1} - \bar{\mathbf{a}}_m^{-1}) q = \sup_{q \in B_1} \left(\nu^*(\square_n, q) - \nu^*(\square_m, q) \right) \leqslant C \sum_{k=n}^{m-1} \tau_k. \tag{6.94}$$

We recall that $\{\bar{\mathbf{a}}(\square_m)\}_{m \geqslant 0}$ is decreasing and $\{\bar{\mathbf{a}}_*(\square_m)\}_{m \geqslant 0}$ is increasing, with the comparison $\bar{\mathbf{a}}_*(\square_m) \leqslant \bar{\mathbf{a}}(\square_m)$. From eq. (6.81), we know that

$$|\bar{\mathbf{a}}(\square_m) - \bar{\mathbf{a}}| \leqslant |\bar{\mathbf{a}}(\square_m) - \bar{\mathbf{a}}_*(\square_m)| \leqslant C \sup_{p \in B_1} \left(J(\square_m, p, \bar{\mathbf{a}}_m p) \right)^{\frac{1}{2}}.$$

From now on, we thus fix $p \in B_1$, and focus on estimating $J(\square_m, p, \bar{\mathbf{a}}_m p)$. We also assume without further notification that m is sufficiently large that $3^m \geqslant R_0$, for the constant R_0 appearing in Proposition 6.3.6. We use $\mathsf{A}_{3^m+2} v(\cdot, \square_{m+1}, p, \bar{\mathbf{a}}_m p)$ to compare with eq. (6.87) and apply the quadratic response eq. (6.89). In the rest of Step 1, we write $v(U)$ as a shorthand for $v(\cdot, U, p, \bar{\mathbf{a}}_m p)$, and decompose

$$\begin{aligned}
(J(\square_m, p, \bar{\mathbf{a}}_m p))^{\frac{1}{2}} &= \left(\mathbb{E}_\rho \left[\frac{1}{\rho|\square_m|} \int_{\square_m} \frac{1}{2} \nabla v(\square_m) \cdot \mathbf{a} \nabla v(\square_m) \, \mathrm{d}\mu \right] \right)^{\frac{1}{2}} \\
&\leqslant \text{eq. (6.95)-a} + \text{eq. (6.95)-b},
\end{aligned} \tag{6.95}$$

with

$$\begin{aligned}&\text{eq. (6.95)-a}\\&=\left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_m|}\int_{\square_m}\frac{1}{2}(\nabla v(\square_m)-\nabla \mathsf{A}_{3^m+2}v(\square_{m+1}))\cdot\mathbf{a}(\nabla v(\square_m)-\nabla \mathsf{A}_{3^m+2}v(\square_{m+1}))\,\mathrm{d}\mu\right]\right)^{\frac{1}{2}},\end{aligned}$$

and

$$\text{eq. (6.95)-b}=\left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_m|}\int_{\square_m}\frac{1}{2}\nabla \mathsf{A}_{3^m+2}v(\square_{m+1})\cdot\mathbf{a}\nabla \mathsf{A}_{3^m+2}v(\square_{m+1})\,\mathrm{d}\mu\right]\right)^{\frac{1}{2}}.$$

We treat the two terms separately. For eq. (6.95)-a, since $\mathsf{A}_{3^m+2}v(\square_{m+1})\in\mathcal{A}(\square_m)$ (see Proposition 6.A.1 for details), we use eq. (6.89) to get

$$\begin{aligned}&|\text{eq. (6.95)-a}|^2\\&\quad=J(\square_m,p,\bar{\mathbf{a}}_m p)-\mathbb{E}_\rho\left[\frac{1}{\rho|\square_m|}\int_{\square_m}\left(-\frac{1}{2}\nabla \mathsf{A}_{3^m+2}v(\square_{m+1})\cdot\mathbf{a}\nabla \mathsf{A}_{3^m+2}v(\square_{m+1})\right)\mathrm{d}\mu\right]\\&\qquad-\mathbb{E}_\rho\left[\frac{1}{\rho|\square_m|}\int_{\square_m}(-p\cdot\mathbf{a}\nabla \mathsf{A}_{3^m+2}v(\square_{m+1})+\bar{\mathbf{a}}_m p\cdot\nabla \mathsf{A}_{3^m+2}v(\square_{m+1}))\,\mathrm{d}\mu\right].\end{aligned}$$

Using Jensen's inequality, we have

$$\begin{aligned}&\mathbb{E}_\rho\left[\int_{\square_m}\left(\frac{1}{2}\nabla \mathsf{A}_{3^m+2}v(\square_{m+1})\cdot\mathbf{a}\nabla \mathsf{A}_{3^m+2}v(\square_{m+1})\right)\mathrm{d}\mu\right]\\&\qquad\leqslant\mathbb{E}_\rho\left[\int_{\square_m}\left(\frac{1}{2}\nabla v(\square_{m+1})\cdot\mathbf{a}\nabla v(\square_{m+1})\right)\mathrm{d}\mu\right],\end{aligned}$$

and the conditional expectation also implies that

$$\begin{aligned}&\mathbb{E}_\rho\left[\int_{\square_m}(-p\cdot\mathbf{a}\nabla \mathsf{A}_{3^m+2}v(\square_{m+1})+\bar{\mathbf{a}}_m p\cdot\nabla \mathsf{A}_{3^m+2}v(\square_{m+1}))\,\mathrm{d}\mu\right]\\&\qquad=\mathbb{E}_\rho\left[\int_{\square_m}(-p\cdot\mathbf{a}\nabla v(\square_{m+1})+\bar{\mathbf{a}}_m p\cdot\nabla v(\square_{m+1}))\,\mathrm{d}\mu\right].\end{aligned}$$

Thus we combine these terms with the quadratic response eq. (6.89) to obtain

$$\begin{aligned}|\text{eq. (6.95)-a}|^2&\leqslant J(\square_m,p,\bar{\mathbf{a}}_m p)-\mathbb{E}_\rho\left[\frac{1}{\rho|\square_m|}\int_{\square_m}\left(-\frac{1}{2}\nabla v(\square_{m+1})\cdot\mathbf{a}\nabla v(\square_{m+1})\right)\mathrm{d}\mu\right]\\&\quad-\mathbb{E}_\rho\left[\frac{1}{\rho|\square_m|}\int_{\square_m}(-p\cdot\mathbf{a}\nabla v(\square_{m+1})+\bar{\mathbf{a}}_m p\cdot\nabla v(\square_{m+1}))\,\mathrm{d}\mu\right]\\&=\mathbb{E}_\rho\left[\frac{1}{\rho|\square_m|}\int_{\square_m}\left(\frac{1}{2}\nabla(v(\square_{m+1})-v(\square_m))\cdot\mathbf{a}\nabla(v(\square_{m+1})-v(\square_m))\right)\mathrm{d}\mu\right],\end{aligned}$$

and we use Lemma 6.4.2 between $\square_m$ and $\square_{m+1}$ to get

$$|\text{eq. (6.95)-a}|^2\leqslant 3^d(J(\square_m,p,\bar{\mathbf{a}}_m p)-J(\square_{m+1},p,\bar{\mathbf{a}}_m p))\leqslant C(d,\Lambda)\tau_m,\tag{6.96}$$

where the quantity τ_m is defined in eq. (6.93).

For the term eq. (6.95)-b, we can apply the modified Caccioppoli inequality eq. (6.50): there exist two finite positive constants $C(d,\Lambda)$ and $\theta(d,\Lambda)\in(0,1)$ such that

$$\begin{aligned}&\mathbb{E}_\rho\left[\frac{1}{\rho|\square_m|}\int_{\square_m}\nabla(\mathsf{A}_{3^m+2}v(\square_{m+1}))\cdot\mathbf{a}\nabla(\mathsf{A}_{3^m+2}v(\square_{m+1}))\,\mathrm{d}\mu\right]\\&\quad\leqslant\frac{C}{3^{2m}\rho|\square_{m+1}|}\mathbb{E}_\rho[(v(\square_{m+1}))^2]+\theta\mathbb{E}_\rho\left[\frac{1}{\rho|\square_{m+1}|}\int_{\square_{m+1}}\nabla v(\square_{m+1})\cdot\mathbf{a}\nabla v(\square_{m+1})\,\mathrm{d}\mu\right].\end{aligned}\tag{6.97}$$

Using eq. (6.87), we see that the averaged gradient term on the right side of eq. (6.97) is $J(\square_{m+1}, p, \bar{\mathbf{a}}_m p)$, and eq. (6.90) asserts that $J(\square_{m+1}, p, \bar{\mathbf{a}}_m p) \leqslant J(\square_m, p, \bar{\mathbf{a}}_m p)$. Therefore, we get the bound for eq. (6.95)-b

$$|\text{eq. (6.95)-b}|^2 \leqslant \frac{C}{3^{2m}\rho|\square_{m+1}|}\mathbb{E}_\rho[(v(\square_{m+1}))^2] + \theta J(\square_m, p, \bar{\mathbf{a}}_m p). \tag{6.98}$$

We put eq. (6.96) and eq. (6.98) back to eq. (6.95), obtaining

$$\begin{aligned}(J(\square_m, p, \bar{\mathbf{a}}_m p))^{\frac{1}{2}} &\leqslant C\tau_m^{\frac{1}{2}} + \left(\frac{C}{3^{2m}\rho|\square_{m+1}|}\|v(\square_{m+1})\|_{\mathscr{L}^2}^2 + \theta J(\square_m, p, \bar{\mathbf{a}}_m p)\right)^{\frac{1}{2}} \\ &\leqslant C\tau_m^{\frac{1}{2}} + \frac{C}{3^m(\rho|\square_{m+1}|)^{\frac{1}{2}}}\|v(\square_{m+1})\|_{\mathscr{L}^2} + \theta^{\frac{1}{2}}(J(\square_m, p, \bar{\mathbf{a}}_m p))^{\frac{1}{2}}.\end{aligned}$$

Since $\theta < 1$, this gives

$$J(\square_m, p, \bar{\mathbf{a}}_m p) \leqslant C\left(\tau_m + \frac{1}{3^{2m}\rho|\square_{m+1}|}\|v(\mu, \square_{m+1}, p, \bar{\mathbf{a}}_m p)\|_{\mathscr{L}^2}^2\right). \tag{6.99}$$

6.5.2 Step 2: flatness estimate

In this step, we estimate the $\mathscr{L}^2$-flatness of optimizers of J. Notice that, using the result of Lemma 6.101 with $v(\cdot, \square_{m+1}, p, \bar{\mathbf{a}}_m p)$, the corresponding affine function is 0 and we obtain from eq. (6.99) that

$$J(\square_m, p, \bar{\mathbf{a}}_m p) \leqslant C\left(3^{-\beta m} + \sum_{n=0}^{m} 3^{-\beta(m-n)}\tau_n\right). \tag{6.100}$$

Lemma 6.5.1 ($\mathscr{L}^2$-flatness estimate)**.** *There exist $\beta(d) > 0$ and $C(d, \Lambda, \rho) < \infty$ such that for every $p, q \in B_1$ and $m \in \mathbb{N}$,*

$$\frac{1}{\rho|\square_{m+1}|}\left\|v(\cdot, \square_{m+1}, p, q) - \ell_{\bar{\mathbf{a}}_m^{-1}q-p, \square_{m+1}}\right\|_{\mathscr{L}^2}^2 \leqslant C3^{2m}\left(3^{-\beta m} + \sum_{n=0}^{m} 3^{-\beta(m-n)}\tau_n\right). \tag{6.101}$$

Proof. In the rest of the proof, we write $v(U) := v(\cdot, U, p, q)$ as we will not change p, q in the proof. Since $\mathbb{E}_\rho\left[v(\square_{m+1}) - \ell_{\bar{\mathbf{a}}_m^{-1}q-p, \square_{m+1}}|\mathcal{G}_{m+1}\right] = 0$, we can use the multiscale Poincaré inequality eq. (6.36)

$$\begin{aligned}&\frac{1}{(\rho|\square_{m+1}|)^{\frac{1}{2}}}\left\|v(\square_{m+1}) - \ell_{\bar{\mathbf{a}}_m^{-1}q-p, \square_{m+1}}\right\|_{\mathscr{L}^2} \\ &\leqslant C\left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_{m+1}|}\int_{\square_{m+1}} |\nabla v(\square_{m+1}) - (\bar{\mathbf{a}}_m^{-1}q - p)|^2 \,\mathrm{d}\mu\right]\right)^{\frac{1}{2}} \\ &\quad + C\sum_{n=0}^{m+1} 3^n\left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_{m+1}|}\int_{\square_{m+1}} |\mathsf{S}_{m+1,n}\nabla v(\square_{m+1}) - (\bar{\mathbf{a}}_m^{-1}q - p)|^2 \,\mathrm{d}\mu\right]\right)^{\frac{1}{2}}.\end{aligned} \tag{6.102}$$

The first term on the right side above is of constant order, by eq. (6.87). For the second term, we use a two-scale comparison for every $0 \leqslant n \leqslant m+1$ that

$$\begin{aligned}
&\left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_{m+1}|}\int_{\square_{m+1}} |\mathsf{S}_{m+1,n}\nabla v(\square_{m+1}) - (\bar{\mathbf{a}}_m^{-1}q - p)|^2 \,\mathrm{d}\mu\right]\right)^{\frac{1}{2}} \\
&\leqslant |\bar{\mathbf{a}}_m^{-1} - \bar{\mathbf{a}}_n^{-1}| + \left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_{m+1}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\int_{z+\square_n} |\mathsf{S}_{m+1,n}\nabla v(\square_{m+1}) - \mathsf{S}_{m+1,n}\nabla v(z+\square_n)|^2 \,\mathrm{d}\mu\right]\right)^{\frac{1}{2}} \\
&\qquad + \left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_{m+1}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\int_{z+\square_n} |\mathsf{S}_{m+1,n}\nabla v(z+\square_n) - (\bar{\mathbf{a}}_n^{-1}q - p)|^2 \,\mathrm{d}\mu\right]\right)^{\frac{1}{2}}.
\end{aligned} \tag{6.103}$$

For the first term $|\bar{\mathbf{a}}_m^{-1} - \bar{\mathbf{a}}_n^{-1}|$ we have

$$|\bar{\mathbf{a}}_m^{-1} - \bar{\mathbf{a}}_n^{-1}|^2 \leqslant C(d,\Lambda)|\bar{\mathbf{a}}_m^{-1} - \bar{\mathbf{a}}_n^{-1}| \leqslant \sum_{k=n}^{m-1} \tau_k.$$

For the second term in eq. (6.103), recalling eq. (6.28), we use Jensen's inequality and eq. (6.91) to get

$$\begin{aligned}
&\mathbb{E}_\rho\left[\frac{1}{\rho|\square_{m+1}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\int_{z+\square_n} |\mathsf{S}_{m+1,n}\nabla v(\square_{m+1}) - \mathsf{S}_{m+1,n}\nabla v(z+\square_n)|^2 \,\mathrm{d}\mu\right] \\
&\leqslant \mathbb{E}_\rho\left[\frac{1}{\rho|\square_{m+1}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\int_{z+\square_n} |\nabla v(\square_{m+1}) - \nabla v(z+\square_n)|^2 \,\mathrm{d}\mu\right] \\
&\leqslant \sum_{k=n}^{m} \tau_k.
\end{aligned}$$

For the third term eq. (6.103), we use eq. (6.30), Jensen's inequality, and stationarity. Here we remark that the operator $\mathsf{S}^z_{n,n}$ is a conditional expectation with more information than $\mathsf{S}_{m+1,n}$.

$$\begin{aligned}
&\mathbb{E}_\rho\left[\frac{1}{\rho|\square_{m+1}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\int_{z+\square_n} |\mathsf{S}_{m+1,n}\nabla v(z+\square_n) - (\bar{\mathbf{a}}_n^{-1}q - p)|^2 \,\mathrm{d}\mu\right] \\
&\leqslant \mathbb{E}_\rho\left[\frac{1}{\rho|\square_{m+1}|}\sum_{z\in\mathcal{Z}_{m+1,n}}\int_{z+\square_n} |\mathsf{S}^z_{n,n}\nabla v(z+\square_n) - (\bar{\mathbf{a}}_n^{-1}q - p)|^2 \,\mathrm{d}\mu\right] \\
&= \mathbb{E}_\rho\left[\frac{1}{\rho|\square_n|}\int_{\square_n} |\mathsf{S}_n\nabla v(\square_n) - (\bar{\mathbf{a}}_n^{-1}q - p)|^2 \,\mathrm{d}\mu\right].
\end{aligned}$$

The estimation of this term is postponed to the next step. We will prove in Lemma 6.5.2 below that

$$\mathbb{E}_\rho\left[\frac{1}{\rho|\square_n|}\int_{\square_n} |\mathsf{S}_n\nabla v(\square_n) - (\bar{\mathbf{a}}_n^{-1}q - p)|^2 \,\mathrm{d}\mu\right] \leqslant C3^{-\beta n} + \sum_{k=0}^{n-1} 3^{-\beta(n-k)}\tau_k.$$

We put these estimates back to eq. (6.102) and obtain that

$$\frac{1}{(\rho|\square_{m+1}|)^{\frac{1}{2}}}\left\| v(\square_{m+1}) - \ell_{\bar{\mathbf{a}}_m^{-1}q-p,\square_{m+1}}\right\|_{\mathscr{L}^2} \leqslant C\sum_{n=0}^{m} 3^n\left(3^{-\beta n} + \sum_{k=0}^{n-1} 3^{-\beta(n-k)}\tau_k + \sum_{k=n}^{m}\tau_k\right)^{\frac{1}{2}}.$$

We square the two sides and use the Cauchy-Schwarz inequality to obtain

$$\begin{aligned}
&\frac{1}{\rho|\square_{m+1}|}\left\|v(\square_{m+1}) - \ell_{\bar{\mathbf{a}}_m^{-1}q-p,\square_{m+1}}\right\|^2_{\mathscr{L}^2} \\
&\leqslant C\left(\sum_{n=0}^{m} 3^n\right)\left(\sum_{n=0}^{m} 3^n\left(3^{-\beta n} + \sum_{k=0}^{n-1} 3^{-\beta(n-k)}\tau_k + \sum_{k=n}^{m}\tau_k\right)\right) \\
&\leqslant C 3^{2m}\left(3^{-\beta m} + \sum_{n=0}^{m} 3^{-\beta(m-n)}\tau_n\right),
\end{aligned}$$

as announced. □

6.5.3 Step 3: variance estimate

In this part, we prove the following variance estimate, which was used in Step 2.

Lemma 6.5.2 (Variance estimate)**.** *There exist $\beta(d) > 0$ and $C(d,\Lambda,\rho) < \infty$ such that for every $p, q \in B_1$ and $n \in \mathbb{N}$,*

$$\mathbb{E}_\rho\left[\frac{1}{\rho|\square_n|}\int_{\square_n}|\mathsf{S}_n\nabla v(\mu,\square_n,p,q) - (\bar{\mathbf{a}}_n^{-1}q - p)|^2\,\mathrm{d}\mu\right] \leqslant C3^{-\beta n} + \sum_{k=0}^{n-1} 3^{-\beta(n-k)}\tau_k. \tag{6.104}$$

Proof. In the rest of the proof, we write $v(U) := v(\cdot, U, p, q)$, as we will not change p, q in the proof. From eq. (6.92), we know that the average slope of $v(\square_n)$ is $(\bar{\mathbf{a}}_n^{-1}q - p)$, and notice that $v(\square_n)$ is $\mathcal{F}_{B_1(\square_n)}$-measurable. Thus the idea is to use $\{v(z+\square_k)\}_{z\in\mathcal{Z}_{n,k}}$ to approximate eq. (6.104) in scale 3^k with some error, and then apply the independence for $v(z+\square_k)$ and $v(z'+\square_k)$ for $\mathrm{dist}(z,z')$ large. However, different from the standard elliptic setting, here we will see a renormalization with random weights.

We start by relaxing eq. (6.104) to $\mathcal{G}_{n,n-2}$. We observe that in fact $\mathsf{S}_n\nabla v(\square_n)$ is constant in $\square_n$, so

$$\int_{\square_n}|\mathsf{S}_n\nabla v(\square_n) - (\bar{\mathbf{a}}_n^{-1}q - p)|^2\,\mathrm{d}\mu = \frac{1}{\mu(\square_n)}\left|\int_{\square_n}\left(\mathsf{S}_n\nabla v(\square_n) - (\bar{\mathbf{a}}_n^{-1}q - p)\right)\mathrm{d}\mu\right|^2.$$

We denote by V_n the left hand side of eq. (6.104). By triangle inequality, we have

$$(V_n)^{\frac{1}{2}} \leqslant \text{eq. (6.105)-a} + \text{eq. (6.105)-b} + \text{eq. (6.105)-c}, \tag{6.105}$$

with

$$\text{eq. (6.105)-a} = |\bar{\mathbf{a}}_n^{-1} - \bar{\mathbf{a}}_{n-2}^{-1}|,$$

$$\text{eq. (6.105)-b} = \left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_n|}\frac{1}{\mu(\square_n)}\left|\sum_{z\in\mathcal{Z}_{n,n-2}}\int_{z+\square_{n-2}}\left(\mathsf{S}_{n,n}\nabla v(\square_n) - \mathsf{S}_{n,n-2}\nabla v(z+\square_{n-2})\right)\mathrm{d}\mu\right|^2\right]\right)^{\frac{1}{2}},$$

$$\text{eq. (6.105)-c} = \left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_n|}\frac{1}{\mu(\square_n)}\left|\sum_{z\in\mathcal{Z}_{n,n-2}}\int_{z+\square_{n-2}}\left(\mathsf{S}_{n,n-2}\nabla v(z+\square_{n-2}) - (\bar{\mathbf{a}}_{n-2}^{-1}q - p)\right)\mathrm{d}\mu\right|^2\right]\right)^{\frac{1}{2}}.$$

The term eq. (6.105)-a can be controlled by eq. (6.94):

$$\text{eq. (6.105)-a} \leqslant C(\tau_{n-2} + \tau_{n-1})^{\frac{1}{2}}. \tag{6.106}$$

For the term eq. (6.105)-b, recalling eq. (6.30) and eq. (6.28), we use Jensen's inequality and the two-scale comparison eq. (6.91) to get

$$\begin{aligned}
\text{eq. (6.105)-b} &\leqslant \left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_n|}\sum_{z\in\mathcal{Z}_{n,n-2}}\int_{z+\square_{n-2}}|\mathsf{S}_{n,n}\nabla v(\square_n)-\mathsf{S}_{n,n-2}\nabla v(z+\square_{n-2})|^2\,\mathrm{d}\mu\right]\right)^{\frac{1}{2}}\\
&\leqslant \left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_n|}\sum_{z\in\mathcal{Z}_{n,n-2}}\int_{z+\square_{n-2}}|\nabla v(\square_n)-\nabla v(z+\square_{n-2})|^2\,\mathrm{d}\mu\right]\right)^{\frac{1}{2}}\\
&\leqslant (\tau_{n-2}+\tau_{n-1})^{\frac{1}{2}}.
\end{aligned} \tag{6.107}$$

The term eq. (6.105)-c is the key for our result. To simplify a little more the notation, we write

$$\begin{cases} X_z := \mathsf{S}_{n,n-2}\nabla v(z+\square_{n-2})(\mu,z)-(\bar{\mathbf{a}}_{n-2}^{-1}q-p),\\ m_z := \mu(z+\square_{n-2}).\end{cases} \tag{6.108}$$

Notice that X_z, m_z are $\mathcal{F}_{z+\square_{n-1}}$-measurable. With this notation in place, we have

$$\int_{z+\square_{n-2}}\left(\mathsf{S}_{n,n-2}\nabla v(z+\square_{n-2})-(\bar{\mathbf{a}}_{n-2}^{-1}q-p)\right)\mathrm{d}\mu = m_zX_z,$$

and by eq. (6.92),

$$\mathbb{E}_\rho[m_zX_z]=0.$$

The term eq. (6.105)-c we want to estimate can be rewritten as

$$\text{eq. (6.105)-c} = \left(\mathbb{E}_\rho\left[\frac{1}{\rho|\square_n|}\frac{\left(\sum_{z\in\mathcal{Z}_{n,n-2}}m_zX_z\right)^2}{\sum_{z\in\mathcal{Z}_{n,n-2}}m_z}\right]\right)^{\frac{1}{2}}.$$

If the coefficients m_z were deterministic, then we would be able to leverage on the finite range of dependence of X_z in this variance term. However, since the number of particles m_z is random, we introduce the event

$$\mathcal{C}_{n,\rho,\delta} := \left\{\mu\in\mathcal{M}_\delta(\mathbb{R}^d):\forall z\in\mathcal{Z}_{n,n-2}, \left|\frac{\mu(z+\square_{n-2})}{\rho|\square_{n-2}|}-1\right|\leqslant\delta,\ \text{and}\ \left|\frac{\mu(\square_n)}{\rho|\square_n|}-1\right|\leqslant\delta\right\}, \tag{6.109}$$

thus we can divide eq. (6.105)-c into two terms

$$\begin{aligned}
&\text{eq. (6.105)-c} \leqslant \text{eq. (6.105)-c1} + \text{eq. (6.105)-c2},\\
&\text{eq. (6.105)-c1} = \left(\mathbb{E}_\rho\left[\frac{\mathbf{1}_{\{(\mathcal{C}_{n,\rho,\delta})^c\}}}{\rho|\square_n|}\frac{\left(\sum_{z\in\mathcal{Z}_{n,n-2}}m_zX_z\right)^2}{\sum_{z\in\mathcal{Z}_{n,n-2}}m_z}\right]\right)^{\frac{1}{2}},\\
&\text{eq. (6.105)-c2} = \left(\mathbb{E}_\rho\left[\frac{\mathbf{1}_{\{\mathcal{C}_{n,\rho,\delta}\}}}{\rho|\square_n|}\frac{\left(\sum_{z\in\mathcal{Z}_{n,n-2}}m_zX_z\right)^2}{\sum_{z\in\mathcal{Z}_{n,n-2}}m_z}\right]\right)^{\frac{1}{2}}.
\end{aligned}$$

For the term eq. (6.105)-c1, we know that $(\mathcal{C}_{n,\rho,\delta})^c$ is not typical in large scales, and we have the Chernoff bound

$$\mathbb{P}_\rho[\mu\notin\mathcal{C}_{n,\rho,\delta}]\leqslant 3^{2d+1}\exp\left(-\frac{\rho|\square_{n-2}|\delta^2}{4}\right).$$

Moreover, by the Cauchy-Schwarz inequality,

$$\frac{\left(\sum_{z\in\mathcal{Z}_{n,n-2}} m_z X_z\right)^2}{\sum_{z\in\mathcal{Z}_{n,n-2}} m_z} \leqslant \sum_{z\in\mathcal{Z}_{n,n-2}} m_z |X_z|^2.$$

We need a bound for the term $|X_z|^2$: recalling the definition in eq. (6.28) and eq. (6.88),

$$\begin{aligned}\mathsf{S}^z_{n-2,n-2}\nabla v(z+\square_{n-2})(\mu,z) &= \mathbb{E}_\rho\left[\fint_{z+\square_{n-2}} \nabla v(z+\square_{n-2})\,\mathrm{d}\mu \;\middle|\; \mathcal{G}^z_{n-2,n-2}\right]\\ &= \mathbf{a}(z+\square_{n-2};\mathcal{G}_{z+\square_{n-2}})^{-1}q - p.\end{aligned}$$

Using the martingale structure of eq. (6.30), we have

$$\begin{aligned}X_z &= \mathsf{S}_{n,n-2}\nabla v(z+\square_{n-2})(\mu,z) - (\bar{\mathbf{a}}^{-1}_{n-2}q - p)\\ &= \mathbb{E}_\rho\left[\fint_{z+\square_{n-2}} \mathsf{S}^z_{n-2,n-2}\nabla v(z+\square_{n-2})\,\mathrm{d}\mu \;\middle|\; \mathcal{G}_{n,n-2}\right] - (\bar{\mathbf{a}}^{-1}_{n-2}q - p)\\ &= \mathbb{E}_\rho\left[\mathbf{a}(z+\square_{n-2};\mathcal{G}_{z+\square_{n-2}})^{-1} - \bar{\mathbf{a}}^{-1}_{n-2} \;\middle|\; \mathcal{G}_{n,n-2}\right]q.\end{aligned}$$

Then we use Jensen's inequality and the bound of $\mathsf{Id} \leqslant \mathbf{a}(z+\square_{n-2};\mathcal{G}_{z+\square_{n-2}}) \leqslant \Lambda\mathsf{Id}$

$$\begin{aligned}|X_z|^2 &= |\mathsf{S}_{n,n-2}\nabla v(z+\square_{n-2}) - (\bar{\mathbf{a}}^{-1}_{n-2}q - p)|^2\\ &= \mathbb{E}_\rho\left[|\mathbf{a}(z+\square_{n-2};\mathcal{G}_{z+\square_{n-2}}) - \bar{\mathbf{a}}^{-1}_{n-2}|^2 \,|\, \mathcal{G}_{n,n-2}\right] \leqslant \Lambda^2.\end{aligned} \tag{6.110}$$

This concludes that

$$\begin{aligned}\text{eq. (6.105)-c1} \leqslant \Lambda^2\mathbb{E}_\rho\left[\frac{\mathbf{1}_{\{(\mathcal{C}_{n,\rho,\delta})^c\}}}{\rho|\square_n|}\mu(\square_n)\right] &\leqslant C(d,\Lambda)\frac{1}{\rho|\square_{n-2}|}\exp\left(-\frac{\rho|\square_{n-2}|\delta^2}{4}\right)\\ &\leqslant C(d,\Lambda,\rho)3^{-dn}.\end{aligned} \tag{6.111}$$

Finally, we treat eq. (6.105)-c2. We calculate eq. (6.105)-c2 at first with the conditional expectation with respect to $\mathcal{G}_{n,n-2}$. Clearly, $\mathcal{C}_{n,\rho,\delta}$ is $\mathcal{G}_{n,n-2}$-measurable, and under this condition $\mu(\square_n) \geqslant (1-\delta)\rho|\square_n|$, so we have

$$\begin{aligned}|\text{eq. (6.105)-c2}|^2 &= \frac{1}{\rho|\square_n|}\mathbb{E}_\rho\left[\frac{\mathbf{1}_{\{\mathcal{C}_{n,\rho,\delta}\}}}{\mu(\square_n)}\mathbb{E}_\rho\left[\left(\sum_{z\in\mathcal{Z}_{n,n-2}} m_z X_z\right)^2 \;\middle|\; \mathcal{G}_{n,n-2}\right]\right]\\ &\leqslant \frac{1}{\rho|\square_n|}\mathbb{E}_\rho\left[\frac{1}{(1-\delta)\rho|\square_n|}\mathbb{E}_\rho\left[\mathbf{1}_{\{\mathcal{C}_{n,\rho,\delta}\}}\left(\sum_{z\in\mathcal{Z}_{n,n-2}} m_z X_z\right)^2 \;\middle|\; \mathcal{G}_{n,n-2}\right]\right].\end{aligned} \tag{6.112}$$

We would like to develop the term $|\sum_{z\in\mathcal{Z}_{n,n-2}} m_z X_z|^2$ and also drop out the indicator term. The argument here is deterministic

$$\begin{aligned}\left|\sum_{z\in\mathcal{Z}_{n,n-2}} m_z X_z\right|^2 &= \sum_{\substack{z,z'\in\mathcal{Z}_{n,n-2}\\ |z-z'|_\infty<3^{n-1}}} m_z m_{z'} X_z\cdot X_{z'} + \sum_{\substack{z,z'\in\mathcal{Z}_{n,n-2}\\ |z-z'|_\infty\geqslant 3^{n-1}}} m_z m_{z'} X_z\cdot X_{z'}\\ &\leqslant \frac{1}{2}\sum_{\substack{z,z'\in\mathcal{Z}_{n,n-2}\\ |z-z'|_\infty<3^{n-1}}} \left((m_z)^2|X_z|^2 + (m_{z'})^2|X_{z'}|^2\right) + \sum_{\substack{z,z'\in\mathcal{Z}_{n,n-2}\\ |z-z'|_\infty\geqslant 3^{n-1}}} m_z m_{z'} X_z\cdot X_{z'},\end{aligned}$$

where $|z-z'|_\infty := \max_{1\leqslant i\leqslant d}|z_i - z'_i|$. We now add back the indicator $\mathbf{1}_{\{\mathcal{C}_{n,\rho,\delta}\}}$ and develop it

$$\begin{aligned}
&\mathbf{1}_{\{\mathcal{C}_{n,\rho,\delta}\}}\left|\sum_{z\in\mathcal{Z}_{n,n-2}} m_z X_z\right|^2 \\
&\leqslant \mathbf{1}_{\{\mathcal{C}_{n,\rho,\delta}\}}\left(\frac{(1+\delta)\rho|\square_{n-2}|}{2}\sum_{\substack{z,z'\in\mathcal{Z}_{n,n-2}\\ |z-z'|<3^{n-1}}}\left(m_z|X_z|^2 + m_{z'}|X_{z'}|^2\right) + \sum_{\substack{z,z'\in\mathcal{Z}_{n,n-2}\\ |z-z'|\geqslant 3^{n-1}}} m_z m_{z'} X_z\cdot X_{z'}\right) \\
&\leqslant \frac{(1+\delta)\rho|\square_{n-2}|}{2}\sum_{\substack{z,z'\in\mathcal{Z}_{n,n-2}\\ |z-z'|<3^{n-1}}}\left(m_z|X_z|^2 + m_{z'}|X_{z'}|^2\right) + \sum_{\substack{z,z'\in\mathcal{Z}_{n,n-2}\\ |z-z'|\geqslant 3^{n-1}}} m_z m_{z'} X_z\cdot X_{z'}.
\end{aligned} \tag{6.113}$$

From the first line to the second line above, we use that $m_z \leqslant (1+\delta)\rho|\square_{n-2}|$ under the event $\mathcal{C}_{n,\rho,\delta}$. We then keep in mind that the quantity in $\left(\cdots\right)$ on the second line of eq. (6.113) is always larger than $\left|\sum_{z\in\mathcal{Z}_{n,n-2}} m_z X_z\right|^2$, so it is nonnegative. Therefore, from the second line to the third line, we can drop the indicator function in front. Inserting this estimate into eq. (6.112), we obtain that

$$\begin{aligned}
|\text{eq. (6.105)-c2}|^2 \leqslant{}& \frac{1}{\rho|\square_n|}\frac{(1+\delta)|\square_{n-2}|}{(1-\delta)|\square_n|}\sum_{\substack{z,z'\in\mathcal{Z}_{n,n-2}\\ |z-z'|_\infty<3^{n-1}}}\mathbb{E}_\rho\left[\frac{1}{2}\left(m_z|X_z|^2 + m_{z'}|X_{z'}|^2\right)\right] \\
&+\frac{1}{\rho|\square_n|}\frac{1}{(1-\delta)|\square_n|}\sum_{\substack{z,z'\in\mathcal{Z}_{n,n-2}\\ |z-z'|_\infty\geqslant 3^{n-1}}}\mathbb{E}_\rho\left[m_z m_{z'} X_z\cdot X_{z'}\right].
\end{aligned}$$

The sum in the second line is 0, because for $|z-z'|_\infty \geqslant 3^{n-1}$, $m_z X_z$ and $m_{z'}X_{z'}$ are independent,

$$\mathbb{E}_\rho\left[m_z m_{z'} X_z\cdot X_{z'}\right] = \mathbb{E}_\rho\left[m_z X_z\right]\cdot\mathbb{E}_\rho\left[m_{z'}X_{z'}\right] = 0.$$

For the sum in the first line, $\mathbb{E}_\rho[m_z|X_z|^2]$ is nothing but

$$\mathbb{E}_\rho\left[\int_{z+\square_{n-2}}|\mathsf{S}_{n,n-2}\nabla v(z+\square_{n-2}) - (\bar{\mathbf{a}}_{n-2}^{-1}q - p)|^2\,d\mu\right].$$

We use Jensen's inequality to shrink the operator to $\mathsf{S}^z_{n-2,n-2}$ that

$$\begin{aligned}
&\mathbb{E}_\rho\left[\int_{z+\square_{n-2}}|\mathsf{S}_{n,n-2}\nabla v(z+\square_{n-2}) - (\bar{\mathbf{a}}_{n-2}^{-1}q - p)|^2\,d\mu\right] \\
&\leqslant \mathbb{E}_\rho\left[\int_{z+\square_{n-2}}|\mathsf{S}^z_{n-2,n-2}\nabla v(z+\square_{n-2}) - (\bar{\mathbf{a}}_{n-2}^{-1}q - p)|^2\,d\mu\right] \\
&= \mathbb{E}_\rho\left[\int_{\square_{n-2}}|\mathsf{S}_{n-2}\nabla v(\square_{n-2}) - (\bar{\mathbf{a}}_{n-2}^{-1}q - p)|^2\,d\mu\right].
\end{aligned}$$

There are at most $9^d\times 5^d$ pairs $z,z'\in\mathcal{Z}_{n,n-2}$ such that $|z-z'|_\infty < 3^{n-1}$; see Figure 6.3 for an illustration. Therefore, we obtain

$$\begin{aligned}
|\text{eq. (6.105)-c2}|^2 &\leqslant \left(\frac{5}{9}\right)^d\left(\frac{1+\delta}{1-\delta}\right)\mathbb{E}_\rho\left[\frac{1}{\rho|\square_{n-2}|}\int_{\square_{n-2}}|\mathsf{S}_{n-2}\nabla v(\square_{n-2}) - (\bar{\mathbf{a}}_{n-2}^{-1}q - p)|^2\,d\mu\right] \\
&= \left(\frac{5}{9}\right)^d\left(\frac{1+\delta}{1-\delta}\right)V_{n-2},
\end{aligned}$$

where we recall that V_n is the left hand side of eq. (6.104). We put this estimate together with eq. (6.106), (6.107), (6.111) back to eq. (6.105) to obtain the recurrence relation

$$(V_n)^{\frac{1}{2}} \leqslant \left(\frac{5}{9}\right)^{\frac{d}{2}} \left(\frac{1+\delta}{1-\delta}\right)^{\frac{1}{2}} (V_{n-2})^{\frac{1}{2}} + C(\tau_{n-2} + \tau_{n-1})^{\frac{1}{2}} + C3^{-dn}.$$

By choosing $\delta(d) > 0$ sufficiently small, we obtain the desired result eq. (6.104). □

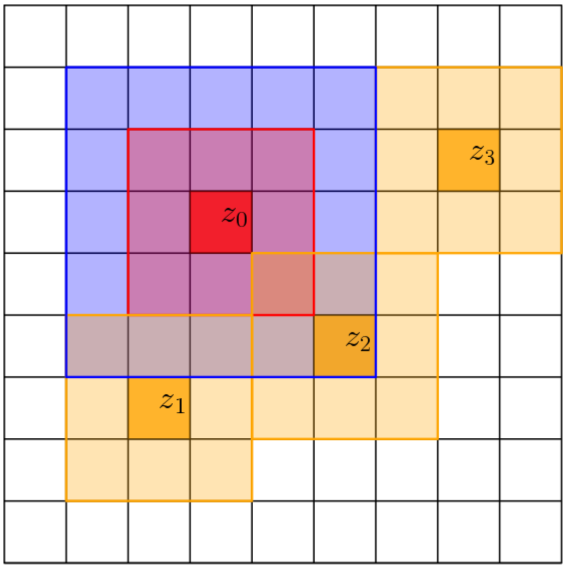

Figure 6.3: In the cube $\square_n$ and all its sub-cubes $\{z + \square_{n-2}\}_{z\in\mathcal{Z}_{n,n-2}}$, for a chosen sub-cube $z_0 + \square_{n-2}$ (the cube in dark red), the support of $v(z_0 + \square_{n-2})$ is in $z_0 + \square_{n-1}$ (the cube in light red), so it has at most 5^d cubes of scale 3^{n-2} whose associated function has a support intersecting with $z_1 + \square_{n-1}$ (the cube in blue). For example, $v(z_2 + \square_{n-2})$ has correlation with $v(z_0 + \square_{n-2})$, while $v(z_1 + \square_{n-2}), v(z_3 + \square_{n-2})$ do not. This gives us the contraction factor $\left(\frac{5}{9}\right)^d$.

6.5.4 Step 4: iterations

Once we obtain the estimate eq. (6.100), it remains to do some numerical iterations, similarly to [25, Page 59-60]. For the reader's convenience, we recall the main steps here. Let $\{e_i\}_{1\leqslant i\leqslant d}$ denote the canonical basis in $\mathbb{R}^d$, and define

$$F_m := \sum_{i=1}^{d} J(\square_m, e_i, \bar{\mathbf{a}}_m e_i).$$

In order to obtain an exponential decay for $(F_m)_{m\geqslant 0}$, we first introduce a weighted version of this quantity:

$$\widetilde{F}_m := \sum_{n=0}^{m} 3^{-\frac{\beta}{2}(m-n)} F_n.$$

Here the exponent β is the same as in eq. (6.100). It is clear that $F_m \leqslant \widetilde{F}_m$, so it suffices to prove an exponential decay for $(\widetilde{F}_m)_{m\geqslant 0}$. We will do so by proving a recurrence equation of type $\widetilde{F}_{m+1} \leqslant C(\widetilde{F}_m - \widetilde{F}_{m+1})$ for some constant $C(d,\Lambda) < \infty$. Thus in the following, we calculate some bounds for $(\widetilde{F}_m - \widetilde{F}_{m+1})$ and $\widetilde{F}_{m+1}$.

Starting with $(\widetilde{F}_m - \widetilde{F}_{m+1})$, we write

$$\widetilde{F}_m - \widetilde{F}_{m+1} \geqslant \sum_{n=0}^{n} 3^{-\frac{\beta}{2}(m-n)}(F_n - F_{n+1}) - C3^{-\frac{\beta m}{2}}.$$

Noticing that $\bar{\mathbf{a}}_{n+1}p$ is the minimizer for the mapping $q \mapsto J(\square_{n+1}, p, q)$ in eq. (6.79), we have

$$F_{n+1} = \sum_{i=1}^{d} J(\square_{n+1}, e_i, \bar{\mathbf{a}}_{n+1}e_i) \leqslant \sum_{i=1}^{d} J(\square_{n+1}, e_i, \bar{\mathbf{a}}_n e_i). \tag{6.114}$$

Using also eq. (6.79), that $\mathsf{Id} \leqslant \bar{\mathbf{a}}_n \leqslant \Lambda \mathsf{Id}$, and that the map $p \mapsto \nu(\square_n, p) - \nu(\square_{n+1}, p)$ and $q \mapsto \nu^*(\square_n, q) - \nu^*(\square_{n+1}, q)$ are positive semidefinite quadratic forms, we get

$$\begin{aligned} F_n - F_{n+1} &\geqslant \sum_{i=1}^{d} \left(J(\square_n, e_i, \bar{\mathbf{a}}_n e_i) - J(\square_{n+1}, e_i, \bar{\mathbf{a}}_n e_i)\right) \\ &= \sum_{i=1}^{d} \left(\nu(\square_n, e_i) - \nu(\square_{n+1}, e_i)\right) + \sum_{i=1}^{d} \left(\nu^*(\square_n, \bar{\mathbf{a}}_n e_i) - \nu^*(\square_{n+1}, \bar{\mathbf{a}}_n e_i)\right) \\ &\geqslant C^{-1} \left(\sup_{p \in B_1} \left(\nu(\square_n, p) - \nu(\square_{n+1}, p)\right) + \sup_{q \in \mathsf{B}_1} \left(\nu^*(\square_n, q) - \nu^*(\square_{n+1}, q)\right) \right) \\ &\geqslant C^{-1} \tau_n, \end{aligned}$$

and thus

$$\widetilde{F}_m - \widetilde{F}_{m+1} \geqslant C^{-1} \sum_{n=0}^{m} 3^{-\frac{\beta}{2}(m-n)} \tau_n - C3^{-\frac{\beta m}{2}}. \tag{6.115}$$

For the upper bound of $\widetilde{F}_{m+1}$, we use eq. (6.114) to see that $F_n \leqslant F_{n+1}$, so

$$\begin{aligned} \widetilde{F}_{m+1} &= 3^{-\frac{\beta}{2}(m+1)} F_0 + \sum_{n=0}^{m} 3^{-\frac{\beta}{2}(m-n)} F_{n+1} \\ &\leqslant C3^{-\frac{\beta m}{2}} + \sum_{n=0}^{m} 3^{-\frac{\beta}{2}(m-n)} F_n. \end{aligned}$$

Then we apply eq. (6.100) into the result above to get

$$\begin{aligned} \widetilde{F}_{m+1} &\leqslant C3^{-\frac{\beta m}{2}} + \sum_{n=0}^{m} 3^{-\frac{\beta}{2}(m-n)} \left(3^{-\beta n} + \sum_{k=0}^{n} 3^{-\beta(n-k)} \tau_k \right) \\ &\leqslant C3^{-\frac{\beta m}{2}} + 3^{-\frac{\beta}{2}m} \sum_{k=0}^{m} \tau_k \sum_{n=k}^{m} 3^{\frac{\beta}{2}(2k-n)} \\ &\leqslant C3^{-\frac{\beta m}{2}} + C \sum_{k=0}^{m} 3^{-\frac{\beta}{2}(m-k)} \tau_k. \end{aligned} \tag{6.116}$$

We combine eq. (6.115) and eq. (6.116), to obtain $C(\widetilde{F}_m - \widetilde{F}_{m+1} + \widetilde{C}3^{-\frac{\beta m}{2}}) \geqslant \widetilde{F}_{m+1}$, which implies

$$\widetilde{F}_{m+1} \leqslant \theta \widetilde{F}_m + C3^{-\frac{\beta m}{2}},$$

for some $\theta(d,\Lambda) \in (0,1)$. We thus conclude for the exponential decay of $(\widetilde{F}_m)_{m \geqslant 0}$, and thus also of F_m, since $F_m \leqslant \widetilde{F}_m$. By eq. (6.81), this completes the proof of Theorem 6.2.1.

6.A Some elementary properties of the function spaces

Lemma 6.A.1 (Canonical projection)**.** *Let $f:\mathcal{M}_\delta(\mathbb{R}^d)\to\mathbb{R}$ be a function, and for every Borel set U, measure $\mu\in\mathcal{M}_\delta(\mathbb{R}^d)$, and $n\in\mathbb{N}$, let $f_n(\cdot,\mu\llcorner U^c)$ denote the (permutation-invariant) function*

$$f_n(\cdot,\mu\llcorner U^c):\begin{cases} U^n & \to \mathbb{R} \\ (x_1,\dots,x_n) & \mapsto f\left(\sum_{i=1}^n \delta_{x_i}+\mu\llcorner U^c\right). \end{cases}$$

The following statements are equivalent.
(1) The function f is $\mathcal{F}$-measurable.
(2) For every $n\in\mathbb{N}$, the function f_n is $\mathcal{B}_U^{\otimes n}\otimes\mathcal{F}_{U^c}$-measurable.

Proof. We start from (1) $\Rightarrow$ (2). Because $\mathcal{F}=\mathcal{F}_U\otimes\mathcal{F}_{U^c}$, it suffices to study the product function

$$f=\mathbf{1}_{\{\mu(V_1)=n_1\}}\mathbf{1}_{\{\mu(V_2)=n_2\}}\mathbf{1}_{\{\mu(U)=n\}},$$

for some Borel sets $V_1\subseteq U, V_2\subseteq U^c$. In this case, we have

$$\begin{aligned}\{f_n=1\}&=\{\mu(V_1)=n_1\}\cap\{\mu(V_2)=n_2\}\cap\{\mu(U)=n\}\\ &=\bigcup_{\sigma\in S_n}\left(\bigcap_{i=1}^{n_1}\{x_{\sigma(i)}\in V_1\}\bigcap_{j=n_1+1}^{n}\{x_{\sigma(j)}\in(U\backslash V_1)\}\bigcap\{\mu(V_2)=n_2\}\right),\end{aligned}$$

where S_n is the symmetric group. This proves that f_n is $\mathcal{B}_U^{\otimes n}\otimes\mathcal{F}_{U^c}$-measurable.

We turn to (2) $\Rightarrow$ (1). Let us pick a suitable f_n and $\mu\llcorner U=\sum_{i=1}^n\delta_{x_i}$, then the main point is to establish the $\mathcal{F}$-measurable property. Since f_n is $\mathcal{B}_U^{\otimes n}\otimes\mathcal{F}_{U^c}$-measurable and permutation-invariant, it suffices to study the function of type

$$f_n=\sum_{\sigma\in S_n}\left(\prod_{i=1}^n\mathbf{1}_{\{x_{\sigma(i)}\in V_i\}}\right)\mathbf{1}_{\{\mu\llcorner U^c(V_0)=n_0\}}\mathbf{1}_{\{\mu(U)=n\}},\tag{6.117}$$

for $\{V_i\}_{0\leqslant i\leqslant n}$ Borel sets. This is still a complicated function, but we can add one more condition

$$\forall 1\leqslant i,j\leqslant n,\quad V_i=V_j \text{ or } V_i\cap V_j=\varnothing.\tag{6.118}$$

For example, let $\{\widetilde{V}_j\}_{0\leqslant j\leqslant m}$ be all the different elements in $\{V_i\}_{0\leqslant i\leqslant n}$, and $\widetilde{V}_j$ appears n_j times. For the functions of type eq. (6.117) satisfying the condition eq. (6.118), the $\mathcal{F}$-measurable property is easy to treat since we have

$$\sum_{\sigma\in S_n}\left(\prod_{i=1}^n\mathbf{1}_{\{x_{\sigma(i)}\in V_i\}}\right)\mathbf{1}_{\{\mu\llcorner U^c(V_0)=n_0\}}\mathbf{1}_{\{\mu(U)=n\}}=\left(\prod_{j=1}^m\mathbf{1}_{\{\mu(\widetilde{V}_j)=n_j\}}\right)\mathbf{1}_{\{\mu\llcorner U^c(V_0)=n_0\}}\mathbf{1}_{\{\mu(U)=n\}},$$

which is an $\mathcal{F}$-measurable function.

Finally, let us conclude that for a general f_n in eq. (6.117), they can be decomposed into the sum of the functions with the propriety eq. (6.118). Let us see the case $n=2$, where we have the following decomposition

$$\begin{aligned}\mathbf{1}_{\{x_1\in V_1\}}\mathbf{1}_{\{x_2\in V_2\}}&=\left(\mathbf{1}_{\{x_1\in(V_1\backslash V_2)\}}+\mathbf{1}_{\{x_1\in(V_1\cap V_2)\}}\right)\left(\mathbf{1}_{\{x_2\in(V_2\backslash V_1)\}}+\mathbf{1}_{\{x_2\in(V_1\cap V_2)\}}\right)\\ &=\mathbf{1}_{\{x_1\in(V_1\backslash V_2)\}}\mathbf{1}_{\{x_2\in(V_2\backslash V_1)\}}+\mathbf{1}_{\{x_1\in(V_1\backslash V_2)\}}\mathbf{1}_{\{x_2\in(V_1\cap V_2)\}}\\ &\quad+\mathbf{1}_{\{x_1\in(V_1\cap V_2)\}}\mathbf{1}_{\{x_2\in(V_2\backslash V_1)\}}+\mathbf{1}_{\{x_1\in(V_1\cap V_2)\}}\mathbf{1}_{\{x_2\in(V_1\cap V_2)\}}.\end{aligned}$$

For a general n, one can use induction and this concludes the proof. ◻

Proposition 6.A.1. *For every $s > 0$ and $f \in \mathscr{H}^1(Q_s)$, we have $\mathsf{A}_s f \in \mathscr{H}^1(Q_s)$, and for every $x \in \operatorname{supp}(\mu) \cap Q_s$*

$$\nabla(\mathsf{A}_s f)(\mu, x) = \mathsf{A}_s(\nabla f)(\mu, x). \tag{6.119}$$

Moreover, if $s > 2$ and $f \in \mathcal{A}(Q_s)$, then $\mathsf{A}_s f \in \mathcal{A}(Q_{s-2})$.

Proof. At first, we should remark the well-definedness of the right side of eq. (6.119). Notice that the Poisson measure can be decomposed as a sum of the independent parts $\mu = \mu \llcorner \overline{Q}_s + \mu \llcorner \overline{Q}_s^c$, we have

$$\mathsf{A}_s f = \int_{\mathcal{M}_\delta(\mathbb{R}^d)} f(\mu \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c) \, \mathrm{d}\mathbb{P}_\rho(\mu').$$

Thus the right-hand side of eq. (6.119) is defined as

$$\mathsf{A}_s(\nabla f)(\mu, x) = \int_{\mathcal{M}_\delta(\mathbb{R}^d)} \nabla f(\mu \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c, x) \, \mathrm{d}\mathbb{P}_\rho(\mu'). \tag{6.120}$$

We prove eq. (6.119) and $\mathsf{A}_s f \in \mathscr{H}^1(Q_s)$ for the functions in $\mathscr{C}^\infty(Q_s) \cap \mathscr{H}^1(Q_s)$ as they are dense, and we can focus on the case $\mu(Q_s) = n$ fixed. We use Lemma 6.A.1 to write $f(\mu)\mathbf{1}_{\{\mu(Q_s)=n\}} = f_n(x_1, \cdots, x_n, \mu \llcorner \overline{Q}_s^c)\mathbf{1}_{\{\mu(Q_s)=n\}}$, and observe that

$$\begin{aligned}
\|\nabla f_n\|_{L^\infty((Q_s)^n)} &= \sup_{(Q_s)^n} \left(\sum_{k=1}^n |\nabla_{x_k} f_n(x_1, \cdots, x_n, \mu \llcorner \overline{Q}_s^c)|^2 \right)^{\frac{1}{2}} \\
&= \sup_{(\mathbb{Q} \cap Q_s)^n} \left(\sum_{k=1}^n |\nabla_{x_k} f_n(x_1, \cdots, x_n, \mu \llcorner \overline{Q}_s^c)|^2 \right)^{\frac{1}{2}},
\end{aligned}$$

is finite and $\mathcal{F}_{\overline{Q}_s^c}$-measurable. Thus we can define a cut-off version of f that

$$f^M = f_n \mathbf{1}_{\{\mu(Q_s)=n\}} \mathbf{1}_{\left\{\|\nabla f_n\|_{L^\infty((Q_s)^n)} \leqslant M\right\}},$$

and then we can establish eq. (6.119) for f^M

$$\begin{aligned}
&\partial_k(\mathsf{A}_s f^M)(\mu, x) \\
&= \lim_{h \to 0} \int_{\mathcal{M}_\delta(\mathbb{R}^d)} \frac{f_n((\mu - \delta_x + \delta_{x+he_k}) \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c) - f_n(\mu \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c)}{h} \\
&\qquad \times \mathbf{1}_{\left\{\|\nabla f_n\|_{L^\infty((Q_s)^n)} \leqslant M\right\}} \mathrm{d}\mathbb{P}_\rho(\mu') \mathbf{1}_{\{\mu(Q_s)=n\}},
\end{aligned}$$

for h small enough. Since $f_n \in C^\infty((Q_s)^n)$, we use the mean value principle

$$\frac{f_n(\mu - \delta_x + \delta_{x+he_k}) - f_n(\mu)}{h} = \partial_k f_n(\mu - \delta_x + \delta_{x+\theta e_k}, x + \theta e_k),$$

for some $\theta \in (0,1)$. With the indicator $\mathbf{1}_{\left\{\|\nabla f_n\|_{L^\infty((Q_s)^n)} \leqslant M\right\}}$, this term is bounded by M, so we can use the dominated convergence theorem that

$$\begin{aligned}
&\partial_k(\mathsf{A}_s f^M)(\mu, x) \\
&= \int_{\mathcal{M}_\delta(\mathbb{R}^d)} \lim_{h \to 0} \frac{f_n((\mu - \delta_x + \delta_{x+he_k}) \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c) - f_n(\mu \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c)}{h} \\
&\qquad \times \mathbf{1}_{\left\{\|\nabla f_n\|_{L^\infty((Q_s)^n)} \leqslant M\right\}} \mathrm{d}\mathbb{P}_\rho(\mu') \mathbf{1}_{\{\mu(Q_s)=n\}} \\
&= \int_{\mathcal{M}_\delta(\mathbb{R}^d)} \partial_k f_n(\mu \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c, x) \mathbf{1}_{\left\{\|\nabla f_n\|_{L^\infty((Q_s)^n)} \leqslant M\right\}} \mathrm{d}\mathbb{P}_\rho(\mu') \mathbf{1}_{\{\mu(Q_s)=n\}},
\end{aligned}$$

which establishes the eq. (6.119) in the sense eq. (6.120). Then by Jensen's inequality and Fubini's lemma we observe that

$$\begin{aligned}
&\mathbb{E}_\rho\left[\int_{Q_s} |\nabla(\mathsf{A}_s f^M)|^2(\mu,x)\,\mathrm{d}\mu(x)\right] \\
&= \mathbb{E}_\rho\left[\int_{Q_s} \left|\int_{\mathcal{M}_\delta(\mathbb{R}^d)} \nabla f^M(\mu \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c, x)\,\mathrm{d}\mathbb{P}_\rho(\mu')\right|^2 \mathrm{d}\mu(x)\right] \\
&\leqslant \mathbb{E}_\rho\left[\int_{Q_s} \int_{\mathcal{M}_\delta(\mathbb{R}^d)} \left|\nabla f^M(\mu \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c, x)\right|^2 \mathrm{d}\mathbb{P}_\rho(\mu')\,\mathrm{d}\mu(x)\right] \\
&= \mathbb{E}_\rho\left[\int_{Q_s} \left|\nabla f^M\right|^2(\mu,x)\,\mathrm{d}\mu(x)\right].
\end{aligned}$$

This implies that $\mathsf{A}_s f^M \in \mathscr{H}^1(Q_s)$. Then we use once again Jensen's inequality for f^M and $f^{M'}$ with $M < M'$

$$\begin{aligned}
&\mathbb{E}_\rho\left[\int_{Q_s} |\nabla(\mathsf{A}_s f^M) - \nabla(\mathsf{A}_s f^{M'})|^2(\mu,x)\,\mathrm{d}\mu(x)\right] \\
&\leqslant \mathbb{E}_\rho\left[\int_{Q_s} \left|\nabla f^M - \nabla f^{M'}\right|^2(\mu,x)\,\mathrm{d}\mu(x)\right] \\
&= \mathbb{E}_\rho\left[\int_{Q_s} |\nabla f_n|^2(\mu,x)\,\mathrm{d}\mu(x)\mathbf{1}_{\{\mu(Q_s)=n\}}\mathbf{1}_{\left\{M<\|\nabla f_n\|_{L^\infty((Q_s)^n)}\leqslant M'\right\}}\right].
\end{aligned}$$

So $\{f^M\}_{M\geqslant 0}$ gives a Cauchy sequence in $\mathscr{H}^1(Q_s)$, and the only candidate is f because it is the limit in $\mathscr{L}^2$. By this, we also establish eq. (6.119) for f in $\mathscr{C}^\infty(Q_s)\cap\mathscr{H}^1(Q_s)$, and we can then extend to a general function in $\mathscr{H}^1(Q_s)$ by the density argument.

For the part of $\mathbf{a}$-harmonic function, we suppose $f \in \mathcal{A}(Q_s)$ and test $\phi \in \mathscr{H}_0^1(Q_{s-2})$ with eq. (6.119),

$$\begin{aligned}
&\mathbb{E}_\rho\left[\int_{Q_{s-2}} (\nabla \mathsf{A}_s f)(\mu,x)\cdot\mathbf{a}(\mu,x)\nabla\phi(\mu,x)\,d\mu(x)\right] \\
&= \mathbb{E}_\rho\left[\int_{Q_{s-2}} \mathsf{A}_s(\nabla f)(\mu,x)\cdot\mathbf{a}(\mu,x)\nabla\phi(\mu,x)\,d\mu(x)\right] \\
&= \mathbb{E}_\rho\left[\int_{Q_{s-2}} \left(\int_{\mathcal{M}_\delta(\mathbb{R}^d)} \nabla f(\mu \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c, x)\,\mathrm{d}\mathbb{P}_\rho(\mu')\right)\cdot\mathbf{a}(\mu,x)\nabla\phi(\mu,x)\,d\mu(x)\right].
\end{aligned}$$

Restricted on $x \in Q_{s-2}$, we have $\mathbf{a}(\mu,x), \nabla\phi(\mu,x)$ are $\mathcal{F}_{Q_s}\otimes\mathcal{B}_{Q_s}$-measurable, so we have

$$\forall x \in \operatorname{supp}(\mu)\cap Q_{s-2}, \qquad \mathbf{a}(\mu,x)\nabla\phi(\mu,x) = \mathbf{a}(\mu \llcorner \overline{Q}_s, x)\nabla\phi(\mu \llcorner \overline{Q}_s, x).$$

We can enter the part in the integration, and then use Fubini's lemma

$$\begin{aligned}
&\mathbb{E}_\rho\left[\int_{Q_{s-2}} (\nabla \mathsf{A}_s f)(\mu,x)\cdot\mathbf{a}(\mu,x)\nabla\phi(\mu,x)\,d\mu(x)\right] \\
&= \mathbb{E}_\rho\left[\int_{Q_{s-2}} \left(\int_{\mathcal{M}_\delta(\mathbb{R}^d)} \nabla f(\mu \llcorner \overline{Q}_s + \mu' \llcorner \overline{Q}_s^c, x)\cdot\mathbf{a}(\mu \llcorner \overline{Q}_s, x)\nabla\phi(\mu \llcorner \overline{Q}_s, x)\,\mathrm{d}\mathbb{P}_\rho(\mu')\right) d\mu(x)\right] \\
&= \mathbb{E}_\rho\left[\int_{Q_{s-2}} \nabla f(\mu,x)\cdot\mathbf{a}(\mu,x)\nabla\phi(\mu,x)\,d\mu(x)\right] \\
&= 0.
\end{aligned}$$

In the last step, we use $f \in \mathcal{A}(Q_s)$ and this finishes the proof. □

Bibliography

[1] M. Aizenman, F. Delyon, and B. Souillard. Lower bounds on the cluster size distribution. *J. Statist. Phys.*, 23(3):267–280, 1980.

[2] S. Albeverio, Y. G. Kondratiev, and M. Röckner. Canonical Dirichlet operator and distorted Brownian motion on Poisson spaces. *C. R. Acad. Sci. Paris Sér. I Math.*, 323(11):1179–1184, 1996.

[3] S. Albeverio, Y. G. Kondratiev, and M. Röckner. Differential geometry of Poisson spaces. *C. R. Acad. Sci. Paris Sér. I Math.*, 323(10):1129–1134, 1996.

[4] S. Albeverio, Y. G. Kondratiev, and M. Röckner. Analysis and geometry on configuration spaces. *J. Funct. Anal.*, 154(2):444–500, 1998.

[5] S. Albeverio, Y. G. Kondratiev, and M. Röckner. Analysis and geometry on configuration spaces: the Gibbsian case. *J. Funct. Anal.*, 157(1):242–291, 1998.

[6] G. Allaire. Homogenization and two-scale convergence. *SIAM J. Math. Anal.*, 23(6):1482–1518, 1992.

[7] G. Allaire. A brief introduction to homogenization and miscellaneous applications. In *Mathematical and numerical approaches for multiscale problem*, volume 37 of *ESAIM Proc.*, pages 1–49. EDP Sci., Les Ulis, 2012.

[8] G. Allaire and M. Amar. Boundary layer tails in periodic homogenization. *ESAIM Control Optim. Calc. Var.*, 4:209–243, 1999.

[9] C. Alves and A. Sapozhnikov. Decoupling inequalities and supercritical percolation for the vacant set of random walk loop soup. Preprint, arXiv:1808.01277.

[10] S. Andres, M. T. Barlow, J.-D. Deuschel, and B. M. Hambly. Invariance principle for the random conductance model. *Probab. Theory Related Fields*, 156(3-4):535–580, 2013.

[11] S. Andres, A. Chiarini, J.-D. Deuschel, and M. Slowik. Quenched invariance principle for random walks with time-dependent ergodic degenerate weights. *Ann. Probab.*, 46(1):302–336, 2018.

[12] S. Andres, J.-D. Deuschel, and M. Slowik. Invariance principle for the random conductance model in a degenerate ergodic environment. *Ann. Probab.*, 43(4):1866–1891, 2015.

[13] S. Andres, J.-D. Deuschel, and M. Slowik. Harnack inequalities on weighted graphs and some applications to the random conductance model. *Probab. Theory Related Fields*, 164(3-4):931–977, 2016.

[14] S. Andres, J.-D. Deuschel, and M. Slowik. Heat kernel estimates for random walks with degenerate weights. *Electron. J. Probab.*, 21, 2016.

[15] S. Andres, J.-D. Deuschel, and M. Slowik. Heat kernel estimates and intrinsic metric for random walks with general speed measure under degenerate conductances. *Electron. Commun. Probab.*, 24:Paper No. 5, 17, 2019.

[16] S. Andres, J.-D. Deuschel, and M. Slowik. Green kernel asymptotics for two-dimensional random walks under random conductances. *Electron. Commun. Probab.*, 25:Paper No. 58, 14, 2020.

[17] P. Antal and A. Pisztora. On the chemical distance for supercritical Bernoulli percolation. *Ann. Probab.*, 24(2):1036–1048, 1996.

[18] S. Armstrong, A. Bordas, and J.-C. Mourrat. Quantitative stochastic homogenization and regularity theory of parabolic equations. *Analysis & PDE*, 11(8):1945–2014, 2018.

[19] S. Armstrong and P. Dario. Elliptic regularity and quantitative homogenization on percolation clusters. *Commun. Pure Appl. Math.*, 71(9):1717–1849, 2018.

[20] S. Armstrong, S. Ferguson, and T. Kuusi. Homogenization, linearization and large-scale regularity for nonlinear elliptic equations. *Comm. Pure Appl. Math.*, to appear.

[21] S. Armstrong, S. J. Ferguson, and T. Kuusi. Higher-order linearization and regularity in nonlinear homogenization. *Arch. Ration. Mech. Anal.*, 237(2):631–741, 2020.

[22] S. Armstrong, A. Hannukainen, T. Kuusi, and J.-C. Mourrat. An iterative method for elliptic problems with rapidly oscillating coefficients. *ESAIM Math. Model. Numer. Anal.*, 55(1):37–55, 2021.

[23] S. Armstrong, T. Kuusi, and J.-C. Mourrat. Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Comm. Math. Phys.*, 347(2):315–361, 2016.

[24] S. Armstrong, T. Kuusi, and J.-C. Mourrat. The additive structure of elliptic homogenization. *Invent. Math.*, 208(3):999–1154, 2017.

[25] S. Armstrong, T. Kuusi, and J.-C. Mourrat. *Quantitative stochastic homogenization and large-scale regularity*, volume 352 of *Grundlehren der mathematischen Wissenschaften.* Springer Nature, 2019.

[26] S. Armstrong, T. Kuusi, J.-C. Mourrat, and C. Prange. Quantitative analysis of boundary layers in periodic homogenization. *Arch. Ration. Mech. Anal.*, 226(2):695–741, 2017.

[27] S. Armstrong and J. Lin. Optimal quantitative estimates in stochastic homogenization for elliptic equations in nondivergence form. *Arch. Ration. Mech. Anal.*, 225(2):937–991, 2017.

[28] S. Armstrong and S. Serfaty. Local laws and rigidity for Coulomb gases at any temperature. *Ann. Probab.*, 49(1):46–121, 2021.

[29] S. Armstrong and W. Wu. C^2 regularity of the surface tension for the $\nabla\phi$ interface model. Preprint, arXiv:1909.13325.

[30] S. N. Armstrong and J.-C. Mourrat. Lipschitz regularity for elliptic equations with random coefficients. *Arch. Ration. Mech. Anal.*, 219(1):255–348, 2016.

[31] S. N. Armstrong and C. K. Smart. Quantitative stochastic homogenization of convex integral functionals. *Ann. Sci. Éc. Norm. Supér. (4)*, 49(2):423–481, 2016.

[32] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, 73:890–896, 1967.

[33] M. Avellaneda and F.-H. Lin. Compactness methods in the theory of homogenization. *Comm. Pure Appl. Math.*, 40(6):803–847, 1987.

[34] M. Avellaneda and F.-H. Lin. Un théorème de Liouville pour des équations elliptiques à coefficients périodiques. *C. R. Acad. Sci. Paris Sér. I Math.*, 309(5):245–250, 1989.

[35] M. Avellaneda and F.-H. Lin. L^p bounds on singular integrals in homogenization. *Comm. Pure Appl. Math.*, 44(8-9):897–910, 1991.

[36] I. Babuska and R. Lipton. Optimal local approximation spaces for generalized finite element methods with application to multiscale problems. *Multiscale Model. Simul.*, 9(1):373–406, 2011.

[37] G. Barbatis. Spectral stability under L^p-perturbation of the second-order coefficients. *J. Differential Equations*, 124(2):302–323, 1996.

[38] G. Barbatis. Stability of weighted Laplace-Beltrami operators under L^p-perturbation of the Riemannian metric. *J. Anal. Math.*, 68:253–276, 1996.

[39] M. T. Barlow. Random walks on supercritical percolation clusters. *Ann. Probab.*, 32(4):3024–3084, 2004.

[40] M. T. Barlow and J.-D. Deuschel. Invariance principle for the random conductance model with unbounded conductances. *Ann. Probab.*, 38(1):234–276, 2010.

[41] M. T. Barlow and B. M. Hambly. Parabolic Harnack inequality and local limit theorem for percolation clusters. *Electron. J. Probab.*, 14:no. 1, 1–27, 2009.

[42] M. Bebendorf and W. Hackbusch. Existence of $\mathscr{H}$-matrix approximants to the inverse FE-matrix of elliptic operators with L^∞-coefficients. *Numer. Math.*, 95(1):1–28, 2003.

[43] P. Bella, B. Fehrman, and F. Otto. A Liouville theorem for elliptic systems with degenerate ergodic coefficients. *Ann. Appl. Probab.*, 28(3):1379–1422, 2018.

[44] P. Bella and M. Schäffner. Quenched invariance principle for random walks among random degenerate conductances. Preprint, arXiv:1902.05793.

[45] I. Benjamini, H. Duminil-Copin, G. Kozma, and A. Yadin. Disorder, entropy and harmonic functions. *Ann. Probab.*, 43(5):2332–2373, 2015.

[46] I. Benjamini and E. Mossel. On the mixing time of a simple random walk on the supercritical percolation cluster. *Probab. Theory Related Fields*, 125(3):408–420, 2003.

[47] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures. Reprint of the 1978 original with corrections and bibliographical additions.* Providence, RI: AMS Chelsea Publishing, reprint of the 1978 original with corrections and bibliographical additions edition, 2011.

[48] A. Bensoussan, J.-L. Lions, and G. C. Papanicolaou. Boundary layers and homogenization of transport processes. *Publ. Res. Inst. Math. Sci.*, 15(1):53–157, 1979.

[49] N. Berger and M. Biskup. Quenched invariance principle for simple random walk on percolation clusters. *Probab. Theory Related Fields*, 137(1-2):83–120, 2007.

[50] N. Berger, M. Biskup, C. E. Hoffman, and G. Kozma. Anomalous heat-kernel decay for random walk among bounded random conductances. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(2):374–392, 2008.

[51] C. Bernardin. Regularity of the diffusion coefficient for lattice gas reversible under Bernoulli measures. *Stochastic Process. Appl.*, 101(1):43–68, 2002.

[52] L. Bertini and B. Zegarlinski. Coercive inequalities for Kawasaki dynamics. The product case. *Markov Process. Related Fields*, 5(2):125–162, 1999.

[53] M. Biskup. Recent progress on the random conductance model. *Probab. Surv.*, 8:294–373, 2011.

[54] M. Biskup and O. Boukhadra. Subdiffusive heat-kernel decay in four-dimensional i.i.d. random conductance models. *J. Lond. Math. Soc. (2)*, 86(2):455–481, 2012.

[55] M. Biskup and T. M. Prescott. Functional CLT for random walk among bounded random conductances. *Electron. J. Probab.*, 12:no. 49, 1323–1348, 2007.

[56] T. Bodineau, I. Gallagher, and L. Saint-Raymond. The Brownian motion as the limit of a deterministic system of hard-spheres. *Invent. Math.*, 203(2):493–553, 2016.

[57] T. Bodineau, I. Gallagher, and L. Saint-Raymond. From hard sphere dynamics to the Stokes-Fourier equations: an L^2 analysis of the Boltzmann-Grad limit. *Ann. PDE*, 3(1):Paper No. 2, 118, 2017.

[58] T. Bodineau, I. Gallagher, and L. Saint-Raymond. Derivation of an Ornstein-Uhlenbeck process for a massive particle in a rarified gas of particles. *Ann. Henri Poincaré*, 19(6):1647–1709, 2018.

[59] T. Bodineau, I. Gallagher, L. Saint-Raymond, and S. Simonella. Statistical dynamics of a hard sphere gas: fluctuating Boltzmann equation and large deviations. Preprint, arXiv:2008.10403.

[60] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities.* Oxford University Press, Oxford, 2013.

[61] O. Boukhadra. Heat-kernel estimates for random walk among random conductances with heavy tail. *Stochastic Process. Appl.*, 120(2):182–194, 2010.

[62] O. Boukhadra. Standard spectral dimension for the polynomial lower tail random conductances model. *Electron. J. Probab.*, 15:no. 68, 2069–2086, 2010.

[63] A. Bourgeat and A. Piatnitski. Approximations of effective coefficients in stochastic homogenization. *Ann. Inst. H. Poincaré Probab. Statist.*, 40(2):153–165, 2004.

[64] W. L. Briggs, V. E. Henson, and S. F. McCormick. *A multigrid tutorial. 2nd ed.* Philadelphia, PA: SIAM, 2nd ed. edition, 2000.

[65] W. L. Briggs, V. E. Henson, and S. F. McCormick. *A multigrid tutorial.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 2000.

[66] T. Brox and H. Rost. Equilibrium fluctuations of stochastic particle systems: the role of conserved quantities. *Ann. Probab.*, 12(3):742–759, 1984.

[67] R. M. Burton and M. Keane. Density and uniqueness in percolation. *Comm. Math. Phys.*, 121(3):501–505, 1989.

[68] N. Cancrini, F. Cesi, and C. Roberto. Diffusive long-time behavior of Kawasaki dynamics. *Electron. J. Probab.*, 10:no. 7, 216–249, 2005.

[69] C. C. Chang. Equilibrium fluctuations of gradient reversible particle systems. *Probab. Theory Related Fields*, 100(3):269–283, 1994.

[70] C.-C. Chang. Equilibrium fluctuations of nongradient reversible particle systems. In *Nonlinear stochastic PDEs (Minneapolis, MN, 1994)*, volume 77 of *IMA Vol. Math. Appl.*, pages 41–51. Springer, New York, 1996.

[71] C. C. Chang and H.-T. Yau. Fluctuations of one-dimensional Ginzburg-Landau models in nonequilibrium. *Comm. Math. Phys.*, 145(2):209–234, 1992.

[72] Y. Chang. Supercritical loop percolation on $\mathbb{Z}^d$ for $d \geqslant 3$. *Stochastic Process. Appl.*, 127(10):3159–3186, 2017.

[73] X. Chen, Z.-Q. Chen, T. Kumagai, and J. Wang. Homogenization of Symmetric Stable-like Processes in Stationary Ergodic Media. *SIAM J. Math. Anal.*, 53(3):2957–3001, 2021.

[74] X. Chen, Z.-Q. Chen, T. Kumagai, and J. Wang. Periodic homogenization of non-symmetric lévy-type processes. Preprint, arXiv:2007.03388.

[75] X. Chen, Z.-Q. Chen, T. Kumagai, and J. Wang. Quenched invariance principle for long range random walks in balanced random environments. Preprint, arXiv:2007.03391.

[76] Z.-Q. Chen, Z. Qian, Y. Hu, and W. Zheng. Stability and approximations of symmetric diffusion semigroups and kernels. *J. Funct. Anal.*, 152(1):255–280, 1998.

[77] A. Chiarini and J.-D. Deuschel. Invariance principle for symmetric diffusions in a degenerate and unbounded stationary and ergodic random medium. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1535–1563, 2016.

[78] A. Cholaquidis, A. Cuevas, and R. Fraiman. On Poincaré cone property. *Ann. Statist.*, 42(1):255–284, 2014.

[79] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to algorithms. 3rd ed.* Cambridge, MA: MIT Press, 3rd ed. edition, 2009.

[80] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization. *Ann. Mat. Pura Appl. (4)*, 144:347–389, 1986.

[81] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization and ergodic theory. *J. Reine Angew. Math.*, 368:28–42, 1986.

[82] P. Dario. Quantitative homogenization of the disordered $\nabla\phi$, model. *Electron. J. Probab.*, 24:Paper No. 90, 99p., 2019.

[83] P. Dario. Optimal corrector estimates on percolation cluster. *Ann. Appl. Probab.*, 31(1):377–431, 2021.

[84] P. Dario. Quantitative homogenization of differential forms. *Ann. Inst. Henri Poincaré Probab. Stat.*, 57(2):1157–1202, 2021.

[85] P. Dario and C. Gu. Quantitative homogenization of the parabolic and elliptic Green's functions on percolation clusters. *Ann. Probab.*, 49(2):556–636, 2021.

[86] P. Dario and W. Wu. Massless phases for the Villain model in $d \geqslant 3$. Preprint, arXiv:2002.02946.

[87] E. B. Davies. Large deviations for heat kernels on graphs. *J. London Math. Soc. (2)*, 47(1):65–72, 1993.

[88] E. De Giorgi and S. Spagnolo. Sulla convergenza degli integrali dell'energia per operatori ellittici del secondo ordine. *Boll. Un. Mat. Ital. (4)*, 8:391–411, 1973.

[89] A. De Masi, P. A. Ferrari, S. Goldstein, and W. D. Wick. An invariance principle for reversible Markov processes. Applications to random motions in random environments. *J. Statist. Phys.*, 55(3-4):787–855, 1989.

[90] A. De Masi, P. A. Ferrari, and J. L. Lebowitz. Reaction-diffusion equations for interacting particle systems. *J. Statist. Phys.*, 44(3-4):589–644, 1986.

[91] A. De Masi, E. Presutti, H. Spohn, and W. D. Wick. Asymptotic equivalence of fluctuation fields for reversible exclusion processes with speed change. *Ann. Probab.*, 14(2):409–423, 1986.

[92] T. Delmotte. Parabolic Harnack inequality and estimates of Markov chains on graphs. *Rev. Mat. Iberoamericana*, 15(1):181–232, 1999.

[93] J.-D. Deuschel. Algebraic L^2 decay of attractive critical processes on the lattice. *Ann. Probab.*, 22(1):264–283, 1994.

[94] J.-D. Deuschel, T. A. Nguyen, and M. Slowik. Quenched invariance principles for the random conductance model on a random graph with degenerate ergodic weights. *Probab. Theory Related Fields*, 170(1-2):363–386, 2018.

[95] F. Du, R. C. Scogna, W. Zhou, S. Brand, J. E. Fischer, and K. I. Winey. Nanotube networks in polymer nanocomposites: rheology and electrical conductivity. *Macromolecules*, 37(24):9048–9055, 2004.

[96] M. Duerinckx and A. Gloria. Analyticity of homogenized coefficients under Bernoulli perturbations and the Clausius-Mossotti formulas. *Arch. Ration. Mech. Anal.*, 220(1):297–361, 2016.

[97] H. Duminil-Copin. Sixty years of percolation. In *Proceedings of the International Congress of Mathematicians—Rio de Janeiro 2018. Vol. IV. Invited lectures*, pages 2829–2856. World Sci. Publ., Hackensack, NJ, 2018.

[98] R. Durrett. *Brownian motion and martingales in analysis.* Wadsworth Mathematics Series. Wadsworth International Group, Belmont, CA, 1984.

[99] W. E. *Principles of multiscale modeling.* Cambridge University Press, Cambridge, 2011.

[100] W. E, P. Ming, and P. Zhang. Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *J. Amer. Math. Soc.*, 18(1):121–156, 2005.

[101] Y. Efendiev, J. Galvis, and T. Y. Hou. Generalized multiscale finite element methods (GMsFEM). *J. Comput. Phys.*, 251:116–135, 2013.

[102] A.-C. Egloffe, A. Gloria, J.-C. Mourrat, and T. N. Nguyen. Random walk in random environment, corrector equation and homogenized coefficients: from theory to numerics, back and forth. *IMA J. Numer. Anal.*, 35(2):499–545, 2015.

[103] B. Engquist and E. Luo. New coarse grid operators for highly oscillatory coefficient elliptic problems. *J. Comput. Phys.*, 129(2):296–306, 1996.

[104] B. Engquist and E. Luo. Convergence of a multigrid method for elliptic equations with highly oscillatory coefficients. *SIAM J. Numer. Anal.*, 34(6):2254–2273, 1997.

[105] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, second edition, 2010.

[106] P. A. Ferrari, E. Presutti, and M. E. Vares. Nonequilibrium fluctuations for a zero range process. *Ann. Inst. H. Poincaré Probab. Statist.*, 24(2):237–268, 1988.

[107] J. Fischer. The choice of representative volumes in the approximation of effective properties of random materials. Preprint, arXiv:1807.00834.

[108] F. Flegel, M. Heida, and M. Slowik. Homogenization theory for the random conductance model with degenerate ergodic weights and unbounded-range jumps. Preprint, arXiv:1702.02860.

[109] M. Fukushima, Y. Ōshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes*, volume 19 of *De Gruyter Studies in Mathematics.* Walter de Gruyter & Co., Berlin, 1994.

[110] T. Funaki. Equilibrium fluctuations for lattice gas. In *Itô's stochastic calculus and probability theory*, pages 63–72. Springer, Tokyo, 1996.

[111] T. Funaki, K. Uchiyama, and H. T. Yau. Hydrodynamic limit for lattice gas reversible under Bernoulli measures. In *Nonlinear stochastic PDEs (Minneapolis, MN, 1994)*, volume 77 of *IMA Vol. Math. Appl.*, pages 1–40. Springer, New York, 1996.

[112] M. Giaquinta and E. Giusti. On the regularity of the minima of variational integrals. *Acta Math.*, 148:31–46, 1982.

[113] M. Giaquinta and G. Modica. Regularity results for some classes of higher order nonlinear elliptic systems. *J. Reine Angew. Math.*, 311/312:145–169, 1979.

[114] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order.* Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[115] A. Giunti, C. Gu, and J.-C. Mourrat. Quantitative homogenization of interacting particle systems. Preprint, arXiv:2011.06366.

[116] A. Giunti, Y. Gu, and J.-C. Mourrat. Heat kernel upper bounds for interacting particle systems. *Ann. Probab.*, 47(2):1056–1095, 2019.

[117] A. Giunti, R. Höfer, and J. Velázquez. Homogenization for the Poisson equation in randomly perforated domains under minimal assumptions on the size of the holes. *Comm. Partial Differential Equations*, 43(9):1377–1412, 2018.

[118] E. Giusti. *Direct methods in the calculus of variations.* World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[119] A. Gloria. Numerical approximation of effective coefficients in stochastic homogenization of discrete elliptic equations. *ESAIM, Math. Model. Numer. Anal.*, 46(1):1–38, 2012.

[120] A. Gloria, S. Neukamm, and F. Otto. An optimal quantitative two-scale expansion in stochastic homogenization of discrete elliptic equations. *ESAIM Math. Model. Numer. Anal.*, 48(2):325–346, 2014.

[121] A. Gloria, S. Neukamm, and F. Otto. Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. *Invent. Math.*, 199(2):455–515, 2015.

[122] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators. *Milan J. Math.*, 88(1):99–170, 2020.

[123] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856, 2011.

[124] A. Gloria and F. Otto. An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28, 2012.

[125] A. Gloria and F. Otto. Quantitative results on the corrector equation in stochastic homogenization. *J. Eur. Math. Soc. (JEMS)*, 19(11):3489–3548, 2017.

[126] A. Gloria and F. Otto. The corrector in stochastic homogenization: optimal rates, stochastic integrability, and fluctuations. Preprint, arXiv:1510.08290.

[127] L. Grafakos. *Classical and modern Fourier analysis.* Pearson Education, Inc., Upper Saddle River, NJ, 2004.

[128] L. Grasedyck, I. Greff, and S. Sauter. The AL basis for the solution of elliptic problems in heterogeneous media. *Multiscale Model. Simul.*, 10(1):245–258, 2012.

[129] M. Griebel and S. Knapek. A multigrid-homogenization method. In *Modeling and computation in environmental sciences (Stuttgart, 1995)*, volume 59 of *Notes Numer. Fluid Mech.*, pages 187–202. Friedr. Vieweg, Braunschweig, 1997.

[130] A. A. Grigor′ yan. The heat equation on noncompact Riemannian manifolds. *Mat. Sb.*, 182(1):55–87, 1991.

[131] G. Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1999.

[132] G. R. Grimmett and I. Manolescu. Inhomogeneous bond percolation on square, triangular and hexagonal lattices. *Ann. Probab.*, 41(4):2990–3025, 2013.

[133] C. Gu. Uniform estimate of an iterative method for elliptic problems with rapidly oscillating coefficients. *Stoch PDE: Anal Comp.*, 8:797–818, 2020.

[134] C. Gu. An efficient algorithm for solving elliptic problems on percolation clusters. Preprint, arXiv:1907.13571.

[135] C. Gu. Decay of semigroup for an infinite interacting particle system on continuum configuration spaces. Preprint, arXiv:2007.04058.

[136] M. Z. Guo, G. C. Papanicolaou, and S. R. S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. *Comm. Math. Phys.*, 118(1):31–59, 1988.

[137] W. Hackbusch. *Multigrid methods and applications*, volume 4 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 1985.

[138] A. Hannukainen, J.-C. Mourrat, and H. Stoppels. Computing homogenized coefficients via multiscale representation and hierarchical hybrid grids. Preprint, arXiv:1905.06751.

[139] I. S. Helland. Central limit theorems for martingales with discrete or continuous time. *Scand. J. Statist.*, 9(2):79–94, 1982.

[140] R. A. Holley and D. W. Stroock. Generalized Ornstein-Uhlenbeck processes and infinite particle branching Brownian motions. *Publ. Res. Inst. Math. Sci.*, 14(3):741–788, 1978.

[141] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.

[142] E. Janvresse, C. Landim, J. Quastel, and H. T. Yau. Relaxation to equilibrium of conservative dynamics. I. Zero-range processes. *Ann. Probab.*, 27(1):325–360, 1999.

[143] M. Jara. Finite-dimensional approximation for the diffusion coefficient in the simple exclusion process. *Ann. Probab.*, 34(6):2365–2381, 2006.

[144] M. Jara and O. Menezes. Non-equilibrium fluctuations of interacting particle systems. Preprint, arXiv:1810.09526.

[145] V. V. Jikov, S. M. Kozlov, and O. A. Oleunik. *Homogenization of differential operators and integral functionals*. Springer-Verlag, Berlin, 1994. Translated from the Russian by G. A. Yosifian.

[146] C. Kenig, F. Lin, and Z. Shen. Periodic homogenization of green and neumann functions. *Communications on Pure and Applied Mathematics*, 67(8):1219–1262, 2014.

[147] C. E. Kenig, F. Lin, and Z. Shen. Convergence rates in L^2 for elliptic homogenization problems. *Arch. Ration. Mech. Anal.*, 203(3):1009–1036, 2012.

[148] C. E. Kenig, F. Lin, and Z. Shen. Homogenization of elliptic systems with Neumann boundary conditions. *J. Am. Math. Soc.*, 26(4):901–937, 2013.

[149] H. Kesten. *Percolation theory for mathematicians.*, volume 2. Birkhäuser, Boston/Basel/Stuttgart, 1982.

[150] H. Kesten and Y. Zhang. The probability of a large finite cluster in supercritical Bernoulli percolation. *Ann. Probab.*, 18(2):537–555, 1990.

[151] J. F. C. Kingman. *Poisson processes*, volume 3 of *Oxford Studies in Probability.* The Clarendon Press, Oxford University Press, New York, 1993.

[152] C. Kipnis and C. Landim. *Scaling limits of interacting particle systems*, volume 320 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, Berlin, 1999.

[153] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Commun. Math. Phys.*, 104:1–19, 1986.

[154] S. Knapek. Matrix-dependent multigrid homogeneization for diffusion problems. *SIAM J. Sci. Comput.*, 20(2):515–533, 1998.

[155] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent RV'-s, and the sample DF. I. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 32(1-2):111–131, 1975.

[156] J. Komlós, P. Major, and G. Tusnády. An approximation of partial sums of independent RV's, and the sample DF. II. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 34(1):33–58, 1976.

[157] T. Komorowski, C. Landim, and S. Olla. *Fluctuations in Markov processes*, volume 345 of *Grundlehren der Mathematischen Wissenschaften.* Springer, Heidelberg, 2012.

[158] V. Konakov and E. Mammen. Local approximations of Markov random walks by diffusions. *Stochastic Process. Appl.*, 96(1):73–98, 2001.

[159] R. Kornhuber and H. Yserentant. Numerical homogenization of elliptic multiscale problems by subspace decomposition. *Multiscale Model. Simul.*, 14(3):1017–1036, 2016.

[160] S. M. Kozlov. Averaging of differential operators with almost periodic rapidly oscillating coefficients. *Mat. Sb. (N.S.)*, 107(149)(2):199–217, 317, 1978.

[161] S. M. Kozlov. Averaging of random operators. *Math. USSR, Sb.*, 37:167–180, 1980.

[162] T. Kumagai. *Random walks on disordered media and their scaling limits*, volume 2101 of *Lecture Notes in Mathematics.* Springer, Cham, 2014.

[163] A. V. Kyrylyuk and P. van der Schoot. Continuum percolation of carbon nanotubes in polymeric and colloidal media. *Proceedings of the National Academy of Sciences*, 105(24):8221–8226, 2008.

[164] A. Lamacz, S. Neukamm, and F. Otto. Moment bounds for the corrector in stochastic homogenization of a percolation model. *Electron. J. Probab.*, 20:30, 2015. Id/No 106.

[165] C. Landim, S. Olla, and S. R. S. Varadhan. Symmetric simple exclusion process: regularity of the self-diffusion coefficient. *Comm. Math. Phys.*, 224(1):307–321, 2001. Dedicated to Joel L. Lebowitz.

[166] C. Landim, S. Olla, and S. R. S. Varadhan. Finite-dimensional approximation of the self-diffusion coefficient for the exclusion process. *Ann. Probab.*, 30(2):483–508, 2002.

[167] C. Landim, S. Olla, and S. R. S. Varadhan. On viscosity and fluctuation-dissipation in exclusion processes. *J. Statist. Phys.*, 115(1-2):323–363, 2004.

[168] C. Landim and H. T. Yau. Convergence to equilibrium of conservative particle systems on $\mathbb{Z}^d$. *Ann. Probab.*, 31(1):115–147, 2003.

[169] G. Last and M. Penrose. *Lectures on the Poisson process*, volume 7 of *Institute of Mathematical Statistics Textbooks.* Cambridge University Press, Cambridge, 2018.

[170] G. F. Lawler and V. Limic. *Random walk: A modern introduction.*, volume 123. Cambridge: Cambridge University Press, 2010.

[171] G. Leoni. *A first course in Sobolev spaces*, volume 181 of *Graduate Studies in Mathematics.* American Mathematical Society, Providence, RI, second edition, 2017.

[172] T. M. Liggett. *Interacting particle systems*, volume 276 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, New York, 1985.

[173] T. M. Liggett. L_2 rates of convergence for attractive reversible nearest particle systems: the critical case. *Ann. Probab.*, 19(3):935–959, 1991.

[174] S. L. Lu. Equilibrium fluctuations of a one-dimensional nongradient Ginzburg-Landau model. *Ann. Probab.*, 22(3):1252–1272, 1994.

[175] X. Luo and P. T. Mather. Conductive shape memory nanocomposites for high speed electrical actuation. *Soft Matter*, 6(10):2146–2149, 2010.

[176] T. J. Lyons and T. S. Zhang. Note on convergence of Dirichlet processes. *Bull. London Math. Soc.*, 25(4):353–356, 1993.

[177] Z. M. Ma and M. Röckner. *Introduction to the theory of (nonsymmetric) Dirichlet forms.* Universitext. Springer-Verlag, Berlin, 1992.

[178] A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. *Math. Comp.*, 83(290):2583–2603, 2014.

[179] P. Mathieu. Quenched invariance principles for random walks with random conductances. *J. Stat. Phys.*, 130(5):1025–1046, 2008.

[180] P. Mathieu and A. Piatnitski. Quenched invariance principles for random walks on percolation clusters. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 463(2085):2287–2307, 2007.

[181] P. Mathieu and E. Remy. Isoperimetry and heat kernel decay on percolation clusters. *Ann. Probab.*, 32(1A):100–128, 2004.

[182] J.-C. Mourrat. Variance decay for functionals of the environment viewed by the particle. *Ann. Inst. Henri Poincaré Probab. Stat.*, 47(1):294–327, 2011.

[183] J.-C. Mourrat. First-order expansion of homogenized coefficients under Bernoulli perturbations. *J. Math. Pures Appl. (9)*, 103(1):68–101, 2015.

[184] J.-C. Mourrat. Efficient methods for the estimation of homogenized coefficients. *Found. Comput. Math.*, 19(2):435–483, 2019.

[185] J.-C. Mourrat. An informal introduction to quantitative stochastic homogenization. *J. Math. Phys.*, 60(3):031506, 11, 2019.

[186] F. Murat. Compacité par compensation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 5(3):489–507, 1978.

[187] F. Murat and L. Tartar. H-convergence. In *Topics in the mathematical modelling of composite materials*, volume 31 of *Progr. Nonlinear Differential Equations Appl.*, pages 21–43. Birkhäuser Boston, Boston, MA, 1997.

[188] A. Naddaf and T. Spencer. Estimates on the variance of some homogenization problems. 1998, unpublished preprint.

[189] Y. Nagahata. Regularity of the diffusion coefficient matrix for the lattice gas with energy. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(1):45–67, 2005.

[190] Y. Nagahata. Regularity of the diffusion coefficient matrix for generalized exclusion process. *Stochastic Process. Appl.*, 116(6):957–982, 2006.

[191] Y. Nagahata. Regularity of the diffusion coefficient matrix for lattice gas reversible under Gibbs measures with mixing condition. *Comm. Math. Phys.*, 273(3):637–650, 2007.

[192] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958.

[193] S. Neukamm, M. Schaffner, and A. Schlomerkemper. Stochastic homogenization of nonconvex discrete energies with degenerate growth. *SIAM J. Math. Anal.*, 49(3):1761–1809, 2017.

[194] G. Nguetseng. A general convergence result for a functional related to the theory of homogenization. *SIAM J. Math. Anal.*, 20(3):608–623, 1989.

[195] H. Osada. Homogenization of diffusion processes with random stationary coefficients. In *Probability theory and mathematical statistics (Tbilisi, 1982)*, volume 1021 of *Lecture Notes in Math.*, pages 507–517. Springer, Berlin, 1983.

[196] H. Owhadi. Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games. *SIAM Rev.*, 59(1):99–149, 2017.

[197] H. Owhadi, L. Zhang, and L. Berlyand. Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization. *ESAIM Math. Model. Numer. Anal.*, 48(2):517–552, 2014.

[198] G. C. Papanicolaou and S. R. S. Varadhan. Boundary value problems with rapidly oscillating random coefficients. In *Random fields, Vol. I, II (Esztergom, 1979)*, volume 27 of *Colloq. Math. Soc. János Bolyai*, pages 835–873. North-Holland, Amsterdam, 1981.

[199] G. A. Pavliotis and A. M. Stuart. *Multiscale methods*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008. Averaging and homogenization.

[200] M. Penrose and A. Pisztora. Large deviations for discrete and continuous percolation. *Adv. in Appl. Probab.*, 28(1):29–52, 1996.

[201] A. Pisztora. Surface order large deviations for Ising, Potts and percolation models. *Probab. Theory Related Fields*, 104(4):427–466, 1996.

[202] E. Presutti and H. Spohn. Hydrodynamics of the voter model. *Ann. Probab.*, 11(4):867–875, 1983.

[203] E. Procaccia, R. Rosenthal, and A. Sapozhnikov. Quenched invariance principle for simple random walk on clusters in correlated percolation models. *Probab. Theory and Related Fields*, 166(3-4):619–657, 2016.

[204] J. Quastel. Diffusion of color in the simple exclusion process. *Comm. Pure Appl. Math.*, 45(6):623–679, 1992.

[205] M. Reed and B. Simon. *Methods of modern mathematical physics. I. Functional analysis*. Academic Press, New York-London, 1972.

[206] M. Röckner. Stochastic analysis on configuration spaces: basic ideas and recent results. In *New directions in Dirichlet forms*, volume 8 of *AMS/IP Stud. Adv. Math.*, pages 157–231. Amer. Math. Soc., Providence, RI, 1998.

[207] Y. Saad. *Iterative methods for sparse linear systems*. Society for Industrial and Applied Mathematics, Philadelphia, PA, second edition, 2003.

[208] L. Saloff-Coste. A note on Poincaré, Sobolev, and Harnack inequalities. *Internat. Math. Res. Notices*, (2):27–38, 1992.

[209] A. Sapozhnikov. Random walks on infinite percolation clusters in models with long-range correlations. *Ann. Probab.*, 45(3):1842–1898, 2017.

[210] Z. Shen. *Periodic homogenization of elliptic systems*, volume 269 of *Operator Theory: Advances and Applications*. Birkhäuser/Springer, Cham, 2018. Advances in Partial Differential Equations (Basel).

[211] V. Sidoravicius and A.-S. Sznitman. Quenched invariance principles for walks on clusters of percolation or among random conductances. *Probab. Theory Related Fields*, 129(2):219–244, 2004.

[212] S. Spagnolo. Sulla convergenza di soluzioni di equazioni paraboliche ed ellittiche. *Ann. Scuola Norm. Sup. Pisa (3) 22 (1968), 571-597; errata, ibid. (3)*, 22:673, 1968.

[213] S. Spagnolo. Convergence in energy for elliptic operators. In *Numerical Solution of Partial Differential Equations–III*, pages 469–498. 1976.

[214] H. Spohn. Equilibrium fluctuations for interacting Brownian particles. *Comm. Math. Phys.*, 103(1):1–33, 1986.

[215] H. Spohn. *Large scale dynamics of interacting particles*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 1991.

[216] E. M. Stein. *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[217] M. Sued. Regularity properties of the diffusion coefficient for a mean zero exclusion process. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(1):1–33, 2005.

[218] L. Tartar. Compensated compactness and applications to partial differential equations. In *Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV*, volume 39 of *Res. Notes in Math.*, pages 136–212. Pitman, Boston, Mass.-London, 1979.

[219] L. Tartar. *The general theory of homogenization. A personalized introduction.*, volume 7. Berlin: Springer, 2009.

[220] S. Torquato. *Random heterogeneous materials*, volume 16 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 2002. Microstructure and macroscopic properties.

[221] S. R. S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. II. In *Asymptotic problems in probability theory: stochastic models and diffusions on fractals (Sanda/Kyoto, 1990)*, volume 283 of *Pitman Res. Notes Math. Ser.*, pages 75–128. Longman Sci. Tech., Harlow, 1993.

[222] S. R. S. Varadhan. Regularity of self-diffusion coefficient. In *The Dynkin Festschrift*, volume 34 of *Progr. Probab.*, pages 387–397. Birkhäuser Boston, Boston, MA, 1994.

[223] P. Wesseling. *An introduction to multigrid methods.* Pure and Applied Mathematics (New York). John Wiley & Sons, Ltd., Chichester, 1992.

[224] H.-T. Yau. Relative entropy and hydrodynamics of Ginzburg-Landau models. *Lett. Math. Phys.*, 22(1):63–80, 1991.

[225] V. V. Yurinskiĭ. Averaging of symmetric diffusion in a random medium. *Sibirsk. Mat. Zh.*, 27(4):167–180, 215, 1986.

[226] V. V. Zhikov and A. L. Pyatniskii. Homogenization of random singular structures and random measures. *Izv. Math.*, 70(1):19–67, 2006.

RÉSUMÉ

Cette thèse étudie l'interaction entre la théorie de l'homogénéisation quantitative et deux modèles stochastiques : le modèle de percolation surcritique et les systèmes de particules en interaction. L'homogénéisation stochastique se concentre sur les propriétés de grande échelle dans l'environnement aléatoire, et ces deux modèles représentent respectivement la généralisation dans l'environnement aléatoire dégénéré et dans l'environnement aléatoire dynamique.
Dans les chapitres 2 et 3, nous étudions un algorithme efficace pour le problème de Dirichlet avec des coefficients aléatoires. Cet algorithme est proposé par Armstrong, Hannukainen, Kuusi et Mourrat, et permet une approximation dans H^1 pour la solution avec une grande précision et un faible coût de calcul. Nous confirmons sa cohérence au chapitre 2, puis nous l'étendons au modèle de percolation surcritique au chapitre 3.
Le chapitre 4 est consacré à l'homogénéisation quantitative du semigroupe pour la marche aléatoire sur l'amas de percolation surcritique infini. Son interprétation probabiliste est un théorème limite central local quantitatif, et ce résultat implique également le taux de convergence de la fonction de Green elliptique. La preuve dans ce chapitre combine plusieurs estimations quantitatives sur le modèle de percolation : les correcteurs de premier ordre, le flux, le développement à deux échelles, et aussi la concentration de la densité de l'amas.
Dans les chapitres 5 et 6, nous développons la théorie de l'homogénéisation pour un système de particules en interaction sans condition de gradient dans l'espace de configuration continu. Dans le chapitre 5, nous construisons ce modèle et prouvons une décroissance de la variance de type gaussien. Dans le chapitre 6, nous étudions son coefficient de diffusion global, et obtenons un taux de convergence pour l'approximation en volume fini. Notre stratégie est l'approche de sous-additivité et renormalisation, et nous développons également de nouvelles inégalités fonctionnelles adaptées à cette situation de dimension infinie.

MOTS CLÉS

Homogénéisation, algorithmes numériques, percolation, systèmes de particules en interaction.

ABSTRACT

This thesis studies the interaction between quantitative homogenization theory and two stochastic models: the supercritical percolation model and interacting particle systems. Stochastic homogenization focuses on large-scale properties in the random environment, and these two models represent generalization in the degenerate random environment and the dynamic random environment, respectively.
In chapters 2 and 3, we study an efficient algorithm for the Dirichlet problem with random coefficients. This algorithm is proposed by Armstrong, Hannukainen, Kuusi and Mourrat, which allows an approximation in H^1 for the solution with high precision and low computational cost. We confirm its consistency in chapter 2, then extend it to the supercritical percolation model in chapter 3.
Chapter 4 is devoted to the quantitative homogenization of the semigroup for the random walk on the infinite supercritical percolation cluster. Its probabilistic interpretation is as a quantitative local central limit theorem, and it also implies the convergence rate of the elliptic Green's function. The proof in this chapter combines several quantitative estimates on the percolation model: first-order correctors, flux, two-scaled expansion, and also the cluster density concentration.
In chapters 5 and 6, we develop the homogenization theory for an interacting particle system without gradient condition in continuum configuration space. In chapter 5, we construct this model and prove its variance decay of Gaussian type. In chapter 6, we study its bulk diffusion coefficient and obtain a rate of convergence for the finite-volume approximation. Our strategy is the subadditivity-renormalization approach, and we also develop new functional inequalities adapted to this infinite-dimensional setting.

KEYWORDS

Homogenization, numerical algorithms, percolation, interacting particle systems.