MASSLESS PHASES FOR THE VILLAIN MODEL IN $d \geq 3$

PAUL DARIO, WEI WU

Abstract. A major open question in statistical mechanics, known as the Gaussian spin wave conjecture, predicts that the low temperature phase of the Abelian spin systems with continuous symmetry behave like Gaussian free fields. In this paper we consider the classical Villain rotator model in $\mathbb{Z}^d, d \geq 3$ at sufficiently low temperature, and prove that the truncated two-point function decays asymptotically as $|x|^{2-d}$, with an algebraic rate of convergence. We also obtain the same asymptotic decay separately for the transversal two-point functions. This quantifies the spontaneous magnetization result for the Villain model at low temperatures and constitutes a first step toward a more precise understanding of the spin-wave conjecture. We believe that our method extends to finite range interactions, and to other Abelian spin systems and Abelian gauge theory in $d \geq 3$. We also develop a quantitative perspective on homogenization of uniformly convex gradient Gibbs measures.

Contents

1. Introduction

1.1. **Rotator models and the spin wave picture.** Rotator models, such as the XY and the Villain models, have drawn considerable attention from distinct research communities in mathematical and theoretical physics. They are of much interest in statistical mechanics, as they exhibit new types of phase transition for ferromagnetic systems and can be applied to the design of novel materials. A canonical rotator model is the XY model defined as follows: given a finite set $U \subseteq \mathbb{Z}^d$, we assign to each function $\theta : U \to (-\pi, \pi]$ satisfying $\theta = 0$ on the external vertex boundary ∂U the energy

$$H_U^{XY}(\theta) := -\sum_{\substack{x,y \in U^+ \\ x \sim y}} \cos(\theta(x) - \theta(y)),$$

where $U^+ := U \cap \partial U$ and the notation $x \sim y$ means that the points x and y are nearest neighbor in the lattice $\mathbb{Z}^d$. The Gibbs measure of the XY model with zero boundary condition at inverse temperature $\beta > 0$ is then defined the probability distribution

$$d\mu_{\beta,U}^{XY}(d\theta) := \frac{1}{Z_{\beta,U}^{XY}} \exp\left(-\beta H_U^{XY}(\theta)\right) \prod_{x \in U} \mathrm{d}\theta(x) \mathbb{1}_{\theta|_{\partial U} = 0}. \tag{1.1}$$

Date: March 27, 2023.

The XY model can be equivalently seen as a spin system with spin valued in the circle $\mathbb{S}^1$ by setting $S_x := e^{i\theta_x}$. In this article, we will be interested in another closely related rotator model, the Villain model [90] is defined by the Gibbs weight

$$d\mu^{\mathrm{Vill}}_{\beta,U}(d\theta) := \frac{1}{Z^{\mathrm{Vill}}_U} \prod_{x\sim y} v_\beta(\theta(x)-\theta(y))\, \mathrm{d}\theta(x) \mathbb{1}_{\theta|_{\partial U}=0}, \tag{1.2}$$

where is the heat kernel on $\mathbb{S}^1$ defined according to the identity

$$v_\beta(\theta) := \sum_{m\in\mathbb{Z}} \exp\left(-\frac{\beta}{2}(\theta+2\pi m)^2\right). \tag{1.3}$$

The two models belong to the class of spin systems with continuous Abelian symmetry. They exhibit a similar behavior and have been extensively studied in the literature. We collect below some of their main features.

Since the spins take values in the compact space $\mathbb{S}^1$, the existence of a thermodynamic limit for the XY model (i.e., an infinite-volume limit as $U\to\infty$) is guaranteed along subsequences by standard compactness arguments. It is additionally known that this limit is unique, and we denote it by μ^{XY}_β (see [78]). The Griffiths correlation inequalities [57, 25, 78] imply that that the expected value of the spins and the two-point function are monotone in the domain U and in particular show the convergences

$$\langle S_x\rangle_{\mu^{XY}_{\beta,U}} \underset{U\uparrow\mathbb{Z}^d}{\longrightarrow} \langle S_x\rangle_{\mu^{XY}_\beta} \quad \text{and} \quad \langle S_x\cdot S_y\rangle_{\mu^{XY}_{\beta,U}} \underset{U\uparrow\mathbb{Z}^d}{\longrightarrow} \langle S_x\cdot S_y\rangle_{\mu^{XY}_\beta}.$$

The same results hold for the Villain model, and we denote by μ^V_β the corresponding thermodynamic limit[1].

In two dimensions, the Mermin-Wagner theorem [77] shows that there is no continuous symmetry breaking at any temperature, i.e., for any $\beta>0$,

$$\langle S_x\rangle_{\mu^{XY}_\beta} = 0. \tag{1.4}$$

In particular, the system does not undergo an order/disorder phase transition. Nevertheless, the system is known to exhibit a phase transition of a different type, characterized by a different asymptotic behavior of the correlation function: there exists a critical inverse temperature $\beta_c\in(0,\infty)$ such that in the low temperature regime ($\beta>\beta_c$), the two-point function $\langle S_x\cdot S_0\rangle_{\mu_\beta}$ decays polynomially fast (which characterizes a so-called topological long-range order [71]), while, in the high temperature regime ($\beta\le\beta_c$), the two-point function decays exponentially fast. This phase transition is known as the Berezinskii–Kosterlitz–Thouless transition became the basis of the Nobel prize in Physics in 2016 to Haldane, Kosterlitz and Thouless. From a mathematical perspective, the existence of this transition was established in the celebrated work of Fröhlich and Spencer [49], and has been the subject of recent developments [74, 89, 4].

In the low temperature regime ($\beta>\beta_c$), additional predictions can be made regarding the behavior of the model. A simple heuristics suggests that, as the temperature goes to zero, the spins tend to align with each other so as to minimize the Hamiltonian. Using the approximations, when $|\delta\theta|\ll 1$,

$$\exp(\beta\cos(\delta\theta)) \approx \sum_{m\in\mathbb{Z}} \exp\left(-\frac{\beta}{2}(\delta\theta+2\pi m)^2\right) \quad \text{and} \quad \cos\left(2\pi\left(\delta\theta\right)\right) \approx 1-\left(\delta\theta\right)^2/2, \tag{1.5}$$

it is expected that at low temperature, both the XY and the Villain Gibbs measures on large scales behave like the Gaussian measure

$$\mu^{GFF}_\beta(d\phi) := \frac{1}{Z}\exp\left(-\frac{\beta}{2}\sum_{x\sim y}(\phi(x)-\phi(y))^2\right)\prod_x d\phi(x). \tag{1.6}$$

The Gibbs measure (1.6) is the Gaussian free field, and its law is fully characterized by its covariance matrix given by the lattice Green's function. This heuristic computation is the starting point of the celebrated spin wave picture originating in the work of Dyson [42] (see also [77]). The spin wave conjecture predicts that at low temperatures both the XY and the Villain Gibbs measures behave on large scales like a Gaussian free field of the form (1.6) with a notable difference: since the approximations (1.5) are not exact (and does not recover the information of the periodized field in (1.1) and (1.2)), a corrective term, corresponding to the so-called vortex lines, has to be taken into account in the analysis, and the limiting Gaussian free field describing the

[1]The monotonicity of the correlation function and the uniqueness of the infinite volume Gibbs state were first established for the XY model [78]. However, the Villain model can be represented as a metric graph limit of the XY model [84, 49]. By taking this limit, we obtain the corresponding monotonicity and the uniqueness of Gibbs state for the Villain model.

large-scale behaviors of the XY and Villain models should display an effective temperature $\beta_{\text{eff}} \neq \beta$ (with $\beta_{\text{eff}} = (1 + o(1))\beta$ as $\beta \to \infty$).

More precisely, the spin wave picture in the case of the two-point function asserts that, for $d = 2$ and $\beta > \beta_c$, there exists an effective inverse temperature $\beta_{\text{eff}} > 0$ (with $\beta_{\text{eff}} \neq \beta$) such that

$$\left\langle e^{i(\theta(0)-\theta(x))} \right\rangle_{\mu_\beta^V} = \left\langle e^{i(\phi(0)-\phi(x))} \right\rangle_{\mu_{\beta_{\text{eff}}}^{GFF}} (1 + o(1)) = |x|^{-\frac{1}{2\pi\beta_{\text{eff}}}} + o\left(|x|^{-\frac{1}{2\pi\beta_{\text{eff}}}}\right). \tag{1.7}$$

Rigorous (but non-optimal) power law upper and lower bounds for the two-point function were established in the 1980s in the celebrated works of McBryan-Spencer [76] and Fröhlich-Spencer [48] in the low temperature regime, namely, for $\beta \gg 1$,

$$c_1 |x|^{-\frac{1}{2\pi\beta_1}} \leq \langle S_0 \cdot S_x \rangle_{\mu_\beta^V} \leq c_2 |x|^{-\frac{1}{2\pi\beta}},$$

where $\beta_1 = \beta_1(\beta)$ and satisfies $\beta_1 = \beta(1 + o(1))$ as $\beta \to \infty$. For a closely related model, the two dimensional two-component Coulomb gas with small activity, Falco justified the spin wave picture (with an effective β_{eff} in the exponent) for all the inverse temperatures in the Kosterlitz-Thouless phase, in a series of impressive works [44, 45]. For the two dimensional XY and Villain models, the asymptotic two point function (1.7) still remains an important open question.

In three dimensions and higher, the breakthrough work of Fröhlich, Simon and Spencer [47] shows that these models undergo an order/disorder phase transitions: there exists an critical inverse temperature $\beta_c > 0$ such that

$$\text{for any } \beta > \beta_c, \ \langle S_x \rangle_{\mu_\beta} \neq 0 \quad \text{and} \quad \text{for any } \beta < \beta_c, \ \langle S_x \rangle_{\mu_\beta} = 0.$$

In the low temperature phase ($\beta > \beta_c$), the spin wave picture predicts that there exist two coefficients c_1, c_2 such that

$$\langle S_0 \cdot S_x \rangle_{\mu_\beta^V} = c_1 + \frac{c_2}{|x|^{d-2}} + o\left(\frac{1}{|x|^{d-2}}\right). \tag{1.8}$$

Considerable progress towards quantitative information for the XY/Villain model at low temperature were made in the 1980s. In dimensions $d \geq 3$, the best known result is the one of Fröhlich and Spencer [50] who observed that the classical Villain model in $\mathbb{Z}^d$ can be mapped, via duality, to a statistical mechanical model of lattice Coulomb gas. They obtained the following next order description of the correlation function at low temperature.

Proposition 1.1 (Fröhlich and Spencer [50]). *Let μ_β^V be the thermodynamic limit of the Villain model in $\mathbb{Z}^d$, for $d \geq 3$. There exist constants $\beta_0 = \beta_0(d)$, $c_0 = c_0(\beta, d)$, such that for all $\beta > \beta_0$,*

$$\langle S_0 \cdot S_x \rangle_{\mu_\beta^V} = c_0 + O\left(\frac{1}{|x|^{d-2}}\right).$$

Moreover, denote by G the lattice Green's function in $\mathbb{Z}^d$, then we have as $\beta \to \infty$,

$$\exp\left(\frac{1}{\beta}(G(0) - G(x))\right) \geq \langle S_0 \cdot S_x \rangle_{\mu_\beta^V} \geq \exp\left(\left(\frac{1}{\beta} + o\left(\frac{1}{\beta}\right)\right)(G(0) - G(x))\right).$$

This suggests that the truncated two-point function may be related to a massless free field in $\mathbb{R}^d$, which corresponds to the emergence of a (conjectured) Goldstone boson. Similar results were also obtained for the Abelian gauge theory in four dimensions (see [50, 65]). Kennedy and King in [69] obtained a similar low temperature expansion for the Abelian Higgs model, which couples an XY model with a gauge fixing potential. Their proofs rely on a different approach, via a transformation introduced by [14] and a polymer expansion.

It is also of much interest to justify the spin wave conjecture separately for the longitudinal and transversal two-point functions of the rotator models, i.e., observables of the form $\langle \cos\theta(0) \cos\theta(x) \rangle_{\mu_\beta^{XY}}$ and $\langle \sin\theta(0) \sin\theta(x) \rangle_{\mu_\beta^{XY}}$. The best known result is due to Bricmont, Fontaine, Lebowitz, Lieb, and Spencer [26], where, relies on a combination of the infrared bound [47], a Mermin-Wagner type argument, and correlation inequalities, they perform a low temperature expansion of the truncated correlation function of the XY model and obtain the following expansion.

Proposition 1.2 (Bricmont, Fontaine, Lebowitz, Lieb, and Spencer [26]). *There exist an inverse temperature $\beta_1 < \infty$ and two constants $c_1 > c_2 > 0$ such that, for any $\beta \geq \beta_1$,*

$$\frac{c_2}{\beta |x|^{d-2}} \leq \langle \sin\theta(0) \sin\theta(x) \rangle_{\mu_\beta^{XY}} \leq \frac{c_1}{\beta |x|^{d-2}}.$$

Despite these considerable progress, the rigorous derivations of the spin wave conjecture (1.8) remain largely open. The main result of our paper, stated below, identifies the next-order term for the Villain model in dimensions three and higher, by obtaining the precise asymptotics of the two-point functions at low temperature.

Theorem 1. *For any dimension $d \geq 3$, there exist $\beta_0 = \beta_0(d)$ and $\alpha = \alpha(d) > 0$ such that, for any $\beta \geq \beta_0$, there exist constants $c_0 = c_0(\beta, d), c_1 = c_1(\beta, d), c_2 = c_2(\beta, d)$, and such that, for all $\beta > \beta_0$, the transversal two-point function has the asymptotics*

$$\langle \sin\theta(0) \sin\theta(x) \rangle_{\mu_\beta^V} = \frac{c_2}{|x|^{d-2}} + O\left(\frac{1}{|x|^{d-2+\alpha}}\right), \tag{1.9}$$

and the spin-spin correlation function satisfies

$$\langle S_0 \cdot S_x \rangle_{\mu_\beta^V} = c_0 + \frac{c_1}{|x|^{d-2}} + O\left(\frac{1}{|x|^{d-2+\alpha}}\right). \tag{1.10}$$

Remark 1.3. The proof of Theorem 1 yields the following characterization for the constant c_0

$$c_0 = \langle S_0 \rangle_{\mu_\beta^V}^2 .$$

Regarding the constants c_1 and c_2, the free field computation (1.5) indicates that they should be close to the constant

$$C = -\frac{1}{\beta} \exp\left(G(0)/\beta\right) \frac{\Gamma(d/2-1)}{4\pi^{d/2}},$$

where Γ is the standard Gamma function. The constant C is defined so as to satisfy

$$\left\langle e^{i(\phi(0)-\phi(x))} \right\rangle_{\mu_\beta^{GFF}} = \exp\left(\frac{1}{\beta}(G(0) - G(x))\right) = \exp\left(G(0)/\beta\right) + \frac{C}{|x|^{d-2}} + O\left(\frac{1}{|x|^{d-1}}\right).$$

In this direction, the proof of Theorem 1 yields the identities

$$c_1 = C + O(e^{-c\beta}) \quad \text{and} \quad c_2 = -C + O(e^{-c\beta}).$$

Remark 1.4. It follows from (1.9) and (1.10) that the two-point correlation function is asymptotically rotation invariant. Indeed, the proof yields rotation invariance for the Villain Gibbs measures that are invariant under the $\pi/2$-degree rotations and the reflections of the lattice. For more general Villain models, i.e., replacing the potential (1.3) by

$$v_{\beta,x,y}(\theta) \coloneqq \sum_{m\in\mathbb{Z}} \exp\left(-\frac{\beta J_{x,y}}{2}(\theta + 2\pi m)^2\right),$$

for strictly positive, nearest neighbor and periodic coupling constants $J_{x,y}$, one expects the second order term to take the form of a more general $(2-d)$-homogeneous function.

We remark here that an alternative approach, based on elaborate renormalization group analysis, was developed in a series of works of Balaban, and culminated in [15]. They studied a class of Euclidean field theories that are invariant under the $O(N)$ symmetry group, for $N \geq 2$, and obtained results similar to Theorem 1 for these models.

We conclude the introduction by mentioning two open questions. The Gaussian spin wave approximation predicts that the two-point function of the XY model in $d \geq 3$ also admits a low temperature expansion like that stated in Theorem 1. The main challenge is a technical one: in the first step of the proof (described in Section 1.2 below), a duality transformation and a cluster expansion step are used to prove that the model can be expressed as gradient model with a strictly convex potential (this part of the proof follows well-known arguments [50, 16]). The specific structure of the Hamiltonian of the Villain model (1.2) allows an exact factorisation (in particular, the two-point function can be factorized as a Gaussian contribution and a vortex contribution, see Section 3, (3.6)). Such an exact factorization does not hold for the XY model and a new idea for renormalization is required to implement the argument.

The spin wave conjecture and the asymptotic two-point function (1.7) remains open for the XY and Villain model in $d = 2$. The renormalization argument developed by Falco [44, 45] does not directly apply, because by applying a duality transform to the XY and Villain model, one obtains a lattice Coulomb gas with infinite activity (instead of small activity). Building new insights into the renormalization group analysis, Bauerschmidt, Park and Rodriguez showed recently that the scaling limit of the two-dimensional Discrete Gaussian at high temperature is a continuous Gaussian free field (with an effective inverse temperature) in [17, 18]. Their result makes another progress toward the spin-wave conjecture for the two-dimensional Villain model in the low temperature regime ($\beta \gg 1$). Resolving the conjecture requires extending the results of [17, 18] to more singular test functions.

1.2. **Strategy of the proof.** We initiate a renormalization-Helffer-Sjöstrand-homogenization program to prove Theorem 1. The periodic potential of the XY and Villain model makes the interaction highly non-convex, and poses significant challenges to study their large scale behavior. Indeed, the ground states at zero temperature already leads to highly nontrivial variational problems (see, e.g. [5]). To overcome these difficulties, we start from the insight of Fröhlich and Spencer [50] (see also [16, Section 5]), applying a duality transformation and a cluster expansion to the Villain Gibbs measure. In the low temperature regime ($\beta \gg 1$), this argument shows that two-point function can be expressed as a non-linear and non-local observable of a uniformly convex gradient model (or uniformly convex $\nabla\phi$ model). Contrary to the Villain and XY models, tools from PDE and homogenization theory can be applied to study the behavior over large-scales of the uniformly convex $\nabla\phi$ model (see Section 1.2.2), which can thus be used to study the Villain model via the duality transformation of [50]. The general strategy described above encounters two difficulties. Firstly the convex model is not nearest neighbor, and has an infinite-range with exponential tail. Secondly the two-point function of the Villain model is mapped via the duality transform to a non-linear and non-local observable (see Proposition 3.1). Understanding the behavior of this observable requires a precise, quantitative theory to describe the large-scale behavior of the convex gradient model.

This first part of the proof thus consists of applying a duality transformation and cluster expansion to relate the Villain model to a uniformly convex $\nabla\phi$ model. It is the subject of Section 3 and mostly follows [50] and [16, Section 5]. The second part of the proof consists of studying quantitatively the large-scale behavior of the convex gradient Gibbs model and treating the non-linear, non-local observable arising from the arguments of [50] and [16, Section 5], and is the subject of the remaining sections.

One of the main tools to study $\nabla\phi$ model is the so-called Helffer-Sjöstrand equation, originally introduced by Helffer and Sjöstrand [67], Naddaf and Spencer [81] and Giacomin, Olla and Spohn [54] to identify the scaling limit of the model. The main insight of [67, 81] is that the large-scale behavior of the $\nabla\phi$ model is related to the large-scale behavior of the solutions of an infinite-dimensional elliptic equation called the Helffer-Sjöstrand equation. The crucial observation of [81] is that the large-scale behavior of these solutions can be studied using techniques of homogenization.

At a high level, the proof of Theorem 1 consists of developing a quantitative homogenization theory for the Helffer-Sjöstrand equation and exploits the insights of the following three works: the work of Naddaf and Spencer [81], that relates large-scale behavior of the convex gradient Gibbs measure to an elliptic homogenization problem for the Helffer-Sjöstrand equation; the quantitative theory for homogenization by Armstrong, Kuusi and Mourrat [8, 7]; and the application of quantitative homogenization to the $\nabla\phi$ model by Armstrong and Wu [9]. However there is a distinct difference of our method compared to [81, 8, 9]. Firstly, the results of [81] are qualitative, and a quantitative theory is required to understand the behavior of the Villain model. To obtain a quantitative rate of homogenization it is crucial to have some decorrelation of the underlying random field. In [8], a straightforward mixing condition of the coefficient field is assumed. The argument in [9] relies on couplings based on the probabilistic interpretation of the equation to obtain decorrelation of the gradient field. In the present paper, we rely on the observation that this information can be obtained by studying another infinite-dimensional equation, the *second-order Helffer-Sjöstrand equation* (see [29, (2.12)] or Section 1.2.4); in particular, the decorrelation is a consequence of the decay estimates for the Green's function associated with the second-order Helffer-Sjöstrand operator. We note that the second-order equation appears in the work [29], and is closely related to techniques used to develop a quantitative theory of stochastic homogenization in [60, 61, 58, 59].

The following subsections provide a more detailed outline of the argument.

1.2.1. *Sine-Gordon representation and polymer expansion.* The spin wave computation (1.8) is only heuristic and does not give the correct constants C_1, C_2. The main problem for the spin wave heuristics (1.8) is that it ignores the formation of *vortices*, which are defined on the faces of $\mathbb{Z}^d$. Kosterlitz and Thouless [71] gave a heuristic argument, indicating that the vortices interact like a neutral Coulomb gas taking integer-valued charges.

Our proof of Theorem 1 starts from an insight of Fröhlich and Spencer [50], which makes this observation rigorous. In particular, the correlation function of the Villain model in $\mathbb{Z}^d$, $d \geq 3$ can be mapped, by duality, to a statistical mechanical model with integer-valued and locally neutral charges on discrete 2-forms $\Lambda^2(\mathbb{Z}^d)$, interacting with Coulomb potential (see Section 3.1). By performing a Fourier transform of this Gibbs measure with respect to the charge variable, we obtain a helpful random field representation of the Coulomb gas, known as the sine-Gordon representation (see e.g. [48, 49]). When the temperature is low enough, opposite charges tend to bind together into neutral (short range) dipoles, therefore on large scales this Coulomb gas behaves like an effective dipole gas with a reduced effective activity of the charges. This can be formalized by applying a one-step renormalization argument and a cluster expansion, following the presentation of [16, Chapter 5]. The renormalized Gibbs measure (see (3.16)) is a vector-valued random interface model in $\Lambda^2(\mathbb{Z}^d)$ with infinite range and uniformly convex potential. The question of the asymptotic behavior of the Villain correlation function is thus reduced to the question of the quantitative understanding of the large-scale properties of the random interface model.

1.2.2. *Random surfaces and Helffer-Sjöstrand equation.* Our study of the large-scale properties of the random interface model starts from the insight of Naddaf and Spencer [81] that the fluctuations of the field are closely related to an elliptic homogenization problem for the Helffer-Sjöstrand equation [67, 88]. This approach has been used by Giacomin, Olla and Spohn in [54] to prove that the large-scale space-time fluctuations of the field is described by an infinite-dimensional Ornstein-Uhlenbeck process and by Deuschel, Giacomin and Ioffe to establish concentration properties and large deviation principles on the random surface (we also refer to [87, 21, 22, 31, 30] for extension of these results to some non-convex potentials, and [72] for a study of a more general class of Hamiltonians). The strategy presented in many of the aforementioned articles relies on a probabilistic approach: one can, through the Helffer-Sjöstrand representation, reduce the problem to a question of random walk in dynamic random environment, and then prove properties on this object, e.g. invariance principles, using the results Kipnis and Varadhan [70], or annealed upper bounds on the heat kernel, using Delmotte and Deuschel [41]. However, the results obtained so far using this probabilistic approach are not quantitative. A more analytical approach was developed by Armstrong and Wu in [9], where they extend and quantify the homogenization argument of Naddaf and Spencer [81], resolved an open question posed by Funaki and Spohn [52] regarding the C^2 regularity of surface tension, and the fluctuation-dissipation conjecture of [54].

Besides the approach based on the Helffer-Sjöstrand equation and the random walk representation, various techniques have been successfully used on the model. Funaki and Spohn [52] established the hydrodynamic limit of the model relying on methods developed in the setting of the Ginzburg-Landau equation with a conserved order parameter [64]. A renormalization group approach has been implemented in the works of Adams, Kotecký, Müller [3] and Adams, Buchholz, Kotecký, Müller [2]. In these contributions, the authors study the $\nabla\phi$ model for a general class of (perturbative) *non-convex potentials* (in a low temperature regime) and establish (among other results) regularity properties as well as the strict convexity of the surface tension of the model. The articles [3, 2] differ from ours in various aspects. In [3, 2], the authors consider a nonconvex perturbation of Gaussian, and proved after successive renormalizations the surface tension (i.e., the log partition function under different tilts) gains sufficient regularity and convexity. In the present article, the gradient-type model obtained from the Villain model by duality is uniformly convex, and the main difficulty relies on the specific structure of the model: the Hamiltonian has infinite-range, the observable we wish to study is highly non-linear and non-local. Therefore it is not enough to prove the Gibbs measure converges to a Gaussian free field in the scaling limit, and we need to estimate the correlation of nonlinear functions of the field with high precision, which we do by implementing methods from PDE and homogenization theory.

On a high level, we follow the analytical approach, namely the program developed in [81, 9] on homogenization for the random interface models. Since the sine-Gordon representation and the polymer expansion give a random interface model valued in the vector space $\mathbb{R}^{\binom{d}{2}}$ with long range and uniformly convex potential, an application of the strategy of Naddaf and Spencer [81] to this model leads to the Helffer-Sjöstrand operator

$$\mathcal{L} := -\Delta_\phi + \mathcal{L}_{\mathrm{spat}}, \tag{1.11}$$

which is an infinite-dimensional elliptic operator acting on functions defined in the space $\Omega \times \mathbb{Z}^d$ where Ω is the set of functions $\phi : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}}$ (see (3.57) for the precise definition of this operator), where Ω is the space of functions from $\mathbb{Z}^d$ to $\mathbb{R}^{\binom{d}{2}}$ in which the vector-valued random interface considered in this article takes its values. The operator Δ_ϕ is the (infinite-dimensional) Laplacian computing derivatives with respect to the height of the random surface and $\mathcal{L}$ is an operator associated with a uniformly elliptic *system of equations* with infinite range (and with exponential decay on the size of the long range coefficients) on the discrete lattice $\mathbb{Z}^d$. The analysis of these systems requires to overcome some difficulties; a number of properties which are valid for elliptic equations, and used to study the random interface models, are known to be false for elliptic systems. It is for instance the case for the maximum principle, which is used to obtain a random walk representation, the De Giorgi-Nash-Moser regularity theory for uniformly elliptic and parabolic PDE (see [83, 39], [56, Section 8] and the counterexample of De Giorgi [40]) and the Nash-Aronson estimate on the heat kernel (see [11]).

To resolve this lack of regularity, we rely on a perturbative argument, and make use of ideas from *Schauder theory* (see [66, Section 3]), as well as the ones from *the large-scale regularity in homogenization* (see Avellaneda, Lin [12, 13] and Armstrong, Smart [10]); we leverage on the fact that the inverse temperature β is chosen very large so that the elliptic operator $\mathcal{L}$ can be written

$$\mathcal{L}_{\mathrm{spat}} := -\frac{1}{2\beta}\Delta + \mathcal{L}_{\mathrm{pert}},$$

where the operator $\mathcal{L}_{\mathrm{pert}}$ is a perturbative term; its typical size is of order $\beta^{-\frac{3}{2}} \ll \beta^{-1}$. One can thus prove that any solution u of the equation (1.11) is well-approximated on every scale by a solution $\overline{u}$ of the equation $-\Delta_\phi - \frac{1}{2\beta}\Delta$ for which the regularity can be easily established. It is then possible to borrow the strong regularity properties of the function $\overline{u}$ and transfer it to the solution of (1.11). This strategy is implemented in Section 5 and allows us to prove the $C^{0,1-\varepsilon}$-regularity of the solution of the Helffer-Sjöstrand equation, and to deduce from this regularity property various estimates on other quantities of interest (e.g, decay estimates on the heat kernel in dynamic random environment, decay and regularity for the Green's matrix associated with the Helffer-Sjöstrand operator). The regularity exponent ε depends on the dimension d and the inverse temperature β, and tends to 0 as β tends to infinity; in the perturbative regime, the result turns out to be much stronger than the $C^{0,\alpha}$-regularity provided by the De Giorgi-Nash-Moser theory (for some tiny exponent $\alpha > 0$) in the case of elliptic equations, and allows to quantify (precisely) the mixing properties of the random field.

1.2.3. *Stochastic homogenization.* The main difficulty to establish Theorem 1 is that since the Villain model is not exactly solvable, the dependence of the constants c_1 and c_2 on the dimension d and the inverse temperature β is highly non explicit; one does not expect to have a simple formula for these coefficients. However, it is necessary to analyze them in order to prove the expansions (1.9) and (1.10); this is achieved by using tools from the quantitative theory of stochastic homogenization.

This theory is typically interested in the understanding of the large-scale behavior of the solutions of the elliptic equation

$$-\nabla \cdot \mathbf{a}(x)\nabla u = 0 \text{ in } \mathbb{R}^d, \tag{1.12}$$

where $\mathbf{a}$ is a random, uniformly elliptic coefficient field that is stationary and ergodic. The general objective is to prove that, on large scales, the solutions of (1.12) behave like the solutions of the elliptic equation

$$-\nabla \cdot \bar{\mathbf{a}}\nabla u = 0 \text{ in } \mathbb{R}^d, \tag{1.13}$$

where $\bar{\mathbf{a}}$ is a constant uniformly elliptic coefficient called *the homogenized matrix.* The theory was initially developed in the 80's, in the works of Kozlov [73], Papanicolaou and Varadhan [85], and Yurinskiĭ [91]. Dal Maso and Modica [32, 33] extended these results a few years later to non-linear equations using variational arguments inspired by Γ-convergence. All of these results rely on the ergodic theorem, and are therefore purely qualitative.

The main difficulty in the establishment of a quantitative theory is to transfer the quantitative ergodicity encoded in the coefficient field $\mathbf{a}$ to the solutions of the equation. This problem was addressed in a satisfactory fashion for the first time by Gloria and Otto in [60, 61], where, building upon the ideas of [82], they used spectral gap inequalities (or concentration inequalities) to transfer the quantitative ergodicity of the coefficient field to the solutions of (1.12). These results were then further developed in [62, 63, 58, 59].

Another approach, which is the one pursued in this article, was initiated by Armstrong and Smart in [10], who extended the techniques of Avellaneda and Lin [12, 13], the ones of Dal Maso and Modica [32, 33] and obtained an algebraic, suboptimal rate of convergence for the homogenization error of the Dirichlet problem associated with the non-linear version of the equation (1.12). These results were then improved in [6, 7, 8] to obtain optimal rates. Their approach relies on mixing conditions on the coefficient fields and on the quantification of the closeness of dual monotone quantities (see Section 6). An extension of the techniques of [8] to the setting of differential forms (which also appear in this article in the dual Villain model) can be found in [35], and to the uniformly convex gradient field model in [34]. In [80], Mourrat and Otto study the correlation structure of the corrector and prove that it is similar, in the large-scale limit, to the one of a variant of a Gaussian free field. Their strategy shares some similarities with ours: under some suitable assumptions on the coefficient field, they use a Helffer-Sjöstrand representation formula to study the correlation of the corrector, and reduce the problem to the question of the quantitative homogenization of the Green's function associated with the heterogeneous operator (1.12).

To prove Theorem 1, we apply the techniques of [8] to the Helffer-Sjöstrand equation to prove the quantitative homogenization of the mixed derivative of the Green's matrix associated with this operator. The strategy can be decomposed into two steps.

The first one relies on the variational structure of the Helffer-Sjöstrand operator and is the main subject of Section 6: following the arguments of [8, Section 2], we define two subadditive quantities, denoted by ν and ν^*. The first one corresponds to the energy of the Dirichlet problem associated with the Helffer-Sjöstrand operator (1.11) in a domain $U \subseteq \mathbb{Z}^d$ and subject to affine boundary condition, the second one corresponds to the energy of the Neumann problem of the same operator with an affine flux. Each of these two quantities depends on two parameters: the domain of integration U and the slope of the affine boundary condition, denoted by p (for ν) and p^* (for ν^*). These energies are quadratic, uniformly convex with respect to the variables p and p^*, and are approximately convex dual to one another. They additionally satisfy a subadditivity property with respect to the domain U, and one can show that they converge as the size of the domain tends to infinity to a pair of quadratic, convex dual functions, i.e., there exists a positive definite matrix $\bar{\mathbf{a}}$ such that

$$\nu(U,p) \underset{|U|\to\infty}{\longrightarrow} \frac{1}{2} p \cdot \bar{\mathbf{a}} p \quad \text{and} \quad \nu^*(U,p^*) \underset{|U|\to\infty}{\longrightarrow} \frac{1}{2} p^* \cdot \bar{\mathbf{a}}^{-1} p^*.$$

The matrix $\bar{\mathbf{a}}$ plays a similar role as the homogenized matrix in (1.13); in the case of the present random interface model, it gives the covariance matrix of the continuous (homogenized) Gaussian free field which describes the large-scale behavior of the random surface as established in [81]. The objective of the proofs of Section 6 is to quantify this convergence and to obtain an algebraic rate: we show that, for large β, there exists an exponent $\alpha > 0$ depending only on the dimension d such that for any cube $\square \subseteq \mathbb{Z}^d$ of size $R > 0$,

$$\left| \nu(\square, p) - \frac{1}{2} p \cdot \bar{\mathbf{a}} p \right| + \left| \nu(\square, p^*) - \frac{1}{2} p^* \cdot \bar{\mathbf{a}}^{-1} p^* \right| \leq C R^{-\alpha}. \tag{1.14}$$

The strategy to prove the quantitative rate (1.14) relies on the approximate convex duality of the maps $p \mapsto \nu(U,p)$ and $p^* \mapsto \nu^*(U,p^*)$. Following [8], we use a multiscale argument to prove that, as one passes to a larger scale, the *convex duality defect*

$$p \mapsto \inf_{p^* \in \mathbb{R}^d} \left[\nu(\square, p) + \nu^*(\square, p^*) - p \cdot p^* \right],$$

must contract by a multiplicative factor strictly smaller than 1, and thus it is equal to 0 in the infinite volume limit. More precisely we show that the convex duality defect can be controlled by the subadditivity defect, and then iterate the result over all the scales from 1 to R to obtain (1.14) (see Section 6.1.3). As a byproduct of the proof, we obtain a quantitative control on the sublinearity of the finite-volume corrector defined as the solution of the Dirichlet problem: given an affine function l_p of slope p, and a cube $\square_R := [-R,R]^d \cap \mathbb{Z}^d$ of size R,

$$\begin{cases} \mathcal{L}(l_p + \chi_{R,p}) = 0 \text{ in } \square_R \times \Omega, \\ \qquad\quad \chi_{R,p} = 0 \text{ on } \partial \square_R \times \Omega. \end{cases}$$

This estimate takes the following form

$$\|\chi_{R,p}\|_{\underline{L}^2(\square_R, \mu_\beta)} \leq \frac{C}{R^{1-\alpha}}, \tag{1.15}$$

where the average L^2-norm is considered over both the spatial variable and the random field (see (2.5)).

The second step in the argument, which extends the results of [9], is to prove quantitative homogenization of the mixed derivative of the Green's matrix associated with the Helffer-Sjöstrand operator (1.11); it is the subject of Section 7. In the setting of the divergence form elliptic operator (1.12), the properties of the Green's function are well-understood: moment bounds on the Green's function, its gradient and mixed derivative are proved in [41, 19, 28], and quantitative homogenization estimates are proved in [8, Sections 8 and 9] and in [20]. The argument used here relies on a common strategy in stochastic homogenization: the two-scale expansion. It is implemented as follows: the large-scale behavior of the fundamental solution $\mathcal{G} : \Omega \times \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2} \times \binom{d}{2}}$ of the elliptic system

$$\mathcal{L}\mathcal{G} = \delta_0 \text{ in } \mathbb{Z}^d \times \Omega,$$

is described by the (deterministic) fundamental solution $\overline{G} : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}\times\binom{d}{2}}$ of the homogenized elliptic system

$$-\nabla \cdot \bar{\mathbf{a}} \nabla \overline{G} = \delta_0 \text{ in } \mathbb{Z}^d.$$

The proof of this result relies on a two-scale expansion for systems of equations: we select a suitable cube $\square \subseteq \mathbb{R}^d$ and define the function, for any $k \in \{1, \dots, \binom{d}{2}\}$

$$\mathcal{H}_{\cdot k} := \overline{G}_{\cdot k} + \sum_{i=1}^{d} \sum_{j=1}^{\binom{d}{2}} \chi_{R, e_{ij}} \nabla_i \overline{G}_{jk}.$$

We then compute the value of $\mathcal{L}\mathcal{H}$ and prove, by using the quantitative information obtained on the corrector (1.15), that this value is small in a suitable functional space. This argument shows that the function $\mathcal{H}$ (resp. its gradient) is quantitatively close to the functions $\mathcal{G}$ (resp. its gradient). Once this is achieved, we can iterate the argument to obtain a quantitative homogenization result for the mixed derivative of the Green's matrix. The overall strategy is similar to the one in the case of the divergence form elliptic equations (1.12) but a number of technicalities need to be treated along the way pertaining to either the Witten Laplacian Δ_ϕ (this difficulty has been successfully addressed in [9]), and the infinite range of the elliptic operator $\mathcal{L}_{\mathrm{spat}}$ (using the exponential decay of the interaction is enough to adapt the arguments developed in the nearest-neighbor setting).

1.2.4. *Second-order Helffer-Sjöstrand equation.* As we mentioned, the method pursued in this paper differs from [8] and [9] and is based on the regularity theory of the second-order Helffer-Sjöstrand equation. We note that, contrary to the case of the homogenization of the elliptic equation (1.12), the subadditive quantities are deterministic objects and are applied to the operator (1.11) which is essentially infinite-dimensional. To quantify the subadditive ergodic theorem and obtain the rate of convergence (1.14), it is crucial that the random fields $\nabla\phi$ that appears in the definition of ν and ν^* decorrelates (see Definition 6.4). While the proofs of quantitative rate of convergence in [8, Section 2] rely on a finite range dependence assumption of the coefficient field, we rely here on the regularity properties of the Helffer-Sjöstrand operator to prove sufficient decorrelation estimates on the field. The same issues were addressed in the work of Armstrong and Wu [9], to study the $\nabla\phi$ model and prove C^2-regularity of the surface tension conjectured by Funaki and Spohn [52]; the arguments presented there are different as they rely on couplings based on the probabilistic interpretation of the equation to obtain sufficient decorrelation of the discrete gradient $\nabla\phi$. In the present paper, we rely on the observation of Conlon and Spencer [29] that if u is a solution to the Helffer-Sjöstrand equation (1.11), then the derivative of the function u with respect to the field ϕ, i.e., the map $v : (x, \phi, y) \mapsto \partial_y u(x, \phi)$, for $x, y \in \mathbb{Z}^d$ and $\phi \in \Omega$, solves a second-order Helffer-Sjöstrand equation of the form

$$\Delta_\phi v(x, y, \phi) + \mathcal{L}_{\mathrm{spat},x} v(x, y, \phi) + \mathcal{L}_{\mathrm{spat},y} v(x, y, \phi) + (\partial_y \mathcal{L})\, v = 0 \text{ in } \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega. \tag{1.16}$$

We refer to Section 5.4 for a precise definition. This operator is then used in [29] to obtain uniform third moment bounds for the $\nabla\phi$ Gibbs measure. We note that this strategy is very similar to the one developed in stochastic homogenization in [60, 61, 58, 59]. In this paper we exploit more precise information of the operator, and apply the $C^{0,1-\varepsilon}$ regularity theory to obtain decay estimates on the Green's function associated with (1.16). In particular, we obtain the regularity theory for the second-order Helffer-Sjöstrand operator for large β, namely, the off-diagonal decay of the associated Green's matrix, its gradient, and its mixed derivative (see Corollary 5.13). These properties can be used to quantify the ergodicity of the Helffer-Sjöstrand equation and obtain the quantitative rate of convergence (1.14).

The second-order Helffer-Sjöstrand equation also plays a crucial role to derive Theorem 1 from the homogenization results. Applying the duality, we map the two-point function of the Villain model to a non-local observable (see Proposition 3.1). This non-local observable is then analyzed by repeated applications of the Helffer-Sjöstrand representation to single out the main contribution (thus the second-order Helffer-Sjöstrand operator emerges), and the $C^{0,1-\varepsilon}$ regularity theory is crucially applied to control the remainder terms (see Section 4.4 and Section 4.5 for the details).

1.2.5. *First order expansion of the two-point functions.* The first order expansion of the two-point function stated in Theorem 1 is obtained by post-processing all the arguments above. We first use the sine-Gordon representation and the polymer expansion to reduce the question to the understanding of the large scale behavior of a vector-valued random surface model, whose Hamiltonian is a perturbation of the one of a Gaussian free field, and use the properties of the Helffer-Sjöstrand equation to treat the problem. The proof of Theorem 1 is decomposed into three parts:

- We establish a $C^{0,1-\varepsilon}$-regularity theory for the solutions of the Helffer-Sjöstrand and second-order Helffer-Sjöstrand operators by using the techniques of Schauder regularity (through a perturbative argument) in order to obtain a precise understanding of the correlation structure of the random field, this is done in Section 5;
- We prove a quantitative homogenization theorem for the mixed derivative associated with the Helffer-Sjöstrand operator (Theorem 2), this is done in Sections 6 and 7;
- We post-process the results of the two arguments above to prove Theorem 1. The proof relies on the study of the non-local observable introduced in Proposition 3.1; it requires to analyze a number of terms, to isolate the leading order terms, and to estimate quantitatively the lower order ones. It is rather technical and is split into two sections: in Section 4, we present a detailed sketch of the argument, isolate the leading order from the lower order terms, and state the estimates on each of these terms. Section 8 is devoted to the proof of the technical estimates.

1.3. **Organization of the paper.** This article is the short version of the v1 of arxiv preprint [36], which contains in addition some detailed but standard computations which are recalled here without a proof. In the next section, we introduce some preliminary notation and results. In Section 3, we recall the dual formulation of the Villain model in terms of a vector-valued random interface model, based on the ideas of Fröhlich and Spencer [50] and following the presentation of Bauerschmidt [16]. We then derive the Helffer-Sjöstrand equation for the renormalized measure and state the main regularity estimates on the Green's matrix proved in Section 5, and the quantitative homogenization of the mixed derivative of the Green's matrix proved in Sections 6 and 7. In Section 4, we sketch the proof of the main theorem, assuming the $C^{0,1-\varepsilon}$ regularity for the solutions of the Helffer-Sjöstrand equation (established in Section 5), and the quantitative homogenization of the mixed derivative of the Green's matrix (established in Sections 7 and 8). Finally in Section 8, we give detailed proofs of the claims in Section 4.

Acknowledgments. P.D. is supported by the Israel Science Foundation grants 861/15 and 1971/19 and by the European Research Council starting grant 678520 (LocalOrder). W.W. is supported in part by the EPSRC grant EP/T00472X/1. We thank T. Spencer for many insightful discussions that inspired the project, R. Bauerschmidt for kindly explaining the arguments in [16], and S. Armstrong for many helpful discussions. We also thank S. Armstrong and J.-C. Mourrat for helpful feedbacks on a previous version of the paper.

2. Preliminaries

2.1. **Notation and assumptions.**

2.1.1. *General notation.* We work on the Euclidean lattice $\mathbb{Z}^d$ in dimension $d \geq 3$, and denote by $|\cdot|$ the Euclidean norm on the lattice $\mathbb{Z}^d$. We say that two points $x, y \in \mathbb{Z}^d$ are neighbors, and denote it by $x \sim y$, if $|x - y| = 1$. We denote by $e_1, \ldots, e_k$ the canonical basis of $\mathbb{R}^d$. Given a subset $U \subseteq \mathbb{Z}^d$, we define its interior U° and its inner boundary ∂U by the formulae

$$U^\circ := \{x \in U : x \sim y \implies y \in U\} \quad \text{and} \quad \partial U := U \setminus U^\circ.$$

If the subset $U \subseteq \mathbb{Z}^d$ is finite, we denote by $|U|$ its cardinality and refer to this quantity as the volume of U. We denote by $\operatorname{diam} U$ the diameter of U defined by the formula $\operatorname{diam} U := \sup_{x,y \in U} |x - y|$. Given a point $x \in \mathbb{Z}^d$ and a radius $r > 0$, we denote by $B(x, r)$ the discrete Euclidean ball of center x and radius r. We frequently use the notation B_r to mean $B(0, r)$. We also define the annulus $A_R := B_{2R} \setminus B_R$.

A discrete cube $\square$ of $\mathbb{Z}^d$ is a subset of the form

$$\square := x + [-N, N]^d \cap \mathbb{Z}^d \text{ with } x \in \mathbb{Z}^d \text{ and } N \in \mathbb{N}. \tag{2.1}$$

We refer to the point x as the center of the cube $\square$, and to the integer $2N + 1$ as its length. For $L \in \mathbb{N}$, we also denote by $\square_L := [-N, N]^d \cap \mathbb{Z}^d$.

Given three real numbers $X, Y \in \mathbb{R}$ and $\kappa \in [0, \infty)$, we write

$$X = Y + O(\kappa) \quad \text{if and only if} \quad |X - Y| \leq \kappa.$$

We frequently consider functions defined from $\mathbb{Z}^d$ and valued in $\mathbb{R}$ of the form $x \to |x|^{-k}$. We implicitly extend these functions at the point $x = 0$ by the value 1 so that they are defined on the entire lattice $\mathbb{Z}^d$.

2.1.2. *Notation for vector-valued functions.* For each integer $k \in \mathbb{N}$, we let $\mathcal{F}\left(\mathbb{Z}^d, \mathbb{R}^k\right)$ be the set of functions defined on $\mathbb{Z}^d$ and taking values in $\mathbb{R}^k$. Given a function $g \in \mathcal{F}\left(\mathbb{Z}^d, \mathbb{R}^k\right)$, we denote by $g_1, \ldots, g_k$ its components on the canonical basis of $\mathbb{R}^k$ and write $g = (g_1, \ldots, g_k)$. We define the support of the function g to be the set

$$\operatorname{supp} g := \left\{x \in \mathbb{Z}^d : g(x) \neq 0\right\}.$$

For each integer $i \in \{1, \ldots, d\}$, we define its discrete i-th derivative $\nabla_i g : \mathbb{Z}^d \to \mathbb{R}^k$ and its adjoint $\nabla_i^* g : \mathbb{Z}^d \to \mathbb{R}^k$ by the formulae, for each $x \in \mathbb{Z}^d$,

$$\nabla_i g(x) := g(x + e_i) - g(x) \quad \text{and} \quad \nabla_i^* g(x) := g(x) - g(x - e_i).$$

The discrete gradient $\nabla g : \mathbb{Z}^d \to \mathbb{R}^{d \times k}$ is then defined by

$$\nabla g(x) = (\nabla_i g_j(x))_{1 \leq i \leq d, 1 \leq j \leq k} \quad \text{and} \quad \nabla^* g(x) = (\nabla_i^* g_j(x))_{1 \leq i \leq d, 1 \leq j \leq k}. \tag{2.2}$$

We define similarly the divergence, for any function $g : \mathbb{Z}^d \to \mathbb{R}^d$,

$$\nabla \cdot g(x) = \sum_{i=1}^d g_i(x) - g_i(x - e_i),$$

We extend this definition to a more general class of vector-valued functions as follows: for an integer $k \in \mathbb{N}$, and a function $g = (g_{ij})_{1 \leq i \leq d, 1 \leq j \leq k} : \mathbb{Z}^d \to \mathbb{R}^{d \times k}$, we define $\nabla \cdot g : \mathbb{Z}^d \to \mathbb{R}^k$ by the identity

$$\nabla \cdot g(x) = \left(\sum_{i=1}^d g_{i,1}(x) - g_{i,1}(x - e_i), \ldots, \sum_{i=1}^d g_{i,k}(x) - g_{i,k}(x - e_i)\right).$$

The Laplacian is then defined by the identity $\Delta = \nabla \cdot \nabla$ and is equivalently given by the explicit formula: for any $g : \mathbb{Z}^d \to \mathbb{R}^k$,

$$\Delta g(x) = \sum_{y \sim x} (g(y) - g(x)). \tag{2.3}$$

Given two functions $f, g : \mathbb{Z}^d \to \mathbb{R}^k$ and a point $x \in \mathbb{Z}^d$, we define the scalar product $f(x) \cdot g(x) := \sum_{i=1}^d f_i(x) g_i(x)$. To ease the notation, we may write $f(x) g(x)$ to mean $f(x) \cdot g(x)$. Given a finite subset $U \subseteq \mathbb{Z}^d$, we define the L^2-scalar products $(\cdot, \cdot)$ and $(\cdot, \cdot)_U$ according to the formulae

$$(f, g) = \sum_{x \in \mathbb{Z}^d} f(x) g(x) \quad \text{and} \quad (f, g)_U := \sum_{x \in U} f(x) g(x). \tag{2.4}$$

For each subset $U \subseteq \mathbb{Z}^d$, we define the $L^2(U)$-norm

$$\|g\|_{L^2(U)} := \left(\sum_{x \in U} |g(x)|^2 \right)^{\frac{1}{2}}.$$

where the notation $|\cdot|$ refers to the Euclidean norm on $\mathbb{R}^k$. Given a bounded subset $U \subseteq \mathbb{Z}^d$, we denote by $\underline{L}^2(U)$ the normalized norm

$$\|g\|_{\underline{L}^2(U)} := \left(\frac{1}{|U|} \sum_{x \in \mathbb{Z}^d} |g(x)|^2 \right)^{\frac{1}{2}}.$$

We introduce the normalized Sobolev norms $\underline{H}^1(U)$ by the formula

$$\|g\|_{\underline{H}^1(U)} := \frac{1}{\operatorname{diam} U} \|g\|_{\underline{L}^2(U)} + \|\nabla g\|_{\underline{L}^2(U)}.$$

We denote by $H^1_0(U)$ the set of functions from U to $\mathbb{R}^k$ which are equal to 0 outside the set U (by analogy to the Sobolev space).

2.1.3. *Notation for Gibbs measures.* We let Ω be the set of vector-valued functions $\phi : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}}$. We then introduce the set of smooth local and compactly supported functions of the set Ω

$$C^\infty_c(\Omega) := \big\{ F : \Omega \to \mathbb{R} \, : \, \exists \, n \in \mathbb{N}, \exists \, x_1, \ldots, x_n \in \mathbb{Z}^d \text{ and } \exists \, f \in C^\infty_c(\mathbb{R}^n) \\ \text{such that } F(\phi) = f(\phi(x_1), \ldots, \phi(x_n)) \big\}.$$

For $k \in \mathbb{N}$, we extend the previous notation to vector-valued functions $F : \Omega \to \mathbb{R}^k$ and write $F \in C^\infty_c(\Omega)$ if all the components of F belong to $C^\infty_c(\Omega)$ (i.e., if $F = (F_1, \ldots, F_k)$ and for any $i \in \{1, \ldots, k\}$, $F_i \in C^\infty_c(\Omega)$).

Given a probability measure μ on Ω and measurable function $X : \Omega \to \mathbb{R}$ which is integrable with respect to the measure μ, we denote by $\langle X \rangle_\mu$ and $\operatorname{var}_\mu[X]$ its expectation and variance respectively. As before, we extend the notation to vector-valued functions by writing $\langle X \rangle_\mu = \left(\langle X_1 \rangle_\mu, \ldots, \langle X_k \rangle_\mu \right) \in \mathbb{R}^k$ and $\operatorname{var}_\mu[X] = \sum_{i=1}^k \operatorname{var}_\mu[X_i]$ for $X = (X_1, \ldots, X_k)$.

Fix $u : \Omega \to \mathbb{R}^k$. For $x \in \mathbb{Z}^d$ and each integer $i \in \left\{1, \ldots, \binom{d}{2}\right\}$, we define the differential operators $\partial_{x,i}$ and ∂_x by the formulae

$$\partial_{x,i} u(\phi) := \lim_{h \to 0} \frac{u(\phi + h e_i \mathbb{1}_x) - u(\phi)}{h} \in \mathbb{R}^k \quad \text{and} \quad \partial_x u(\phi) = \left(\partial_{x,1} u, \ldots, \partial_{x, \binom{d}{2}} u \right) \in \mathbb{R}^{\binom{d}{2} \times k},$$

where $\left(e_1, \ldots, e_{\binom{d}{2}} \right)$ is the canonical basis of $\mathbb{R}^{\binom{d}{2}}$. We define the space $H^1(\mu)$ to be the closure of the space $C^\infty_{\mathrm{loc}}(\Omega)$ with respect to the norm

$$\|u\|_{H^1(\mu)} := \|u\|_{L^2(\mu)} + \left(\sum_{x \in \mathbb{Z}^d} \|\partial_x u\|^2_{L^2(\mu)} \right)^{\frac{1}{2}}.$$

For any subset $U \subseteq \mathbb{Z}^d$, we define the $L^2(U, \mu)$ to be the set of functions $u : U \times \Omega \to \mathbb{R}$ such that

$$\|u\|_{L^2(U,\mu)} := \left(\sum_{x \in U} \|u(x, \cdot)\|^2_{L^2(\mu)} \right)^{\frac{1}{2}}$$

If $U \subseteq \mathbb{Z}^d$ is finite, we additionally define

$$\|u\|_{\underline{L}^2(U,\mu)} := \left(\frac{1}{|U|} \sum_{x \in U} \|u(x, \cdot)\|^2_{L^2(\mu)} \right)^{\frac{1}{2}}. \tag{2.5}$$

We similarly define the $H^1(U, \mu)$-norms by the formulae

$$\|u\|_{H^1(U,\mu)} := \left(\sum_{x \in U} \|u(x, \cdot)\|^2_{H^1(\mu)} + \|\nabla u\|^2_{L^2(U,\mu)} \right)^{\frac{1}{2}} \tag{2.6}$$

as well as

$$\|u\|_{\dot{H}^1(U,\mu)} := \left(\sum_{x \in \mathbb{Z}^d} \sum_{y \in U} \|\partial_x u(y, \cdot)\|^2_{L^2(\mu)} + \|\nabla u\|^2_{L^2(U,\mu)} \right)^{\frac{1}{2}}. \tag{2.7}$$

2.1.4. *Discrete differential forms.* For each integer $k \in \{1, \ldots, d\}$, a k-cell of the lattice $\mathbb{Z}^d$ is a set of the form, for a subset $\{i_1, \ldots, i_k\} \subseteq \{1, \ldots, d\}$, and a point $x \in \mathbb{Z}^d$,

$$\left\{ x + \sum_{l=1}^{k} \lambda_l e_{i_l} \in \mathbb{R}^d \, : \, 0 \leq \lambda_1, \ldots, \lambda_k \leq 1 \right\}.$$

We equip the set of k-cells with an orientation induced by the canonical orientation of the lattice $\mathbb{Z}^d$ and denote by $\Lambda^k(\mathbb{Z}^d)$ the set of oriented k-cells of the lattice $\mathbb{Z}^d$. Given a k-cell c_k, we denote by ∂c_k the boundary of the cell; it can be decomposed into a disjoint union of $(k-1)$-cells. The values $k = 0, 1, 2$ are of specific interest to us; they correspond to the set of vertices, edges and faces of the lattice $\mathbb{Z}^d$. We will denote these spaces by $V(\mathbb{Z}^d)$, $E(\mathbb{Z}^d)$ and $F(\mathbb{Z}^d)$ respectively. Given a box $\square \subseteq \mathbb{Z}^d$, we denote by $\Lambda^k(\square)$ the set of oriented k-cells which are included in the cube $\square$, and by $V(\square)$, $E(\square)$ and $F(\square)$ the set of vertices, edges and faces of the cube $\square$ respectively.

For each k-cell c_k, we denote by c_k^{-1} the same k-cell as c_k with reverse orientation and by ∂c_k the boundary this cell. A k-form u is a mapping from $\Lambda^k(\square)$ to $\mathbb{R}$ such that $u\left(c_k^{-1}\right) = -u\left(c_k\right)$.

Given a k-form u, we define its exterior derivative $\mathrm{d}u$ according to the formula, for each oriented $(k+1)$-cell c_{k+1},

$$\mathrm{d}u\left(c_{k+1}\right) = \sum_{c_k \subseteq \partial c_{k+1}} u(c_k), \tag{2.8}$$

where the orientation of the face c_k is given by the orientation of the $(k+1)$-cell c_{k+1}; we set the convention $\mathrm{d}u = 0$ for any d-form u. We define the codifferential d^* according to the formula, for each $(k-1)$-cell c_{k-1} and each k-form $u : \Lambda^k(\square) \to \mathbb{R}$,

$$\mathrm{d}^* u\left(c_{k-1}\right) \coloneqq \sum_{\partial c_k \ni c_{k-1}} u(c_k). \tag{2.9}$$

Clearly, $\mathrm{d}u$ is a $(k+1)$-form and d^*u is a $(k-1)$-form; we set $\mathrm{d}^*u = 0$ for any 0-form u. One also verifies the properties, for each k-form $u : \Lambda^k(\square) \to \mathbb{R}$, $\mathrm{d}\mathrm{d}u = 0$ and $\mathrm{d}^*\mathrm{d}^*u = 0$. For arbitrary k-forms $u, v : \Lambda^k(\mathbb{Z}^d) \to \mathbb{R}$ with finite support, we define the scalar product $(\cdot, \cdot)$ by the formula

$$(u, v) = \sum_{c_k \in \Lambda^k(\mathbb{Z}^d)} u(c_k) v(c_k). \tag{2.10}$$

We may restrict the scalar product $(\cdot, \cdot)$ to forms which are only defined in a cube $\square$; we denote the corresponding scalar product by $(\cdot, \cdot)_\square$. It is defined by the formula, for each pair of k-forms $u, v : \Lambda^k(\square) \to \mathbb{R}$,

$$(u, v) = \sum_{c_k \in \Lambda^k(\square)} u(c_k) v(c_k).$$

The codifferential d^* is the formal adjoint of the exterior derivative d with respect to this scalar product: Given a k-form $u : \Lambda^k(\mathbb{Z}^d) \to \mathbb{R}$ and a $(k+1)$-form $v : \Lambda^{k+1}(\mathbb{Z}^d) \to \mathbb{R}$ with finite supports, one has the identity

$$(\mathrm{d}u, v) = (u, \mathrm{d}^* v). \tag{2.11}$$

For an integer $k \in \{0, \ldots, d-1\}$ and a cube $\square \subseteq \mathbb{Z}^d$, we define the tangential boundary of the cube $\partial_{k,\mathbf{t}}\square$ to be the set of all the k-cells which are included in the boundary of the cube $\square$. Given a k-form $u : \Lambda^k(\square) \to \mathbb{R}$, we define its tangential trace $\mathbf{t}u$ to be the restriction of the form u to the set $\partial_{k,\mathbf{t}}\square$. One has the formula, for each k-form $u : \Lambda^k(\square) \to \mathbb{R}$ such that $\mathbf{t}u = 0$ and each $(k+1)$-form $v : \Lambda^k(\square) \to \mathbb{R}$,

$$(\mathrm{d}u, v)_\square = (u, \mathrm{d}^* v)_\square.$$

2.1.5. *Differential forms as vector-valued functions.* Given a subset $I = (i_1, \ldots, i_k) \subseteq \{1, \ldots, d\}$ of cardinality k. We denote by $\Lambda^k_I(\mathbb{Z}^d)$ the set of oriented k-cells of the hypercubic lattice $\mathbb{Z}^d$ which are parallel to the vectors $(e_{i_1}, \ldots, e_{i_k})$. This set can be characterized as follows: if we let c_I be the k-cell defined by the formula

$$c_I \coloneqq \left\{ \sum_{l=1}^{k} \lambda_l e_{i_l} \in \mathbb{R}^d \, : \, 0 \leq \lambda_1, \ldots, \lambda_k \leq 1 \right\},$$

then we have

$$\Lambda^k_I(\mathbb{Z}^d) = \left\{ x + c_I \, : \, x \in \mathbb{Z}^d \right\}. \tag{2.12}$$

The identity (2.12) allows to identify the vector space of k-forms to the vector space of functions defined on $\mathbb{Z}^d$ and valued in $\mathbb{R}^{\binom{d}{k}}$ according the procedure described below. Note that there are $\binom{d}{k}$ subsets of $\{1, \ldots, d\}$ of cardinality k and consider an arbitrary enumeration $I_1, \ldots, I_{\binom{d}{k}}$ of these sets. To each k-form $\hat{u} : \Lambda^k(\mathbb{Z}^d) \to \mathbb{R}$, we can associate a vector-valued function $u : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{k}}$ defined by the formula, for each point $x \in \mathbb{Z}^d$,

$$u(x) = \left(\hat{u}\left(x + c_{I_1}\right), \ldots, \hat{u}\left(x + c_{I_{\binom{d}{k}}}\right) \right). \tag{2.13}$$

This identification is enforced in most of the article; in fact, except in Section 3.1, we always work with vector-valued functions instead of differential forms. We use the identification (2.13) to extend the formalism described in Section 2.1 to differential forms; we may for instance refer to the gradient of a form, or the Laplacian of a form etc. Reciprocally, we extend the formalism described in Section 2.1.4 to vector-valued functions; given a function $u : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{k}}$, we may refer to the exterior derivative, the codifferential, which we still denote by $\mathrm{d}u$, $\mathrm{d}^* u$ respectively. We note that the two definitions of the scalar products (2.4) for vector valued functions and (2.10) for differential forms coincide through the identification (2.13).

From the definition of the exterior derivative d and the codifferential d^* given in (2.8) and (2.9) and the identification (2.13), one sees that the differential operators d and d^* are linear functionals of the gradient ∇.

We record the following identity which relates the Laplacian Δ defined in (2.3) to the exterior derivative d and the codifferential d^*,

$$-\Delta = \mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}. \tag{2.14}$$

Using the identities $\mathrm{d} \circ \mathrm{d} = 0$ and $\mathrm{d}^* \circ \mathrm{d}^* = 0$, one obtains that the Laplacian commutes with the exterior derivative and codifferential.

2.1.6. *Charges.* An important role is played by the set of integer-valued, compactly supported 2-forms q which satisfy $\mathrm{d}q = 0$ and have connected support. These functions are often called charges in connection with the Coulomb gas of Section 3. We denote by $\mathcal{Q}$ the set of these forms, i.e.,

$$\mathcal{Q} := \left\{ q : \mathbb{Z}^d \to \mathbb{Z}^{\binom{d}{2}} \, : \, |\operatorname{supp} q| < \infty, \ \operatorname{supp} q \text{ is connected and } \mathrm{d}q = 0 \right\}. \tag{2.15}$$

We may restrict our considerations to the charges of $\mathcal{Q}$ whose support is included in a cube $\square \subseteq \mathbb{Z}^d$; to this end, we introduce the notation

$$\mathcal{Q}_\square := \left\{ q : \mathbb{Z}^d \to \mathbb{Z}^{\binom{d}{2}} \, : \, \operatorname{supp} q \subseteq \square, \ \operatorname{supp} q \text{ is connected and } \mathrm{d}q = 0 \right\}.$$

An important result about the exterior derivative is the Poincaré lemma. We will need to use the following version of the lemma in the discrete setting for integer-valued forms. The result is stated in [50, Lemma 1]. A proof can be found in [27, Lemma 2.2]. We mention that the inequality (2.16) is not proved in [27, Lemma 2.2], but can be deduced from the argument (essentially, the inductive argument developed there can be combined with [27, (2.4)] to obtain the result).

Lemma 2.1 (Poincaré for integer-valued forms)**.** *Let k be an integer of the set $\{1, \ldots, d-1\}$ and q be a k-form with values in $\mathbb{Z}$ such that $\mathrm{d}q = 0$, then there exists a $(k-1)$-form n_q with values in $\mathbb{Z}$ such that $q = \mathrm{d}n_q$. Moreover, n_q can be chosen such that $\operatorname{supp} n_q$ is contained in the smallest hypercube containing the support of q and such that*

$$\left\| n_q \right\|_{L^\infty} \leq C \left\| q \right\|_1. \tag{2.16}$$

Given a point $(x, y) \in \mathbb{Z}^d \times \mathbb{Z}^d$, we denote by $\mathcal{Q}_x$ and $\mathcal{Q}_{x,y}$ the set of charges $q \in \mathcal{Q}$ such that the point x and the points x, y belong to the support of n_q respectively, i.e.,

$$\mathcal{Q}_x := \left\{ q \in \mathcal{Q} \, : \, x \in \operatorname{supp} n_q \right\} \quad \text{and} \quad \mathcal{Q}_{x,y} := \left\{ q \in \mathcal{Q} \, : \, x \in \operatorname{supp} n_q \text{ and } y \in \operatorname{supp} n_q \right\}. \tag{2.17}$$

Similarly, we define

$$\mathcal{Q}_{\square,x} := \left\{ q \in \mathcal{Q}_\square \, : \, x \in \operatorname{supp} n_q \right\} \quad \text{and} \quad \mathcal{Q}_{\square,x,y} := \left\{ q \in \mathcal{Q}_\square \, : \, x \in \operatorname{supp} n_q \text{ and } y \in \operatorname{supp} n_q \right\}. \tag{2.18}$$

2.2. **Convention for constants and exponents.** 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 and the inverse temperature β. We use the symbols α, β, γ, δ to denote positive exponents which depend only on the dimension d. 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 $C := C(d,\beta,t)$ to mean that the constant C depends on the parameters d,β and t.

When the constants depend on the charges $q \in \mathcal{Q}$ (see (2.15)), we frequently keep track of their dependence on this parameter; more specifically we need that the growth of the constant C is at most algebraic in the parameter $\|q\|_1$. We usually denote by C_q a constant which depends on the parameters d,β and q and which satisfies the growth condition $C_q \leq C\|q\|_1^k$, for some $C := C(d,\beta) < \infty$ and $k := k(d) < \infty$. We allow the values of C and k to vary from line to line and we may write

$$C_q + C_q \leq C_q \quad \text{or} \quad C_q C_q \leq C_q.$$

We usually do not keep track of the dependence of the constants on the inverse temperature β (even though we believe it should be possible with our techniques) except in Sections 5 and 6. In these two sections, we assume that the constants depend only on the dimension d and make it explicit if they depend on the inverse temperature β.

3. Duality and Helffer-Sjöstrand representation

3.1. **From Villain model to solid on solid model.** In this section we recall the duality relation between the Villain model in $\mathbb{Z}^d$ and a statistical mechanical model of lattice Coulomb gas, with integer-valued and locally neutral charges (which can also be viewed as a solid-on-solid model) defined on $\Lambda^2(\mathbb{Z}^d)$, as observed in [50]. One may then perform a Fourier transform with respect to the charge variable, and obtain a classical random field representation of the Coulomb gas, known as the sine-Gordon representation. When the temperature is low enough, we may apply a one-step renormalization argument, following the presentation of Bauerschmidt [16] (see also [50]), to reduce the effective activity of the charges, thus obtain an effective, real valued random interface model on 2-forms with a uniformly convex potential.

Recall that the partition function for the Villain model in a cube $\square \subseteq \mathbb{Z}^d$ with zero boundary condition is given by

$$Z_{\square,0} := \int \prod_{e \subseteq E(\square)} \sum_{m \in \mathbb{Z}} \exp\left(-\frac{\beta}{2} \left(\nabla \theta(e) - 2\pi m \right)^2 \right) \prod_{x \in \partial \square} \delta_0 \left(\theta(x) \right) \prod_{x \in \square^\circ} \mathbb{1}_{[-\pi,\pi)}(\theta(x)) \, \mathrm{d}\theta(x).$$

Since we will need to use the formalism of discrete differential forms in this section, we note that the function $\theta : \square \mapsto \mathbb{R}$ can be seen as a 0-form, in that case the discrete gradient $\nabla \theta$ can be seen as a 1-form and is equal to the exterior derivative $\mathrm{d}\theta$. We may thus rewrite

$$Z_{\square,0} := \int \prod_{e \subseteq E(\square)} \sum_{m \in \mathbb{Z}} \exp\left(-\frac{\beta}{2} \left(\mathrm{d} \theta(e) - 2\pi m \right)^2 \right) \prod_{x \in \partial \square} \delta_0 \left(\theta(x) \right) \prod_{x \in \square^\circ} \mathbb{1}_{[-\pi,\pi)}(\theta(x)) \, \mathrm{d}\theta(x).$$

Permuting the sum with the product and the integral, we obtain

$$Z_{\square,0} = \sum_{\mathbf{m} \in \mathbb{Z}^{E(\square)}_{\mathbf{t}=0}} \int \prod_{e \subseteq E(\square)} \exp\left(-\frac{\beta}{2} \left(\mathrm{d} \theta(e) - 2\pi \mathbf{m}(e) \right)^2 \right) \prod_{x \in \partial \square} \delta_0 \left(\theta(x) \right) \prod_{x \in \square^\circ} \mathbb{1}_{[-\pi,\pi)}(\theta(x)) \, \mathrm{d}\theta(x), \tag{3.1}$$

where we have used the notation

$$\mathbb{Z}^{E(\square)}_{\mathbf{t}=0} := \left\{ \mathbf{m} : E(\square) \mapsto \mathbb{Z} \, : \, \mathbf{t}\mathbf{m} = 0 \text{ on } \partial\square \right\}.$$

Observe that we may split the sum according to

$$\sum_{\mathbf{m} \in \mathbb{Z}^{E(\square)}_{\mathbf{t}=0}} = \sum_{q \in \mathbb{Z}^{F(\square)}_{\mathbf{t}=0}, \mathrm{d}q = 0} \; \sum_{\mathbf{m} \in \mathbb{Z}^{E(\square)}_{\mathbf{t}=0}, \mathrm{d}\mathbf{m} = q}, \tag{3.2}$$

where we have set

$$\mathbb{Z}^{F(\square)}_{\mathbf{t}=0} := \left\{ q : F(\square) \mapsto \mathbb{Z} \, : \, \mathbf{t} q = 0 \text{ on } \partial\square \right\}.$$

A combination of (3.1) and (3.2) yields

$$Z_{\square,0} = \sum_{q \in \mathbb{Z}^{F(\square)}_{\mathbf{t}=0}, \mathrm{d}q = 0} \; \sum_{\mathbf{m} \in \mathbb{Z}^{E(\square)}_{\mathbf{t}=0}, \mathrm{d}\mathbf{m} = q} \int \prod_{e \subseteq E(\square)} \exp\left(-\frac{\beta}{2} \left(\mathrm{d} \theta(e) - 2\pi \mathbf{m}(e) \right)^2 \right) \prod_{x \in \partial \square} \delta_0 \left(\theta(x) \right) \prod_{x \in \square^\circ} \mathbb{1}_{[-\pi,\pi)}(\theta(x)) \, \mathrm{d}\theta(x).$$

Here $q : F(\square) \to \mathbb{Z}$ is the "vortex charge" on each plaquette of $\square$, which arises, informally, from

$$\oint_F \mathrm{d}\theta(e) = 2\pi q(F).$$

For each $q \in \mathbb{Z}^{F(\square)}_{\mathbf{t}=0}$ satisfying $\mathrm{d}q = 0$, we denote by $\mathbf{n}_q$ an element of $\mathbb{Z}^{E(\square)}_{\mathbf{t}=0}$ such that $\mathrm{d}\mathbf{n}_q = q$, chosen arbitrarily among all the possible candidates (the set of candidates is not empty by Proposition 2.1). Using that each 1-form $\mathbf{m} \in \mathbb{Z}^{E(\square)}_{\mathbf{t}=0}$ satisfying $\mathrm{d}\mathbf{m} = 0$ can be uniquely written $\mathrm{d}w$, for some $w : \square \mapsto \mathbb{Z}$ satisfying $w = 0$ on the boundary $\partial\square$, one can rewrite the previous display according to

$$Z_{\square,0} = \sum_{q \in \mathbb{Z}^{F(\square)}_{\mathbf{t}=0}, \mathrm{d}q = 0} \; \sum_{w \in \mathbb{Z}^{\square}_0} \int \prod_{e \subseteq E(\square)} \exp\left(-\frac{\beta}{2} \left(\mathrm{d} \theta(e) - 2\pi \left(\mathbf{n}_q + \mathrm{d}w \right)(e) \right)^2 \right) \prod_{x \in \partial \square} \delta_0 \left(\theta(x) \right) \prod_{x \in \square^\circ} \mathbb{1}_{[-\pi,\pi)}(\theta(x)) d\theta(x),$$

where we have set

$$\mathbb{Z}^{\square}_0 := \left\{ w : \square \mapsto \mathbb{Z} \, : \, w = 0 \text{ on } \partial\square \right\}.$$

Using the change of variable $\phi := \theta - 2\pi w$, and summing over all the maps $w \in \mathbb{Z}^{\square}_0$, one obtains

$$Z_{\square,0} = \sum_{q \in \mathbb{Z}^{F(\square)}_{\mathbf{t}=0}, \mathrm{d}q = 0} \int_{\mathbb{R}^{\square}} \prod_{e \subseteq E(\square)} \exp\left(-\frac{\beta}{2} \left(\mathrm{d} \phi(e) - 2\pi \mathbf{n}_q(e) \right)^2 \right) \prod_{x \in \partial \square} \delta_0 \left(\phi(x) \right) \prod_{x \in \square^\circ} d\phi(x).$$

We then decompose the function n_q as a sum of an exact and co-exact form. Specifically, one can prove that there exists a function $\phi_{n_q} \in \mathbb{Z}_0^{\square}$ and a two form $\psi_{n_q} \in \mathbb{Z}_{\mathbf{t}=0}^{F(\square)}$ such that

$$\mathbf{n}_q = \mathrm{d}\phi_{\mathbf{n}_q} + \mathrm{d}^* \psi_{n_q}, \tag{3.3}$$

The function ψ_{n_q} can in fact be identified more precisely: there exists a linear operator $(-\Delta_\square)^{-1}$ (which corresponds to inverting the Laplacian in the box $\square$ with the suitable boundary condition as explained below) such that

$$\psi_{n_q} = (-\Delta_\square)^{-1} q.$$

To be more specific, the operator $(-\Delta_\square)^{-1}$ is defined for general k-forms as follows. For each $i \in \{1, \ldots, \binom{d}{2}\}$, let us denote by $\partial_{I_i}\square$ the subset of faces of the boundary $\partial\square$ which are parallel to the cell c_{I_i}, and fix a k-form $q \in \Lambda^k(\square)$. We then let $w := (w_1, \ldots, w_{\binom{d}{k}})$ be the solution of the boundary value problem

$$\begin{cases} -\Delta w_i = q_i \text{ in } \square, \\ w_i = 0 \text{ in } \partial_{I_i}\square, \\ \nabla w_i \cdot \mathbf{n} = 0 \text{ on } \partial\,\square \setminus \partial_{I_i}\square, \end{cases} \tag{3.4}$$

that is, we solve the Laplace equation with Dirichlet boundary condition on the faces parallel to the cell c_{I_i} and Neumann boundary condition on the cells orthogonal to c_{I_i} (see Proposition A.5). We then define

$$(-\Delta_\square)^{-1} q := w.$$

Then using the translation invariance of the Lebesgue measure, we obtain

$$Z_{\square,0} = \sum_{q \in \mathbb{Z}_{\mathbf{t}=0}^{F(\square)}, \mathrm{d}q=0} \int_{\mathbb{R}^\square} \prod_{e \subseteq E(\square)} \exp\left(-\frac{\beta}{2}\left(\mathrm{d}\phi(e) - 2\pi \mathrm{d}^* (-\Delta_\square)^{-1} q(e)\right)^2\right) \prod_{x \in \partial\square} \delta_0(\phi(x)) \prod_{x \in \square} \mathrm{d}\phi(x).$$

The previous identity can be simplified

$$\begin{aligned} Z_{\square,0} &= Z_{GFF} \times Z(0) \\ &:= \int_{\mathbb{R}^\square} \exp\left(-\frac{\beta}{2}(\mathrm{d}\phi, \mathrm{d}\phi)\right) \prod_{x\in\partial\square} \delta_0(\phi(x)) \prod_{x\in\square} \mathrm{d}\phi(x) \times \sum_{q \in \mathbb{Z}_{\mathbf{t}=0}^{F(\square)}, \mathrm{d}q=0} \exp\left(-2\pi^2\beta\left(q, (-\Delta_\square)^{-1} q\right)\right). \end{aligned} \tag{3.5}$$

Using the identity $\mathrm{d}\phi = \nabla\phi$ (valid for 0-forms), we see that the first term in the left hand side of (3.5) is the partition function of the discrete Gaussian free field in the cube $\square$ with Dirichlet boundary condition. In other words, the Villain partition function factorizes into the partition function of a Gaussian free field, and the vortex charges that form a (neutral) Coulomb gas.

One can use the same argument to study the two-point function

$$\left\langle e^{i(\theta(x)-\theta(0))}\right\rangle_{\mu^V_{\beta,\square,0}}.$$

For any point $x \in \square$, define the string observable $h_{0x} : E(\mathbb{Z}^d) \mapsto \mathbb{Z}$ to be the indicator function of a (arbitrarily chosen) line joining 0 to x such that $\mathrm{d}^* h_{0x} = \mathbb{1}_x - \mathbb{1}_0$ and $h_{\{\square,x\}} : E(\square) \mapsto \mathbb{Z}$ be the indicator of the straight line connecting x to the boundary of the box $\square$ in the direction e_1. We have by definition $\mathrm{d}^* h_{0x} = \mathbb{1}_x - \mathbb{1}_0$ in $\mathbb{Z}^d$ and $\mathrm{d}^* h_{\{\square,x\}} = \mathbb{1}_x$ in the box $\square$. With the same computation, we obtain

$$\left\langle e^{i(\theta(x)-\theta(0))}\right\rangle_{\mu^V_{\beta,\square,0}} = \left\langle e^{i(\phi(x)-\phi(0))}\right\rangle_{GFF} \left\langle e^{-2i\pi\left(q, (-\Delta_\square)^{-1} \mathrm{d}h_{0x}\right)}\right\rangle_{\mu_C(\beta)} \tag{3.6}$$

and

$$\left\langle e^{i\theta(x)}\right\rangle_{\mu^V_{\beta,\square,0}} = \left\langle e^{i\phi(x)}\right\rangle_{GFF} \left\langle e^{-2i\pi\left(q, (-\Delta_\square)^{-1} \mathrm{d}h_{\{\square,x\}}\right)}\right\rangle_{\mu_C(\beta)}.$$

Here

$$\left\langle e^{i(\phi(x)-\phi(0))}\right\rangle_{GFF} := Z_{GFF}^{-1} \times \int_{\mathbb{R}^\square} e^{i(\phi(x)-\phi(0))} \exp\left(-\frac{\beta}{2}(\nabla\phi, \nabla\phi)\right) \prod_{x\in\partial\square} \delta_0(\phi(x)) \prod_{x\in\square^\circ} d\phi(x)$$

and

$$\left\langle e^{-2i\pi\left(q, (-\Delta_\square)^{-1} \mathrm{d}h_{0x}\right)}\right\rangle_{\mu_C(\beta)} = Z(0)^{-1} \times \sum_{q \in \mathbb{Z}_{\mathbf{t}=0}^{F(\square)}, \mathrm{d}q=0} e^{-2\pi^2\beta\left(q, (-\Delta_\square)^{-1} q\right)} e^{-2i\pi\left(q, (-\Delta_\square)^{-1} \mathrm{d}h_{0x}\right)}.$$

Following [16], we define the functions

$$\sigma_{\{\square,x\}} := (-\Delta_\square)^{-1} \mathrm{d}h_{\{\square,x\}} \quad \text{and} \quad \sigma_{\{\square,0x\}} := (-\Delta_\square)^{-1} \mathrm{d}h_{0x}.$$

For later purposes, we note that one has the pointwise convergences

$$\sigma_{\{\square,0\}} \underset{\square\uparrow\infty}{\longrightarrow} (-\Delta)^{-1} \mathrm{d}h_0, \quad \sigma_{\{\square,x\}} \underset{\square\uparrow\infty}{\longrightarrow} (-\Delta)^{-1} \mathrm{d}h_x \quad \text{and} \quad \sigma_{\{\square,0x\}} \underset{\square\uparrow\infty}{\longrightarrow} (-\Delta)^{-1} \mathrm{d}h_{0x}, \tag{3.7}$$

where h_0 (resp. h_x) is the indicator of the straight line starting from 0 (resp. x) in the direction e_1. To ease the notation, we denote the limiting functions in (3.7) by

$$\sigma_0 := (-\Delta)^{-1} \mathrm{d}h_0, \quad \sigma_x := (-\Delta)^{-1} \mathrm{d}h_x \quad \text{and} \quad \sigma_{0x} := (-\Delta)^{-1} \mathrm{d}h_{0x}.$$

In particular, using that the Laplacian commutes with the operators d and d*, we obtain

$$\mathrm{d}^*\sigma_0 = (-\Delta)^{-1} \mathrm{d}^*\mathrm{d}h_0 = -h_0 - (-\Delta)^{-1} \mathrm{d}\mathrm{d}^*h_0 = -h_0 - (-\Delta)^{-1} \mathrm{d}\mathbb{1}_0 = -h_0 - \nabla G, \tag{3.8}$$

where G is the standard random walk Green's function on the lattice $\mathbb{Z}^d$. A consequence of the identity (3.8) is the equality

$$e^{-2i\pi(q,\sigma_0)} = e^{-2i\pi(n_q,\nabla G)}. \tag{3.9}$$

Similar statements hold for the maps σ_x and σ_{0x}, and we may write

$$e^{-2i\pi(q,\sigma_x)} = e^{-2i\pi(n_q,\nabla G_x)} \text{ and } e^{-2i\pi(q,\sigma_{0x})} = e^{-2i\pi(n_q,\nabla G_x - \nabla G)} \tag{3.10}$$

where we have used the notation $G_x := G(\cdot - x)$.

We next collect the following identity: for each $q \in \mathbb{Z}^{E(\square)}_{\mathbf{t}=0}$ satisfying $\mathrm{d}q = 0$, one has

$$\left(q, \sigma_{\{\square,x\}} - \sigma_{\{\square,0\}}\right) = \left(q, \sigma_{\{\square,0x\}}\right) \mod \mathbb{Z}. \tag{3.11}$$

To justify the identity (3.11), we note that, with the same argument as in (3.3), we may write

$$h_{\{\square,x\}} - h_{\{\square,0\}} - h_{\{\square,0x\}} = \mathrm{d}\phi + \mathrm{d}^*\sigma_{\{\square,x\}} - \mathrm{d}^*\sigma_{\{\square,0\}} - \mathrm{d}^*\sigma_{\{\square,0x\}},$$

for some field $\phi : \square \to \mathbb{R}$ satisfying $\phi = 0$ on the boundary $\partial\square$. Taking the scalar product with the 1-form $\mathrm{d}\phi$, and performing integrations by parts, we obtain that

$$\left(\mathrm{d}\phi, h_{\{\square,x\}} - h_{\{\square,0\}} - h_{0x}\right) = \left(\phi, \mathrm{d}^*h_{\{\square,x\}} - \mathrm{d}^*h_{\{\square,0\}} - \mathrm{d}^*h_x\right) = \left(\phi, \mathbb{1}_x - \mathbb{1}_0 - (\mathbb{1}_x - \mathbb{1}_0)\right) = 0$$

and

$$\left(\mathrm{d}\phi + \mathrm{d}^*\sigma_{\{\square,x\}} - \mathrm{d}^*\sigma_{\{\square,0\}} - \mathrm{d}^*\sigma_{\{\square,0x\}}, \mathrm{d}\phi\right) = (\mathrm{d}\phi, \mathrm{d}\phi) + \left(\sigma_{\{\square,x\}} - \sigma_{\{\square,0\}} - \sigma_{\{\square,0x\}}, \mathrm{d}\mathrm{d}\phi\right) = (\mathrm{d}\phi, \mathrm{d}\phi).$$

A combination of the two previous displays implies $\mathrm{d}\phi = 0$ and thus $h_{\{\square,x\}} - h_{\{\square,0\}} - h_{0x} = \mathrm{d}^*\sigma_{\{\square,x\}} - \mathrm{d}^*\sigma_{\{\square,0\}} - \mathrm{d}^*\sigma_{\{\square,0x\}}$. We then use that $q = \mathrm{d}n_q$ for some $n_q \in \mathbb{Z}^{E(\square)}_{\mathbf{t}=0}$ to write

$$\left(q, \sigma_{\{\square,x\}} - \sigma_{\{\square,0\}} - \sigma_{\{\square,0x\}}\right) = \left(n_q, \mathrm{d}^*\sigma_{\{\square,x\}} - \mathrm{d}^*\sigma_{\{\square,0\}} - \mathrm{d}^*\sigma_{\{\square,0x\}}\right) = \left(n_q, h_{\{\square,x\}} - h_{\{\square,0\}} - h_{0x}\right) \in \mathbb{Z}.$$

This is (3.11). A consequence of (3.11) is that for each $q \in \mathbb{Z}^{E(\square)}_{\mathbf{t}=0}$ satisfying $\mathrm{d}q = 0$,

$$e^{-2i\pi\left(q,\sigma_{\{\square,x\}} - \sigma_{\{\square,0\}}\right)} = e^{-2i\pi\left(q,\sigma_{\{\square,0x\}}\right)}.$$

We set the notation, for each $\sigma : F(\square) \to \mathbb{R}$,

$$Z(\sigma) := \sum_{q \in \mathbb{Z}^{F(\square)}_{\mathbf{t}=0}, \mathrm{d}q = 0} e^{-2\pi^2\beta\left(q,(-\Delta_\square)^{-1}q\right)} e^{-2i\pi(q,\sigma)}. \tag{3.12}$$

So that

$$\frac{Z\left(\sigma_{\{\square,x\}}\right)}{Z(0)} = \left\langle e^{-2i\pi\left(q,\sigma_{\{\square,x\}}\right)}\right\rangle_{\mu_C(\beta)} \text{ and } \frac{Z\left(\sigma_{\{\square,0x\}}\right)}{Z(0)} = \left\langle e^{-2i\pi\left(q,\sigma_{\{\square,0x\}}\right)}\right\rangle_{\mu_C(\beta)}.$$

Next, we introduce the functional space $C(\square)$ defined as follows

$$C(\square) := \left\{\phi := \left(\phi_1, \ldots, \phi_{\binom{d}{2}}\right) : \square \to \mathbb{R}^{\binom{d}{2}} \,:\, \forall i \in \left\{1, \ldots, \binom{d}{2}\right\},\ \phi_i = 0 \text{ on } \partial\,\square \setminus \partial_{I_i}\square\right\}.$$

and denote by $\phi := \left(\phi_1, \ldots, \phi_{\binom{d}{k}}\right)$ the vector-valued Gaussian free field valued in the space $C(\square)$ (or more specifically, the Gaussian field whose covariance matrix is given by the finite volume Green's function associated with the Laplace equation described in (3.4)). Using the previous definition, we see that, for each $q \in \mathbb{Z}^{E(\square)}$ satisfying $\mathrm{d}q = 0$ and $\mathbf{t}q = 0$,

$$\mathbb{E}\left[e^{2i\pi(q,\phi)}\right] = e^{-2\pi^2\beta\left(q,(-\Delta_\square)^{-1}q\right)} = e^{-2\pi^2\beta\left(q,(-\Delta_\square)^{-1}q\right)}.$$

Consequently,

$$Z\left(\sigma_{\{\square,0x\}}\right) = \sum_{q \in \mathbb{Z}^{F(\square)}_{\mathbf{t}=0}, \mathrm{d}q=0} \mathbb{E}\left[e^{-2i\pi\left(q,\phi+\sigma_{\{\square,0x\}}\right)}\right].$$

Thus the partition function of this lattice Coulomb gas can be represented in terms of a characteristic function with respect to a Gaussian measure. We then claim that, for β sufficiently large, a one-step renormalization maps the Coulomb gas model to an effective one with very small effective activity. Using that the discrete Laplacian is bounded from above, one has that $(-\Delta_\square)^{-1} \geq c$, for some $c := c(d) > 0$. We then choose the inverse temperature β larger than the value c^2 and decompose the Gaussian field ϕ as the sum of two independent Gaussian fields $\phi_1 + \phi_2$, such that ϕ_1 and ϕ_2 have covariance matrices $\beta\left((-\Delta_\square)^{-1} - \beta^{-\frac{1}{2}}\mathrm{Id}\right)$ and $\beta^{\frac{1}{2}}\mathrm{Id}$. We can thus write

$$Z\left(\sigma_{\{\square,0x\}}\right) = \sum_{q \in \mathbb{Z}^{F(\square)}_{\mathbf{t}=0}, \mathrm{d}q=0} \mathbb{E}\left[e^{-2i\pi\left(q,\phi_1+\phi_2+\sigma_{\{\square,0x\}}\right)}\right] = \sum_{q \in \mathbb{Z}^{F(\square)}_{\mathbf{t}=0}, \mathrm{d}q=0} e^{-\pi^2\beta^{1/2}(q,q)}\mathbb{E}_{\mu_1}\left[e^{-2i\pi\left(q,\phi_1+\sigma_{\{\square,0x\}}\right)}\right],$$

where μ_1 is a Gaussian measure on $C(\square)$, given by

$$d\mu_1(\phi_1) = \mathrm{Const} \times \exp\left(-\frac{1}{2\beta}\left(\phi_1, \left((-\Delta_\square)^{-1} - \beta^{-\frac{1}{2}}\mathrm{Id}\right)^{-1}\phi_1\right)\right)d\phi_1,$$

where $d\phi_1$ denotes the Lebesgue measure on the space $C(\square)$. For β sufficiently large, we may expand $\left((-\Delta_\square)^{-1} - \beta^{-\frac{1}{2}}\mathrm{Id}\right)^{-1}$ into a convergent sum

$$\left((-\Delta_\square)^{-1} - \beta^{-\frac{1}{2}}\mathrm{Id}\right)^{-1} = -\Delta + \sum_{n\geq 1}\frac{1}{\beta^{n/2}}(-\Delta)^{n+1},$$

where in the right-hand side, the symbol Δ refers to the discrete Laplacian acting on the space $C(\square)$ (with the corresponding boundary condition so that it can be iterated). Thus

$$d\mu_1(\phi_1) = Z_1^{-1} \times \exp\left(\frac{1}{2\beta}(\phi_1, \Delta\phi_1) - \sum_{n\geq 1}\frac{1}{2\beta}\frac{1}{\beta^{n/2}}(\phi_1, (-\Delta)^{n+1}\phi_1)\right)\mathbb{1}_{\phi_1 \in C(\square)}\, d\phi_1.$$

Following [16], (especially Lemmas 5.14 and 5.15 there), since $e^{-\pi^2\beta^{1/2}(q,q)}$ decays to zero rapidly in $\|q\|_1 := \sum_{x\in F(\square)}|q(x)|$, we may apply a cluster expansion to conclude that for β large enough, one can re-sum $Z(\sigma_{0x})$ as

$$Z(\sigma_{0x}) = \mathbb{E}_{\mu_1}\left[\exp\left(\sum_{q\in\mathcal{Q}_\square} z(\beta,q)e^{-2i\pi\left(q,\phi_1+\sigma_{\{\square,0x\}}\right)}\right)\right] = \mathbb{E}_{\mu_1}\left[\exp\left(\sum_{q\in\mathcal{Q}_\square} z(\beta,q)\cos\left(2\pi\left(q,\phi_1+\sigma_{\{\square,0x\}}\right)\right)\right)\right],$$

where the sum is over all lattice animals $q \in \mathcal{Q}_\square$ with connected support satisfying $\mathrm{d}q = 0$ in the cube $\square$ and $\mathbf{t}q = 0$ on the boundary $\partial\square$ (see (2.15)), and $z(\beta,q)$ is a real number given by the formula (see [16, (5.71)])

$$z(\beta,q) = \sum_{n=1}^{\infty}\frac{1}{n!}I(G(\mathrm{supp}\,q_1,\ldots,\mathrm{supp}\,q_n))\sum_{q_1+\ldots+q_n=q}e^{-\frac{1}{2}c\beta\sum_i(q_i,q_i)}, \tag{3.13}$$

where the sum runs over all the charges $q_1,\ldots,q_n$ with connected support satisfying $\mathrm{d}q_i = 0$, and the combinatorial factor $I(G(\mathrm{supp}\,q_1,\ldots,\mathrm{supp}\,q_n))$ is defined as follows: we let $G(\mathrm{supp}\,q_1,\ldots,\mathrm{supp}\,q_n)$ be the connection graph of the sets $\mathrm{supp}\,q_1,\ldots,\mathrm{supp}\,q_n$ (i.e., the graph whose vertices are $\mathrm{supp}\,q_1,\ldots,\mathrm{supp}\,q_n$, and with an edge between $\mathrm{supp}\,q_i$ and $\mathrm{supp}\,q_j$ if and only if the two sets have nonempty intersection), and for a connected graph G, we define

$$I(G) := \sum_{H\subseteq G}(-1)^{|E(H)|},$$

where the sum runs over all the connected spanning subgraphs of G. These definitions and formulae are the ones of [16, Section 5.5.3]. A few observations can be deduced from them:

- The real number $z(\beta,q)$ depends only on the charge q, and the inverse temperature β in particular, it does not depend on the box $\square$ or on the vertex x;
- The coefficient $z(\beta,q)$ satisfies some invariance properties with respect to the charge q and is not affected by translation or rotations of the charge as well as reflections (in fact the coefficient is invariant under any linear transformation preserving the lattice $\mathbb{Z}^d$ applied to the charge q);
- By [16, Lemma 5.15], one has the estimate

$$|z(\beta,q)| \leq e^{-c\beta^{1/2}\|q\|_1}, \quad \text{for some } c := c(d) > 0. \tag{3.14}$$

Similarly,

$$Z(0) = \mathbb{E}_{\mu_1}\left[\exp\left(\sum_{q\in\mathcal{Q}_\square} z(\beta,q)e^{-2i\pi(q,\phi_1)}\right)\right] = \mathbb{E}_{\mu_1}\left[\exp\left(\sum_{q\in\mathcal{Q}_\square} z(\beta,q)\cos\left(2\pi\left(q,\phi_1\right)\right)\right)\right].$$

Using the trigonometric identity

$$\cos\left(2\pi\left(q,\phi_1+\sigma_{\{\square,0x\}}\right)\right) = \cos\left(2\pi\left(q,\phi_1\right)\right)\cos\left(2\pi\left(q,\sigma_{\{\square,0x\}}\right)\right) - \sin\left(2\pi\left(q,\phi_1\right)\right)\sin\left(2\pi\left(q,\sigma_{\{\square,0x\}}\right)\right),$$

we may write

$$(3.15)\quad \frac{Z(\sigma_{\{\square,0x\}})}{Z(0)}$$
$$= \left\langle \exp\left(\sum_{q\in\mathcal{Q}_\square} z(\beta,q)\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\sigma_{\{\square,0x\}},q)\right) + \sum_{q\in\mathcal{Q}_\square} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(\left(2\pi(\sigma_{\{\square,0x\}},q)\right)\right)-1\right)\right)\right\rangle_{\mu_{\beta,\square}}.$$

Here $\mu_{\beta,\square}$ is defined as a measure on the space $C(\square)$ by

$$(3.16)\quad d\mu_{\beta,\square}(\phi) := \mathrm{Const}\times\exp\left(\frac{1}{2\beta}(\phi,\Delta\phi) - \sum_{n\ge1}\frac{1}{2\beta}\frac{1}{\beta^{n/2}}(\phi,(-\Delta)^{n+1}\phi) + \sum_{q\in\mathcal{Q}_\square} z(\beta,q)\cos\left(2\pi\left(q,\phi\right)\right)\right)d\phi.$$

Combining (3.6) and (3.15), we have the following dual representation for the two-point function of the Villain model. Let $G_\square$ be the Dirichlet Green's function defined on the vertices of the cube $\square$,

$$(3.17)\quad \begin{cases} -\Delta G_\square(\cdot,x) = \delta_x \text{ in } \square, \\ \quad G_\square(\cdot,x) = 0 \text{ on } \partial\square. \end{cases}$$

Proposition 3.1. *There exists an inverse temperature $\beta_1 := \beta_1(d) < \infty$ such that for any $\beta \ge \beta_1$,*

$$(3.18)\quad \left\langle e^{i(\theta(x)-\theta(0))}\right\rangle_{\mu^V_{\beta,\square,0}}\exp\left(\frac{1}{2\beta}\left(G_\square(x,x)+G_\square(0,0)-2G_\square(0,x)\right)\right)$$
$$= \left\langle \exp\left(\sum_{q\in\mathcal{Q}_\square} z(\beta,q)\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\sigma_{\{\square,0x\}},q)\right) + \sum_{q\in\mathcal{Q}_\square} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(\left(2\pi(\sigma_{\{\square,0x\}},q)\right)\right)-1\right)\right)\right\rangle_{\mu_{\beta,\square}}.$$

Following the same argument, we also obtain the dual representation for $\left\langle e^{i(\theta(x)+\theta(0))}\right\rangle_{\mu^V_{\beta,\square,0}}$. Define $\overline{\sigma}_{\{\square,0x\}} := \sigma_{\{\square,0\}} + \sigma_{\{\square,x\}}$. We then have

$$(3.19)\quad \left\langle e^{i(\theta(x)+\theta(0))}\right\rangle_{\mu^V_{\beta,\square,0}}\exp\left(\frac{1}{2\beta}\left(G_\square(x,x)+G_\square(0,0)-2G_\square(0,x)\right)\right)$$
$$= \left\langle \exp\left(\sum_{q\in\mathcal{Q}_\square} z(\beta,q)\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\overline{\sigma}_{\{\square,0x\}},q)\right) + \sum_{q\in\mathcal{Q}_\square} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\overline{\sigma}_{\{\square,0x\}},q)\right)-1\right)\right)\right\rangle_{\mu_{\beta,\square}}.$$

In view of (3.6), to study the two-point function of the (finite-volume) Villain model, it suffices to compute the expectation of a non-linear functional (3.18) with respect to the Gibbs measure $\mu_{\beta,\square}$. Notice that the neutrality condition $\mathrm{d}q = 0$ indicates $\mu_{\beta,\square}$ is a measure of *gradient-type*, i.e., the Hamiltonian only depends on the discrete gradient $\nabla\phi$. Additionally, for β large, the exponential smallness of $z(\beta,q)$ implies that $\mu_{\beta,\square}$ is a smooth perturbation of the discrete Gaussian free field

$$\mu_{GFF}(d\phi) := \mathrm{Const}\times\exp\left(\frac{1}{2\beta}(\phi,\Delta\phi)\right)d\phi.$$

These observations imply that the measure $\mu_{\beta,\square}$ belongs to the class of models in statistical physics known as the uniformly convex $\nabla\phi$ model. This category of models has been extensively studied in the literature, and we refer to [51] for a description of its literature. In particular, one can apply the techniques and tools developed in the context of the $\nabla\phi$ model to study the asymptotic behavior of the measure $\mu_{\beta,\square}$. This is the subject on the next sections where:

- We apply the Brascamp-Lieb inequality [24] to the measure $\mu_{\beta,\square}$ and use it to prove the existence of a thermodynamic limit, denoted by μ_β, for the measure $\mu_{\beta,\square}$ (i.e., the existence of an infinite-volume limiting measure when $|\square|\to\infty$).

- We present the standard tool used to study the macroscopic behavior of the model known as the Helffer-Sjöstrand representation and combine it with quantitative homogenization to show that on large scales the measure μ_β behaves like an *effective* Gaussian free field, with the covariance matrix depending on β.

Remark 3.2. On a heuristical level, the second point above (asserting that the measure μ_β behaves over large scales as an effective Gaussian free field) is sufficient to justify that the subleading order of (3.15) (and therefore, of the truncated two-point function) should decay asymptotically as $C|x|^{2-d}$, for some constant C depending on β.

To see this, we first note that, for β sufficiently large, the inequality (3.14) implies that the coefficient $z(\beta, q)$ decays exponentially fast as the L^1-norm of the charge q increases. On a heuristical level, we may make the following simplifying assumption: we assume that the coefficient $z(\beta, q)$ is equal to 0 for all the charges except on the simplest ones satisfying the neutrality condition, i.e., the charges of the form $q = (\delta_x - \delta_{x+e_i})$, for $i = 1, \ldots, d$ (also called dipoles), for which it takes a nonzero value denoted by $z(\beta)$. Thus the right side of (3.18) is approximately (after taking the limit $|\square| \to \infty$ to replace the finite-volume Gibbs measure $\mu_{\beta,\square}$ by the infinite-volume measure μ_β, the function $\sigma_{\square,0x}$ by σ_{0x} and using the identity (3.10))

$$\left\langle \exp\left(\sum_{e \in E(\square)} z(\beta) \sin\left(2\pi(\nabla\phi(e))\right) \sin\left(2\pi(\nabla G(e) - \nabla G_x(e))\right) \right) \right.$$
$$\left. \times \exp\left(\sum_{e \in E(\square)} z(\beta) \cos\left(2\pi(\nabla\phi(e))\right) \left(\cos\left(2\pi(\nabla G(e) - \nabla G_x(e))\right) - 1\right) \right) \right\rangle_{\mu_\beta}.$$

Since $|\cos\left(2\pi(\nabla G(e) - \nabla G_x(e))\right) - 1| \leq C(\nabla G(e) - \nabla G_x(e))^2$ decays fast away from 0 and x, let us assume for now that the term $\sum_{e \in E(\square)} z(\beta) \cos\left(2\pi(\nabla\phi(e))\right) \left(\cos\left(2\pi(\nabla G(e) - \nabla G_x(e))\right) - 1\right)$ only contributes to the lower order. By further making the approximation $\sin a \approx a$ for small a, we may further approximate the expression above by

$$\left\langle \exp\left(\sum_{e \in E(\square)} z(\beta) 2\pi(\nabla\phi(e)) 2\pi(\nabla G(e) - \nabla G_x(e)) \right) \right\rangle_{\mu_{\beta,\square}}.$$

Using an integration by parts, this equals to $\left\langle \exp(z(\beta) 4\pi^2(\phi(0) - \phi(x))) \right\rangle_{\mu_{\beta,\square}}$. Assuming that over large scales the measure μ_β behaves like a Gaussian free field, we may conclude

$$\left\langle \exp(z(\beta) 4\pi^2(\phi(0) - \phi(x))) \right\rangle_{\mu_{\beta,\square}} \approx \exp\left(\frac{1}{2} \operatorname{var}_{\mu_{\beta,\square}}(z(\beta) 4\pi^2(\phi(0) - \phi(x))) \right) \approx C_0(d,\beta) + C_1(d,\beta)|x|^{2-d}.$$

We remark that the computation above is only heuristical and the constants C_0, C_1 obtained are not the right constants. Indeed, the non-local charges in $\mathcal{Q}_\square$, the non-linear functions $\sin x$ and $\cos x$, and the non-Gaussian field $\mu_{\beta,\square}$ contribute to a nontrivial correction of these constants. Such corrections can be obtained rigorously through the homogenization of the Helffer-Sjöstrand PDE.

3.2. **Brascamp-Lieb inequality.** As we discussed, when β is sufficiently large, the measure $\mu_{\beta,\square}$ is a small smooth perturbation of a discrete Gaussian free field, and is in particular log-concave. In this framework, one is able to apply the celebrated Brascamp-Lieb inequality [24, 23] described below. We let $H : C(\square) \to \mathbb{R}$ be a (strictly) convex function satisfying $\int_{C(\square)} \exp(-H(\phi)) d\phi < \infty$, and introduce the probability measure

$$\mu(d\phi) := \frac{1}{Z} \exp\left(-H(\phi)\right) d\phi.$$

The Brascamp Lieb inequality estimates the variance of a general (differentiable) functional $F : C(\square) \to \mathbb{R}$ of the field ϕ under the measure μ. In order to state it, we will need the following notation

$$\left\langle \partial F, (\operatorname{Hess} H)^{-1} \partial F \right\rangle_\mu := \sum_{x,y \in \square^\circ} \sum_{i=1}^{\binom{d}{2}} \left\langle \partial_{x,i} F \left(\operatorname{Hess} H\right)^{-1}_{(x,i),(y,j)} \partial_{y,j} F \right\rangle_\mu, \tag{3.20}$$

where $(\operatorname{Hess} H)^{-1}$ is the inverse of the Hessian of H defined by $\operatorname{Hess} H := \left(\partial_{x,i}\partial_{y,j} H\right)_{(x,i),(y,j) \in \square \times \binom{d}{2}}$.

Proposition 3.3 (Brascamp-Lieb inequality for log-concave measures [24, 23]). *Let μ be the log-concave measure defined in (3.20). For any smooth and compactly supported function $F : C(\square) \to \mathbb{R}$, one has the upper bound*

$$\mathrm{var}_\mu [F] \leq \left\langle \partial F, (\mathrm{Hess} H)^{-1} \partial F \right\rangle_\mu .$$

We next apply the Brascamp-Lieb inequality to the measure $\mu_{\square,\beta}$. Specifically, we apply it to a class of observables which will be useful to study the Villain model and upgrade it to obtain an estimate on exponential moments (following the techniques of [51, Theorem 4.9]). In order to state the result, we first need to identify the Green's function associated with the Laplace equation (3.4). For each $x \in \square$ and each $i \in \left\{1, \ldots, \binom{d}{2}\right\}$, we let $G_{C(\square),i} : \square \to \mathbb{R}$ be the solution of the equation

$$\begin{cases} -\Delta G_{C(\square),i}(\cdot, x) = \delta_x \text{ in } \square, \\ \qquad G_{C(\square),i}(\cdot, x) = 0 \text{ in } \partial \square \setminus \partial_{I_i} \square, \\ \mathbf{n} \cdot \nabla G_{C(\square),i}(\cdot, x) = 0 \text{ in } \partial_{I_i} \square . \end{cases}$$

We note that, in dimension $d \geq 3$, the Green's function $G_{C(\square),i}(\cdot, x)$ is bounded uniformly in the vertex x, the box $\square$ and the index i. We record two properties of this finite-volume Green's function. First, for any box $\square$, any index $i \in \{1, \ldots, \binom{d}{2}\}$, and any pair of vertices $x, y \in \square$,

$$0 \leq G_{C(\square),i}(y, x) \leq \frac{C}{|x - y|^{d-2}}.$$

In the discrete setting and in dimensions $d \geq 3$, the Green's function is bounded and we have $G_{C(\square),i}(x, x) \leq C$, for any $i \in \left\{1, \ldots, \binom{d}{2}\right\}$, $\square \subseteq \mathbb{Z}^d$, and $x \in \square$. To include this case in the notation, we implicitly extend all the functions of the form $x \mapsto |x|^{-k}$ for $k \geq 0$ by the value 1 when $x = 0$ (as mentioned in Section 2.1).

Additionally, the finite-volume Green's function converges to the infinite-volume one and we have: for any index $i \in \{1, \ldots, \binom{d}{k}\}$, and any pair of vertices $x, y \in \mathbb{Z}^d$,

$$G_{C(\square),i}(y, x) \underset{\square \to \infty}{\longrightarrow} G(y, x).$$

In the case of the finite-volume Green's function defined on a box with Dirichlet boundary condition, they can be found in [75, Section 4.6]. They can be easily extended to the case considered here (where the boundary condition is a combination of the Dirichlet and Neumann boundary conditions).

Proposition 3.4 (Brascamp-Lieb inequality for $\mu_{\beta,\square}$). *Fix two constants $C_0 < \infty$ and $c > 0$. There exists an inverse temperature $\beta_1 := \beta_1(d, C_0, c) < \infty$ and a constant $C := C(d)$ such that for any $\beta \geq \beta_1$, the following two properties hold:*

- *For any vertex $x \in \square$,*

$$\left\langle \exp\left(|\phi(x)|\right) \right\rangle_{\mu_{\beta,\square}} \leq C. \tag{3.21}$$

- *For every collection of coefficients $(g(\beta, q))_{q \in \mathcal{Q}_\square}$ satisfying $|g(\beta, q)| \leq C_0 \exp(-c\beta^{1/2} \|q\|_1)$, if we denote by $Z := \sum_{q \in \mathcal{Q}_\square} g(\beta, q) \sin\left(2\pi(\phi, q)\right)$, then*

$$\left\langle \exp\left(Z\right) \right\rangle_{\mu_{\beta,\square}} \leq \exp\left(\sum_{x,y \in \square} \sum_{i=1}^{\binom{d}{2}} G_{C(\square),i}(x, y) \left\langle \left(\partial_{x,i} Z\right)\left(\partial_{y,i} Z\right) \right\rangle_{\mu_{\beta,\square}} \right). \tag{3.22}$$

Proof. We first apply the Brascamp-Lieb inequality with the map

$$H(\phi) := -\frac{1}{2\beta}(\phi, \Delta\phi) + \sum_{n \geq 1} \frac{1}{2\beta} \frac{1}{\beta^{n/2}} (\phi, (-\Delta)^{n+1} \phi) - \sum_{q \in \mathcal{Q}_\square} z(\beta, q) \cos\left(2\pi\left(q, \phi\right)\right). \tag{3.23}$$

The Hessian of the Hamiltonian H can then be explicitly computed: the first two terms of (3.23) are quadratic (thus their Hessian is constant), and the Hessian of the third term can be obtained by differentiating the cosine twice. We obtain the following identity: for any $\phi, \psi \in C(\square)$,

$$(\psi, \mathrm{Hess} H(\phi) \psi) = -\frac{1}{2\beta}(\psi, \Delta\psi) + \sum_{n \geq 1} \frac{1}{2\beta} \frac{1}{\beta^{n/2}} (\psi, (-\Delta)^{n+1} \psi) - \sum_{q \in \mathcal{Q}_\square} z(\beta, q) \left(q, \psi\right)^2 \cos\left(2\pi\left(q, \phi\right)\right).$$

Using that the second term is nonnegative and that the absolute value of the cosine is always smaller than 1, we deduce that

$$(\psi, \mathrm{Hess} H(\phi)\psi) \geq \frac{1}{2\beta} \|\nabla\psi\|^2_{L^2(\square)} - \sum_{q \in \mathcal{Q}_\square} |z(\beta,q)| \, (q,\psi)^2 .$$

We then estimate the second term in the right-hand side and prove that it is small compared to the first one. To this end, we write

$$\sum_{q \in \mathcal{Q}_\square} |z(\beta,q)| \, (q,\psi)^2 \leq \sum_{q \in \mathcal{Q}_\square} e^{-c\beta\|q\|_1} \, (q,\psi)^2 = \sum_{q \in \mathcal{Q}_\square} e^{-c\beta\|q\|_1} (n_q, \mathrm{d}^*\psi)^2. \tag{3.24}$$

We then use the Cauchy-Schwarz inequality and deduce

$$\begin{aligned} \sum_{q \in \mathcal{Q}_\square} e^{-c\beta\|q\|_1} \left(n_q, \mathrm{d}^*\psi\right)^2 &\leq \sum_{q \in \mathcal{Q}_\square} e^{-c\beta\|q\|_1} \|n_q\|^2_{L^2} \|\mathrm{d}^*\psi\|^2_{L^2(\mathrm{supp}\, n_q)} \\ &= \sum_{q \in \mathcal{Q}_\square} \sum_{y \in \mathrm{supp}\, n_q} e^{-c\beta\|q\|_1} \|n_q\|^2_{L^2} |\mathrm{d}^*\psi(y)|^2 \\ &\leq \sum_{y \in \square} |\mathrm{d}^*\psi(y)|^2 \left(\sum_{q \in \mathcal{Q}_{\square,y}} e^{-c\beta\|q\|_1} \|n_q\|^2_{L^2} \right). \end{aligned}$$

We then observe that the term in the right-hand side can be bounded as follows: one has the estimate, for any $y \in \mathbb{Z}^d$,

$$\sum_{q \in \mathcal{Q}_{\square,y}} e^{-c\beta\|q\|_1} \|n_q\|^2_{L^2} \leq C e^{-c\sqrt{\beta}}. \tag{3.25}$$

To prove this inequality, we first absorb the polynomial factor by using (A.17) and writing

$$\|n_q\|^2_{L^2} e^{-c\beta\|q\|_1} \leq \|q\|_1^{d+2} e^{-c\sqrt{\beta}\|q\|_1} \leq C e^{-c'\sqrt{\beta}\|q\|_1}$$

for some constant $c' \in (0,c)$. We then decompose over the supports of the charges. To this end, let us denote by $\mathcal{A}_y$ the set of the finite connected subsets of $\mathbb{Z}^d$ containing the vertex y. We then write

$$\sum_{q \in \mathcal{Q}_y} e^{-c\sqrt{\beta}\|q\|_1} = \sum_{X \in \mathcal{A}_y} \sum_{\substack{q \in \mathcal{Q}_\square \\ \mathrm{supp}\, q = X}} e^{-c'\sqrt{\beta}\|q\|_1} = \sum_{X \in \mathcal{A}_y} \sum_{\substack{q \in \mathcal{Q}_\square \\ \mathrm{supp}\, q = X}} \left(\prod_{x \in X} e^{-c'\sqrt{\beta}|q(x)|} \right).$$

Exchanging the sum and the product, we see that

$$\sum_{\substack{q \in \mathcal{Q}_\square \\ \mathrm{supp}\, q = X}} \left(\prod_{x \in X} e^{-c'\sqrt{\beta}|q(x)|} \right) \leq \prod_{x \in X} \left(\sum_{q(x)=1}^{\infty} e^{-c'\sqrt{\beta}|q(x)|} \right) = \left(\frac{e^{-c\sqrt{\beta}}}{1 - e^{-c\sqrt{\beta}}} \right)^{|X|}.$$

We thus obtain

$$\sum_{q \in \mathcal{Q}_y} e^{-c\sqrt{\beta}\|q\|_1} \leq C \sum_{X \in \mathcal{A}_y} e^{-c\sqrt{\beta}|X|} = C \sum_{n=1}^{\infty} |\{X \in \mathcal{A}_y \, : \, |X| = n\}| \, e^{-c\sqrt{\beta} n}.$$

We next note that

$$|\{X \in \mathcal{A}_y \, : \, |X| = n\}| \leq e^{Cn}. \tag{3.26}$$

The inequality (3.26) can be established by associating each connected set of n vertices with one of its spanning trees, and then bounding the number of such spanning trees. Choosing the inverse temperature β large enough (i.e., such that $c\sqrt{\beta} \geq 2C$), we deduce that

$$\sum_{q \in \mathcal{Q}_y} e^{-c\sqrt{\beta}\|q\|_1} \leq e^{-c\sqrt{\beta}},$$

where we have reduced the value of the exponent c in the right-hand side. Additionally, we have the estimate $\|\mathrm{d}^*\psi\|^2_{L^2(\square)} \leq C \|\nabla\psi(y)\|^2_{L^2(\square)}$ (this follows form the definition of the codifferential). Combining the previous displays, we obtain

$$\sum_{q \in \mathcal{Q}_\square} |z(\beta,q)| \, (n_q, \mathrm{d}^*\psi)^2 \leq C e^{-c\sqrt{\beta}} \|\nabla\psi\|_{L^2(\square)} .$$

Combining the four previous displays and choosing β small enough, we obtain

$$(\psi, \mathrm{Hess}\, H(\phi)\psi) \geq \frac{1}{2\beta} \|\nabla\psi\|^2_{L^2(\square)} - C e^{-c\sqrt{\beta}} \|\nabla\psi\|_{L^2(\square)} \geq \frac{1}{4\beta} \|\nabla\psi\|^2_{L^2(\square)} = \frac{1}{4\beta} (\psi, -\Delta\psi) . \tag{3.27}$$

We have thus proved the following inequality of symmetric operator on the space $C(\square)$: for any $\phi \in C(\square)$

$$\operatorname{Hess} H(\phi) \geq -\frac{1}{4\beta}\Delta.$$

Noting that the inverse of the discrete Laplacian on the space $C(\square)$ is the Green's function, we obtain: for any $\phi, \psi \in C(\square)$,

$$\left(\psi, \operatorname{Hess} H(\phi)^{-1}\psi\right) \leq 4\beta \sum_{x,y\in\square} \sum_{i=1}^{\binom{d}{2}} \psi_i(x) G_{C(\square),i}(x,y)\psi_i(y).$$

For any $i \in \{1, \ldots, \binom{d}{2}\}$ and any $x \in \square$, we can apply the Brascamp-Lieb inequality with the function $F_{x,i}(\phi) := \phi_i(x)$. We obtain

$$\operatorname{var}\left[\phi_i(x)\right] \leq 4\beta G_{C(\square),i}(x,x) \leq C, \tag{3.28}$$

where in the second inequality, we used that the Green's function is bounded uniformly in the box $\square$ and the vertex $x \in \square$. We next upgrade the estimate (3.28) from an upper bound on the variance to an upper bound on exponential moments. To this end, we follow the techniques of [51, Theorem 4.9] and consider the function

$$t \mapsto \log \left\langle \exp\left(t\phi_i(0)\right)\right\rangle_{\mu_{\beta,\square}}.$$

The second derivative of this map is given by the formula

$$\frac{\partial^2}{\partial t^2} \log \left\langle \exp\left(t\phi_i(0)\right)\right\rangle_{\mu_{\beta,\square}} = \operatorname{var}_{\mu_t}\left(\phi(0)\right),$$

where the measure μ_t is defined via density

$$d\mu_t = \frac{1}{Z} \times \exp\left(H(\phi) + t\phi_i(0)\right) d\phi.$$

We then note that the Hessian of the Hamiltonian $\phi \mapsto H(\phi) + t\phi_i(0)$ is the same as the one of H. We may thus apply the Brascamp-Lieb inequality to the measure μ_t. We obtain, for any $t \in [0,1]$,

$$\frac{\partial^2}{\partial t^2} \log \left\langle \exp\left(t\phi_i(0)\right)\right\rangle_{\mu_{\beta,\square}} = \operatorname{var}_{\mu_t}\left(\phi(0)\right) \leq C.$$

Integrating over $t \in [0,1]$ twice and noting that $\langle \phi(0)\rangle_{\mu_{\beta,\square}} = 0$ (by the $\phi \mapsto -\phi$ symmetry of the field) yields the bound

$$\left\langle \exp\left(\phi_i(0)\right)\right\rangle_{\mu_{\beta,\square}} \leq C.$$

Using once again the $\phi \mapsto -\phi$ symmetry of the field, we also have

$$\left\langle \exp\left(-\phi_i(0)\right)\right\rangle_{\mu_{\beta,\square}} \leq C.$$

Combining that the two previous estimates and noting that they hold for any $i \in \{1, \ldots, \binom{d}{2}\}$ completes the proof of (3.21).

We then prove the inequality (3.22). By the Brascamp-Lieb inequality, we have

$$\operatorname{var}_{\mu_{\beta,\square}}\left[Z\right] \leq 4\beta \sum_{x,y\in\square} \sum_{i=1}^{\binom{d}{2}} G_{C(\square),i}(x,y) \left\langle \left(\partial_{x,i} Z\right)\left(\partial_{y,i} Z\right)\right\rangle_{\mu_{\beta,\square}}.$$

We next upgrade the previous inequality from an estimate on the variance to an estimate on exponential moments using the same strategy as before. We first note that

$$\frac{\partial^2}{\partial t^2} \log \left\langle \exp\left(tZ\right)\right\rangle_{\mu_{\beta,\square}} = \operatorname{var}_{\mu_t}\left(Z\right),$$

where the measure μ_t is defined via the density

$$d\mu_t = \operatorname{Const} \times \exp\left(H(\phi) + tZ(\phi)\right) d\phi.$$

We will then apply the Brascamp-Lieb inequality to the measure μ_t. To this end, we need to show that: for any $\phi, \psi \in C(\square)$ and any $t \in [0,1]$,

$$\left(\psi, \operatorname{Hess}\left(H + tZ\right)(\phi)\psi\right) \geq \frac{1}{8\beta}\left(\psi, (-\Delta)\psi\right). \tag{3.29}$$

From (3.27), we see that it is sufficient to prove

$$|(\psi, \operatorname{Hess} Z(\phi)\psi)| \leq \frac{1}{8\beta}(\psi, (-\Delta)\psi).$$

Using the definition $Z := \sum_{q \in \mathcal{Q}_\square} g(\beta, q) \sin(2\pi(\phi, q))$, we can compute the Hessian of the map X by differentiating the sine twice. We obtain the identity

$$(\psi, \operatorname{Hess} Z(\phi)\psi) = -4\pi^2 \sum_{q \in \mathcal{Q}_\square} g(\beta, q) \sin(2\pi(\phi, q))(q, \psi)^2.$$

Using the assumption on the coefficient $g(\beta, q)$, we deduce that

$$|(\psi, \operatorname{Hess} Z(\phi)\psi)| \leq \sum_{q \in \mathcal{Q}_\square} e^{-c\beta \|q\|_1} (n_q, \mathrm{d}^* \psi)^2.$$

The proof is then identical to the proof (3.21) (specifically the term in the right-hand side appears in (3.24)). Applying the Brascamp-Lieb inequality, we obtain, for any $t \in [0,1]$,

$$\frac{\partial^2}{\partial t^2} \log \langle \exp(tZ) \rangle_{\mu_{\beta,\square}} = \operatorname{var}_{\mu_t}(Z) \leq 8\beta \sum_{x,y \in \square} \sum_{i=1}^{\binom{d}{2}} G_{C(\square),i}(x,y) \langle (\partial_{x,i} Z)(\partial_{y,i} Z) \rangle_{\mu_{\beta,\square}}.$$

Integrating over $t \in [0,1]$ twice and noting that $\langle Z \rangle_{\mu_{\beta,\square}} = 0$ (by the $\phi \mapsto -\phi$ symmetry of the field) completes the proof of (3.22). □

3.3. **Thermodynamic limit.** The Brascamp-Lieb inequality allows us to prove the existence of a thermodynamic limit for the measures $\mu_{\beta,\square}$ as $\square \to \infty$. Specifically, by Proposition 3.4 (since all the constants in the statement do not depend on the volume) and a tightness argument, there exists a sequence of boxes $(\square_{L_k})_{k \in \mathbb{N}}$ centered at 0 and of side length L_k such that L_k tends to infinity as k tends to infinity and such that the sequence of measures $\mu_{\beta,\square_{L_k}}$ converges weakly in the space Ω to an infinite-volume, translation-invariant Gibbs measure denoted by μ_β. By taking the limit in the finite volume identity (3.6), and using that $\mu^V_{\beta,\square,0}$ converges to the unique Gibbs state μ^V_β, we see that any possible limit μ_β gives the same contribution to the correlation functions of the Villain model, thus it suffices to study any of such μ_β.

We record below three properties of the measure μ_β, which are direct consequences of Proposition 3.4 and the definition of μ_β:

- There exists a constant $C := C(d,\beta) < \infty$ such that, for any $x \in \mathbb{Z}^d$,

$$\langle \exp(|\phi(x)|) \rangle_{\mu_\beta} \leq C \text{ and } \langle \phi(x) \rangle_{\mu_\beta} = 0.$$

- For any box $L \in \mathbb{N}$, any collection of coefficients $(g(\beta,q))_{q \in \mathcal{Q}_L}$ satisfying $|g(\beta,q)| \leq C \exp(-c\beta^{1/2}\|q\|_1)$, if we denote by $Z := \sum_{q \in \mathcal{Q}_L} g(\beta,q) \sin(2\pi(\phi,q))$, then

$$\langle \exp(Z) \rangle_{\mu_\beta} \leq \exp\left(\sum_{x,y \in \mathbb{Z}^d} \sum_{i=1}^{\binom{d}{2}} G(x,y) \langle (\partial_{x,i} Z)(\partial_{y,i} Z) \rangle_{\mu_\beta} \right). \tag{3.30}$$

Combining with the thermodynamic limit results for the Villain model [25, 57], we are now ready to state the following dual representation in infinite volume. To this end, for $L \in \mathbb{N}$ denote by $\overline{X}_L : \Omega \to \mathbb{R}$ and $\overline{Y}_L : \Omega \to \mathbb{R}$ the two random variables

$$\begin{cases} \overline{X}_L := \sum_{q \in \mathcal{Q}_{\square_L}} z(\beta,q) \sin(2\pi(\phi,q)) \sin\left(2\pi(\sigma_{\{\square_L,0x\}},q)\right) + \sum_{q \in \mathcal{Q}_{\square_L}} z(\beta,q) \cos(2\pi(\phi,q)) \left(\cos\left(2\pi(\sigma_{\{\square_L,0x\}},q)\right) - 1\right), \\ \overline{Y}_L := \sum_{q \in \mathcal{Q}_{\square_L}} z(\beta,q) \sin(2\pi(\phi,q)) \sin\left(2\pi(\sigma_{\{\square_L,0x\}},q)\right). \end{cases}$$

We first prove that the random variables X_L and Y_L converge in $L^2(\mu_\beta)$ as L tends to infinity.

Proposition 3.5. *There exists an inverse temperature $\beta_1 := \beta_1(d) < \infty$ such that for any $\beta \geq \beta_1$, the sequences of random variables X_L and Y_L converge as $L \to \infty$ in $L^2(\mu_\beta)$.*

Proof. We first introduce the two random variables

$$\begin{cases} X_L := \sum_{q \in \mathcal{Q}_{\square_L}} z(\beta,q) \sin\left(2\pi(\phi,q)\right) \sin\left(2\pi(\sigma_{0x},q)\right) + \sum_{q \in \mathcal{Q}_{\square_L}} z(\beta,q) \cos\left(2\pi(\phi,q)\right) \left(\cos\left(2\pi(\sigma_{0x},q)\right) - 1\right), \\ Y_L := \sum_{q \in \mathcal{Q}_{\square_L}} z(\beta,q) \sin\left(2\pi(\phi,q)\right) \sin\left(2\pi(\sigma_{0x},q)\right). \end{cases}$$

and prove that the function Y_L converges in $L^2(\mu_\beta)$. In the argument below, we will denote by C_x a generic and typically large constant depending on the parameters d, β and on the vertex $x \in \mathbb{Z}^d$ (which is fixed through the proof). By the Brascamp-Lieb inequality, we have

$$\mathrm{var}_{\mu_\beta}\left[Y_{2L} - Y_L\right] \leq \sum_{y,z \in \square_{2L}} 4\beta \sum_{i=1}^{\binom{d}{2}} G_{\square,i}(y,z) \left\langle \left(\partial_{y,i}(Y_{2L} - Y_L)\right)\left(\partial_{z,i}(Y_{2L} - Y_L)\right)\right\rangle_{\mu_\beta}.$$

An explicit computation shows

$$\partial_{y,i}(Y_{2L} - Y_L) = \sum_{q \in \mathcal{Q}_{2L} \setminus \mathcal{Q}_{\square_L}} z(\beta,q) \sin\left(2\pi(\phi,q)\right) \sin\left(2\pi(\sigma_{0x},q)\right) q_i(y).$$

By definition of σ_{0x}, the identity $q = \mathrm{d} n_q$ and the bound $|\sin\theta| \leq |\theta|$ for $\theta \in \mathbb{R}$, we have

$$\begin{aligned} |z(\beta,q) q(y) \sin\left(2\pi(\sigma_{0x},q)\right)| &\leq C e^{-c\sqrt{\beta}\|q\|_1} |q(y)| |(\sigma_{0x},q)| \\ &\leq C e^{-c\sqrt{\beta}\|q\|_1} |q(y)| |(\mathrm{d}^*\sigma_{0x}, n_q)| \\ &\leq C e^{-c\sqrt{\beta}\|q\|_1} |q(y)| \left\|\mathrm{d}^*\sigma_{0x}\right\|_{L^2(\mathrm{supp}\, n_q)} \left\|n_q\right\|_{L^2(\mathbb{Z}^d)} \\ &\leq C e^{-c\sqrt{\beta}\|q\|_1} |q(y)| \left\|\nabla G(\cdot,0) - \nabla G(\cdot,x)\right\|_{L^2(\mathrm{supp}\, n_q)} \left\|n_q\right\|_{L^2(\mathbb{Z}^d)}. \end{aligned} \tag{3.31}$$

We next use the bound $|\nabla\nabla G(0,z)| \leq C|z|^{-d}$ on the mixed derivative of the Green's function and obtain

$$\begin{aligned} \left\|\nabla G(\cdot,0) - \nabla G(\cdot,x)\right\|_{L^2(\mathrm{supp}\, n_q)} &\leq \left|\mathrm{supp}\, n_q\right|^{\frac{1}{2}} \sup_{z \in \mathrm{supp}\, n_q} |\nabla G(z,0) - \nabla G(z,x)| \\ &\leq C_x \left|\mathrm{supp}\, n_q\right|^{\frac{1}{2}} \frac{(\mathrm{diam}\, n_q)^d}{|y|^d}. \end{aligned}$$

Combining the two previous displays and reducing the value of the constant c in the exponential to absorb all the terms involving the diameter, support and L^2-norm of the charge n_q, we obtain

$$e^{-c\sqrt{\beta}\|q\|_1} |q(y)| |(\sigma_{0x},q)| \leq \frac{C_x e^{-c\sqrt{\beta}\|q\|_1} |q(y)|}{|y|^d}. \tag{3.32}$$

Thus,

$$\begin{aligned} |\partial_{y,i}(Y_{2L} - Y_L)| &\leq \frac{C_x}{|y|^d} \sum_{q \in \mathcal{Q}_{2L} \setminus \mathcal{Q}_{\square_L}} e^{-c\sqrt{\beta}\|q\|_1} |q(y)| \\ &\leq \frac{C_x}{|y|^d} e^{-\frac{c}{2}\sqrt{\beta}\,\mathrm{dist}(y,\square_{2L} \setminus \square_L)} \sum_{q \in \mathcal{Q}_{2L} \setminus \mathcal{Q}_{\square_L}} e^{-\frac{c}{2}\sqrt{\beta}\|q\|_1} |q(y)| \\ &\leq \frac{C_x}{|y|^d} e^{-\frac{c}{2}\sqrt{\beta}\,\mathrm{dist}(y,\square_{2L} \setminus \square_L)} \sum_{q \in \mathcal{Q}_y} e^{-\frac{c}{2}\sqrt{\beta}\|q\|_1} |q(y)| \\ &\leq \frac{C_x}{|y|^d} e^{-\frac{c}{2}\sqrt{\beta}\,\mathrm{dist}(y,\square_{2L} \setminus \square_L)}. \end{aligned} \tag{3.33}$$

The second inequality relies on the observation that a charge satisfying $q \in \mathcal{Q}_{2L} \setminus \mathcal{Q}_{\square_L}$ and $q(y) \neq 0$ must have a diameter larger than $\mathrm{dist}(y, \square_{2L} \setminus \square_L)$ (and thus $\|q\|_1 \geq \mathrm{dist}(y, \square_{2L} \setminus \square_L)$ since q is integer-valued with a connected support). The third inequality is a consequence of (A.20) of Appendix A (choosing the value $k = 1$, and noting that, by the definition of the L^1-norm, $|q(y)| \leq \|q\|_1$). Consequently

$$\mathrm{var}_{\mu_\beta}\left[Y_{2L} - Y_L\right] \leq C_x \sum_{y,z \in \mathbb{Z}^d} \frac{1}{|y-z|^{d-2}} \frac{e^{-c\sqrt{\beta}\,\mathrm{dist}(y,\square_{2L} \setminus \square_L)}}{|y|^d} \frac{e^{-c\sqrt{\beta}\,\mathrm{dist}(z,\square_{2L} \setminus \square_L)}}{|z|^d}.$$

Summing over the dyadic scales, we obtain

$$\begin{aligned}(3.34)\qquad \sum_{n=1}^{\infty} \mathrm{var}_{\mu_\beta}\left[Y_{2^{n+1}} - Y_{2^n}\right] &\leq C_x |x|^2 \sum_{n=1}^{\infty} \sum_{y,z\in\mathbb{Z}^d} \frac{1}{|y-z|^{d-2}} \frac{e^{-c\sqrt{\beta}\,\mathrm{dist}(y,\Box_{2^{n+1}}\setminus\Box_{2^n})}}{|y|^d} \frac{e^{-c\sqrt{\beta}\,\mathrm{dist}(z,\Box_{2^{n+1}}\setminus\Box_{2^n})}}{|z|^d} \\ &\leq C_x \sum_{y,z\in\mathbb{Z}^d} \frac{1}{|y-z|^{d-2}} \frac{1}{|y|^d} \frac{1}{|z|^d} \\ &\leq C_x.\end{aligned}$$

Using that, for any $L \in \mathbb{N}$, the $\phi \mapsto -\phi$ invariance of the measure μ_β implies $\langle Y_L \rangle_{\mu_\beta} = 0$, we deduce that the sequence of random variables Y_L converges in $L^2(\mu_\beta)$ to a limit that we denote by

$$(3.35)\qquad Y = \sum_{q\in\mathcal{Q}} z(\beta,q) \sin\left(2\pi(\phi,q)\right) \sin\left(2\pi(\sigma_{0x},q)\right).$$

We next prove the convergence of the random variables X_L. By the definition of σ_{0x}, the identity $q = \mathrm{d}n_q$, the estimate $|\nabla\nabla G(y)| \leq C|y|^{-d}$ and a computation similar to the one of (3.31), we have, for each vertex $y \in \mathbb{Z}^d$, and each charge $q \in \mathcal{Q}_y$,

$$(3.36)\qquad \left|z(\beta,q)\left(\cos\left(2\pi(\sigma_{0x},q)\right) - 1\right)\right| \leq C \exp(-c\sqrt{\beta}\|q\|_1)(\sigma_{0x},q)^2 \leq \frac{C_x e^{-c\sqrt{\beta}\|q\|_1}}{|y|^{2d}}.$$

Since the map $x \to |x|^{-2d}$ is summable in $\mathbb{Z}^d$, we deduce that for any field $\phi \in \Omega$,

$$\begin{aligned}(3.37)\quad \sum_{q\in\mathcal{Q}} \left|z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\sigma_{0x},q)\right)-1\right)\right| &\leq \sum_{y\in\mathbb{Z}^d}\sum_{q\in\mathcal{Q}_y} \left|z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\sigma_{0x},q)\right)-1\right)\right| \\ &\leq C_x \sum_{y\in\mathbb{Z}^d} \frac{1}{|y|^{2d}} \sum_{q\in\mathcal{Q}_y} \exp(-c\sqrt{\beta}\|q\|_1) \\ &\leq C_x \sum_{y\in\mathbb{Z}^d} \frac{1}{|y|^{2d}} \\ &\leq C_x.\end{aligned}$$

The previous inequality implies that the sequence of random variables $\sum_{q\in\mathcal{Q}_{\Box_L}} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\sigma_{0x},q)\right)-1\right)$ converges uniformly over all the possible values of the field $\phi\in\Omega$. In particular it converges in $L^\infty(\mu_\beta)$ (and thus in $L^2(\mu_\beta)$). We denote the limit by

$$(3.38)\qquad X = \sum_{q\in\mathcal{Q}} z(\beta,q)\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\sigma_{0x},q)\right) + \sum_{q\in\mathcal{Q}} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\sigma_{0x},q)\right)-1\right).$$

The proof of the convergences of the random variables Y_L and X_L is complete. To complete the proof of Proposition 3.5, it is sufficient to show that

$$(3.39)\qquad \mathrm{var}_{\mu_\beta}\left[Y_L - \overline{Y}_L\right] \underset{L\to\infty}{\longrightarrow} 0 \quad \text{and} \quad \mathrm{var}_{\mu_\beta}\left[X_L - \overline{X}_L\right] \underset{L\to\infty}{\longrightarrow} 0.$$

We only sketch the proof of the first convergence. Using the definition of the maps $\sigma_{\{\Box_L,0x\}}$ and (standard) regularity estimates on the finite-volume Green's functions, we have the bound

$$(3.40)\qquad \left|\nabla\sigma_{\{\Box_L,0x\}}(y)\right| \leq \frac{C_x}{|y|^d}.$$

Using the previous upper bound and the same computation as the one leading to (3.34), we obtain that for any $\varepsilon > 0$ there exists $R_\varepsilon > 0$ such that for any $L \geq R_\varepsilon$

$$\mathrm{var}_{\mu_\beta}\left[\sum_{q\in\mathcal{Q}_{\Box_L}\setminus\mathcal{Q}_{\Box_{R_\varepsilon}}} z(\beta,q)\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\sigma_{\{\Box_L,0x\}},q)\right)\right] \leq \varepsilon$$

and

$$\mathrm{var}_{\mu_\beta}\left[\sum_{q\in\mathcal{Q}_{\Box_L}\setminus\mathcal{Q}_{\Box_{R_\varepsilon}}} z(\beta,q)\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\sigma_{0x},q)\right)\right] \leq \varepsilon$$

Additionally, using the convergence (3.7), we see that

$$\sup_{\phi\in\Omega}\left|\sum_{q\in\mathcal{Q}_{\square_{R_\varepsilon}}} z(\beta,q)\sin\left(2\pi(\phi,q)\right)\left(\sin\left(2\pi(\sigma_{0x},q)\right)-\sin\left(2\pi(\sigma_{\{\square_L,0x\}},q)\right)\right)\right| \underset{L\to\infty}{\longrightarrow} 0. \tag{3.41}$$

A combination of the three previous displays yields the convergence (3.39) for the variance of the random variable $Y_L - \overline{Y}_L$.

Finally, the convergence of the variance of the random variable $X_L - \overline{X}_L$ can be deduced form the one of the variable $Y_L - \overline{Y}_L$ and the following result: as in (3.37), we can use the convergence (3.7) with the bound (3.40) (together with the bound $|\cos\theta - 1| \leq \theta^2/2$ and the summability of the function $y\mapsto |y|^{-2d}$ on $\mathbb{Z}^d$) to obtain

(3.42)

$$\sup_{\phi\in\Omega}\left|\sum_{q\in\mathcal{Q}_{\square_L}} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\sigma_{\{\square_L,0x\}},q)\right)-1\right) - \sum_{q\in\mathcal{Q}} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\sigma_{0x},q)\right)-1\right)\right| \underset{L\to\infty}{\longrightarrow} 0.$$

□

Using the previous proposition, we are able to establish an infinite-volume version of Proposition 3.1.

Proposition 3.6. *There exists an inverse temperature $\beta_1 := \beta_1(d) < \infty$ such that for any $\beta \geq \beta_1$,*

$$\begin{aligned}\left\langle e^{i(\theta(x)-\theta(0))}\right\rangle_{\mu_\beta^V} &\exp\left(\frac{1}{2\beta}\left(G(0,0)-2G(0,x)\right)\right) \\ &= \left\langle \exp\left(\sum_{q\in\mathcal{Q}} z(\beta,q)\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\sigma_{0x},q)\right) + \sum_{q\in\mathcal{Q}} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\sigma_{0x},q)\right)-1\right)\right)\right\rangle_{\mu_\beta}\end{aligned} \tag{3.43}$$

and

$$\begin{aligned}\left\langle e^{i(\theta(x)+\theta(0))}\right\rangle_{\mu_\beta^V} &\exp\left(\frac{1}{2\beta}G(0,x)\right) \\ &= \left\langle \exp\left(\sum_{q\in\mathcal{Q}} z(\beta,q)\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\overline{\sigma}_{0x},q)\right) + \sum_{q\in\mathcal{Q}} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\overline{\sigma}_{0x},q)\right)-1\right)\right)\right\rangle_{\mu_\beta}.\end{aligned} \tag{3.44}$$

Proof. We give the proof of (3.43) below, (3.44) follows from the same argument. By [25, 57], there exists a thermodynamic limit for the Villain model, denoted by μ_β^V such that $\left\langle e^{i(\theta(x)-\theta(0))}\right\rangle_{\mu^V_{\beta,\square_{L_k}}} \to \left\langle e^{i(\theta(x)-\theta(0))}\right\rangle_{\mu_\beta^V}$ as $k\to\infty$. We also have that the finite-volume Green's function $G_{\square_{L_k}}(\cdot,0)$ converges to $G(\cdot,0)$ as k tends to infinity.

In the argument below, we will (still) denote by C_x a generic and typically large constant depending on the parameters d,β and on the vertex $x\in\mathbb{Z}^d$ (which is fixed through the proof). By Proposition 3.1, it is enough to show the convergence

$$\left\langle \exp\left(\overline{X}_{L_k}\right)\right\rangle_{\mu_{\beta,\square_{L_k}}} \underset{k\to\infty}{\longrightarrow} \left\langle \exp\left(X\right)\right\rangle_{\mu_\beta}. \tag{3.45}$$

We first prove the (simpler) convergence

$$\left\langle \exp\left(X_{L_k}\right)\right\rangle_{\mu_{\beta,\square_{L_k}}} \underset{k\to\infty}{\longrightarrow} \left\langle \exp\left(X\right)\right\rangle_{\mu_\beta}. \tag{3.46}$$

We first fix $k, R\in\mathbb{N}$ satisfying $|x| \ll R \ll L_k$, and write

$$\left\langle \exp(X_{L_k})\right\rangle_{\mu_{\beta,\square_{L_k}}} = \left\langle \exp(X_R)\right\rangle_{\mu_{\beta,\square_{L_k}}} + \left\langle \exp(X_R)\left(\exp(X_{L_k}-X_R)-1\right)\right\rangle_{\mu_{\beta,\square_{L_k}}}$$

and

$$\left\langle \exp(X)\right\rangle_{\mu_\beta} = \left\langle \exp(X_R)\right\rangle_{\mu_\beta} + \left\langle \exp(X_R)\left(\exp(X-X_R)-1\right)\right\rangle_{\mu_\beta}.$$

We then note that, for any $R>0$, the random variable X_R is a bounded Lipschitz function (with a large Lipschitz constant depending on R) which only depends on the values of the field ϕ inside the box $[-R,R]^d$. Thus,

$$\left\langle \exp(X_R)\right\rangle_{\mu_{\beta,\square_{L_k}}} \underset{k\to\infty}{\longrightarrow} \left\langle \exp(X_R)\right\rangle_{\mu_\beta}.$$

We next apply the Hölder inequality and obtain

$$\left\langle \exp(X_R)\left(\exp(X_{L_k}-X_R)-1\right)\right\rangle_{\mu_{\beta,\square_{L_k}}} \leq \left\langle \exp(2X_R)\right\rangle_{\mu_{\beta,\square_{L_k}}}^{1/2} \left\langle \left(\exp(X_{L_k}-X_R)-1\right)^2\right\rangle_{\mu_{\beta,\square_{L_k}}}^{1/2}. \tag{3.47}$$

We estimate the two terms in the right side. For the first one, let us first observe that, by (3.37), there exists a constant $C_x := C_x(d,\beta,x) < \infty$ such that

$$X_R \leq Y_R + C_x. \tag{3.48}$$

Combining the previous estimate with the Brascamp-Lieb inequality (Proposition 3.4) and obtain, for some constant $C_x := C_x(d,\beta,x) < \infty$,

$$\left\langle \exp(2X_R)\right\rangle_{\mu_{\beta,\square_{L_k}}} \leq C \exp\left(\sum_{x,y\in\square^\circ} \sum_{i=1}^{\binom{d}{2}} G_{C(\square),i}(x,y) \left\langle (\partial_{x,i} Y_R)(\partial_{y,i} Y_R)\right\rangle_{\mu_{\beta,\square_{L_k}}}\right).$$

Using the estimates (3.30) (with $Z = Y_R$), (3.48), and an explicit computation, we obtain the upper bound

$$\left\langle \exp(2X_R)\right\rangle_{\mu_{\beta,\square_{L_k}}} \leq C_x. \tag{3.49}$$

There remains to estimate the second term in the right side of (3.47). We claim that

$$\left\langle \left(\exp(X_{L_k}-X_R)-1\right)^2\right\rangle_{\mu_{\beta,\square_{L_k}}} \leq \frac{C_x}{R^{\frac{d}{2}-1}}. \tag{3.50}$$

Using (3.36) and the same computation as in (3.37), we see that, for any $L > R$, and any field $\phi \in \Omega$,

$$\left| \sum_{q\in\mathcal{Q}_L} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\sigma_{0x},q)\right)-1\right) - \sum_{q\in\mathcal{Q}_R} z(\beta,q)\cos\left(2\pi(\phi,q)\right)\left(\cos\left(2\pi(\sigma_{0x},q)\right)-1\right)\right| \leq \frac{C_x}{R^d}.$$

This result implies that to prove the estimate (3.50), it is sufficient to show

$$\left\langle \left(\exp(Y_{L_k}-Y_R)-1\right)^2\right\rangle_{\mu_{\beta,\square_{L_k}}} \leq \frac{C_x}{R^{\frac{d}{2}-1}}. \tag{3.51}$$

We can Taylor expand the left side of (3.51) and use the $\phi \to -\phi$ symmetry of the field to obtain

$$LHS = \left\langle \left(\sum_{k\geq 1} \frac{1}{k!}(Y_{L_k}-Y_R)^k\right)^2\right\rangle_{\mu_{\beta,\square_{L_k}}} = \sum_{l\geq 1}\left\langle \sum_{j=1}^{2l} \frac{1}{j!(2l-j)!}(Y_{L_k}-Y_R)^{2l}\right\rangle_{\mu_{\beta,\square_{L_k}}}.$$

We then apply the exponential Brascamp-Lieb inequality to obtain, for any $C \geq 1$ and β chosen sufficiently large (depending on C),

$$\begin{aligned}\left\langle (Y_{L_k}-Y_R)^{2l}\right\rangle_{\mu_{\beta,\square_{L_k}}} &\leq \frac{(2l)!}{C^{2l}}\left(\left\langle e^{C(Y_{L_k}-Y_R)} + e^{-C(Y_{L_k}-Y_R)}\right\rangle_{\mu_{\beta,\square_{L_k}}} - 2\right) \\ &\leq \frac{(2l)!}{C^{2l}}\left[\exp\left(2C \sum_{y,z\in\square_{L_k}} \sum_{i=1}^{\binom{d}{2}} G_{C(\square),i}(y,z)\left\langle (\partial_{y,i}(Y_{L_k}-Y_R))(\partial_{z,i}(Y_{L_k}-Y_R))\right\rangle_{\mu_{\beta,\square_{L_k}}}\right) - 1\right].\end{aligned}$$

Summing over l, and choosing C large enough (universally), we obtain

$$\left\langle \left(\exp(Y_{L_k}-Y_R)-1\right)^2\right\rangle_{\mu_{\beta,\square_{L_k}}} \leq C\left[\exp\left(2C \sum_{y,z\in\square_{L_k}} \sum_{i=1}^{\binom{d}{2}} G_{C(\square),i}(y,z)\left\langle (\partial_{y,i}(Y_{L_k}-Y_R))(\partial_{z,i}(Y_{L_k}-Y_R))\right\rangle_{\mu_{\beta,\square_{L_k}}}\right) - 1\right]. \tag{3.52}$$

We claim that, the term in the right side of (3.52) is bounded by $C|x|R^{2-d}$. Using the same computation as in (3.33), we have

$$\left|\partial_{y,i}(Y_{L_k}-Y_R)\right| \leq \frac{C_x e^{-\frac{c}{2}\sqrt{\beta}\,\mathrm{dist}(y,\square_{L_k}\setminus\square_R)}}{|y|^d}.$$

Using the estimate (3.32) and the bound $G_{C(\square),i}(y,z) \leq \frac{C}{|y-z|^{d-2}}$, we obtain

$$\begin{aligned}
&\sum_{y,z \in \square_{L_k}} \sum_{i=1}^{\binom{d}{2}} G_{C(\square),i}(y,z) \left\langle \left(\partial_{y,i}(Y_{L_k} - Y_R)\right) \left(\partial_{z,i}(Y_{L_k} - Y_R)\right) \right\rangle_{\mu_{\beta,\square_{L_k}}} \\
&\qquad \leq C_x \sum_{y,z \in \square_{L_k} \setminus \square_R} \frac{1}{|y-z|^{d-2}} \frac{e^{-\frac{c}{2}\sqrt{\beta}\, \mathrm{dist}(y, \square_{L_k} \setminus \square_R)}}{|y|^d} \frac{e^{-\frac{c}{2}\sqrt{\beta}\, \mathrm{dist}(z, \square_{L_k} \setminus \square_R)}}{|z|^d} \\
&\qquad \leq \frac{C_x}{R^{d-2}},
\end{aligned}$$

which, together with (3.47), (3.49), and (3.52), implies

$$\left\langle \exp(X_R) \exp(X - X_R) \right\rangle_{\mu_\beta} \leq \frac{C_x}{R^{\frac{d}{2}-1}}.$$

The same computation yields

$$\left\langle \exp(X_R) \left(\exp(X - X_R) - 1\right) \right\rangle_{\mu_\beta} \leq \frac{C_x}{R^{\frac{d}{2}-1}}.$$

So that the proof of (3.46) is complete. To prove (3.45), it is thus sufficient to show

$$\left\langle \exp\left(\overline{X}_{L_k}\right) \right\rangle_{\mu_{\beta,\square_{L_k}}} - \left\langle \exp\left(X_{L_k}\right) \right\rangle_{\mu_{\beta,\square_{L_k}}} \underset{k \to \infty}{\longrightarrow} 0. \tag{3.53}$$

This is a consequence of the convergences (3.41) and (3.41), the bound (3.51) and the bound: for any $R \geq 0$ and any $L \geq R$,

$$\left\| \exp\left(\sum_{q \in \mathcal{Q}_L \setminus \mathcal{Q}_R} z(\beta,q) \sin\left(2\pi(\phi,q)\right) \sin\left(2\pi(\sigma_{\{\square_L, 0x\}}, q)\right) \right) - 1 \right\|_{L^2(\mu_{\beta,\square_L})}^2 \leq \frac{C_x}{R^{\frac{d}{2}-1}}.$$

The proof of the previous inequality is identical to the proof of (3.51) (the only difference is that the bound (3.40) needs to be used instead of the decay estimates on the Green's function).

□

3.4. **The Helffer-Sjöstrand representation.** Proposition 3.6 shows that, in order to understand the asymptotic behavior as x tends to infinity of the two point function, it is sufficient to understand the behavior of the expectation of the random variable $\exp(X)$ under the measure μ_β as x tends to infinity.

The Gibbs measure μ_β is a specific example of a model of stochastic interface model extensively studied in the literature called the $\nabla\phi$ model [51]. In particular, following the ideas and techniques of [67, 88, 81, 54], the large-scale behavior of the $\nabla\phi$ model can be understood by studying the large-scale behavior of an infinite dimensional PDE called the Helffer-Sjöstrand equation [67, 81]. In this section, we adapt the tools developed by Helffer-Sjöstrand and Naddaf and Spencer [81] to our framework, and introduce the Helffer-Sjöstrand PDE associated with the measure μ_β.

Specifically, in Sections 3.4.1 and 3.4.2, we introduce the Helffer-Sjöstrand PDE, present two equivalent approaches to solve this PDE: the first one is based on variational techniques of [81], the second one is based on a dynamical interpretation of the equation and is the one of [54]. We then show, following [81], how its solutions can be used to identify the covariance of general functionals of the field ϕ distributed according to the measure μ_β. In Section 3.4.4, we introduce the Green's matrix associated with the Helffer-Sjöstrand operator and state a quantitative homogenization theorem for this map. This result is a crucial step in the proof of Theorem 1, and its proof occupies a large part of this article: it is the subject of Sections 6 and 7, where we combine the ideas of [81, 54] with the recent development in quantitative stochastic homogenization of [8, 9].

3.4.1. *The Witten Laplacian.* Following the techniques of [52, 81, 54], we know that the measure μ_β is stationary, ergodic and reversible with respect to the Langevin dynamics defined as follows. We let $\left\{B_t(x) : t \geq 0,\, x \in \mathbb{Z}^d\right\}$ is a collection of independent Brownian motions valued in $\mathbb{R}^{\binom{d}{2}}$ and let $\phi : [0,\infty] \times \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}}$ be the solution of the system of stochastic differential equation: for $t \geq 0$ and $x \in \mathbb{Z}^d$,

$$d\phi_t(x) = -\frac{1}{2\beta}\Delta\phi_t(x) + \sum_{n\geq 0} \frac{1}{2\beta} \frac{1}{\beta^{n/2}} (-\Delta)^{n+1} \phi_t(x) - \sum_{q \in \mathcal{Q}} 2\pi z(\beta,q) q(x) \sin\left(2\pi\left(q,\phi\right)\right) + \sqrt{2}\, dB_t(x). \tag{3.54}$$

We refer to [52] for the justification of this property and a proof of the solvability of the Langevin dynamics (3.54). Following the idea of [81], one observes that the Langevin dynamics is a Markov process whose infinitesimal generator is the operator Δ_ϕ defined on the set of (real-valued) functions $F \in C_c^\infty(\Omega)$ by the formula: for any $\phi \in \Omega$,

$$(3.55)\quad \Delta_\phi F(\phi) := \sum_{x \in \Box^\circ} \partial_x^2 F(\phi) - \sum_{x \in \mathbb{Z}^d} \left[\frac{1}{2\beta} \Delta \phi(x) - \sum_{n \geq 1} \frac{1}{2\beta} \frac{1}{\beta^{n/2}} (-\Delta)^{n+1} \phi(x) - \sum_{q \in \mathcal{Q}} 2\pi z(\beta, q) q(x) \sin\left(2\pi \left(q, \phi\right)\right) \right] \partial_x F(\phi),$$

where the notation ∂_x^2 means $\sum_{i=1}^{\binom{d}{2}} \partial_{x,i}^2$, and we implicitly take the scalar product between the two terms in the right side of (3.55). The operator Δ_ϕ is thus symmetric with respect to the measure $\mu_{\beta,L}$, and one has the identities

$$\langle F \Delta_\phi G \rangle_{\mu_\beta} = \langle G \Delta_\phi F \rangle_{\mu_\beta} = - \sum_{x \in \mathbb{Z}^d} \langle \partial_x F, \partial_x G \rangle_{\mu_\beta}, \quad \forall F, G \in C_c^\infty(\Omega).$$

3.4.2. *Helffer-Sjöstrand operator.* In this section, we introduce the Helffer-Sjöstrand operator. This operator is defined in (3.57) and acts on function defined on $\mathbb{Z}^d \times \Omega$ and valued in $\mathbb{R}^{\binom{d}{2}}$. Its definition requires to introduce a few spaces and definitions. We first introduce the space of smooth and compactly supported functions defined on $\mathbb{Z}^d \times \Omega$

$$C_c^\infty(\mathbb{Z}^d \times \Omega) := \left\{ F : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}} \, : \, \forall x \in \mathbb{Z}^d,\, F(x, \cdot) \in C_c^\infty(\Omega) \text{ and } F(x,\cdot) = 0 \text{ for all but finitely many } x \in \mathbb{Z}^d \right\}.$$

We then extend the domain of the Witten Laplacian $-\Delta_\phi$ to the functions of $C_c^\infty(\mathbb{Z}^d \times \Omega)$ as follows: for any $F \in C_c^\infty(\mathbb{Z}^d \times \Omega)$, any $(y, \phi) \in \mathbb{Z}^d \times \Omega$,

$$(3.56)\quad \Delta_\phi F(y,\phi) := \sum_{x \in \mathbb{Z}^d} \partial_x^2 F(y,\phi) - \sum_{x \in \mathbb{Z}^d} \left[\frac{1}{2\beta} \Delta \phi(x) - \sum_{n \geq 1} \frac{1}{2\beta} \frac{1}{\beta^{n/2}} (-\Delta)^{n+1} \phi(x) - \sum_{q \in \mathcal{Q}} 2\pi z(\beta, q) q(x) \sin\left(2\pi \left(q, \phi\right)\right) \right] \partial_x F(y,\phi),$$

where all the partial derivative are with respect to the field ϕ (for a fixed point $y \in \mathbb{Z}^d$). We next extend the definition of the discrete Laplacian Δ to functions of $C_c^\infty(\mathbb{Z}^d \times \Omega)$ by setting: for any $(y, \phi) \in \mathbb{Z}^d \times \Omega$,

$$\Delta F(y, \phi) := \sum_{x \sim y} (F(x, \phi) - F(y, \phi)).$$

We similarly define the iteration of the Laplacian $(-\Delta)^k$ by iterating the previous definition. For $q \in \mathcal{Q}$, and $\phi \in \Omega$, we define the coefficient

$$\mathbf{a}_q(\phi) := 4\pi^2 z\left(\beta, q\right) \cos\left(2\pi \left(\phi, q\right)\right).$$

Given a function $F \in C_c^\infty(\mathbb{Z}^d \times \Omega)$ and $\phi \in \Omega$, we introduce the notation

$$\nabla_q F(\phi) := (q, F(\cdot, \phi)).$$

We finally combine the two previous definitions, and introduce the operator

$$\nabla_q^* \cdot \mathbf{a}_q \nabla_q F(\phi, x) = 4\pi^2 z\left(\beta, q\right) \cos\left(2\pi \left(\phi, q\right)\right) \left(F(\cdot, \phi), q\right) q(x).$$

The notation is motivated by the following symmetry property satisfied by the operator $\nabla_q^* \cdot \mathbf{a}_q \nabla_q$: for any $F, G \in C_c^\infty(\mathbb{Z}^d \times \Omega)$,

$$\begin{aligned} \sum_{x \in \mathbb{Z}^d} \left\langle \nabla_q^* \cdot \mathbf{a}_q \nabla_q F(\cdot, x) G(\cdot, x) \right\rangle_{\mu_\beta} &= \sum_{x \in \mathbb{Z}^d} \left\langle \nabla_q^* \cdot \mathbf{a}_q \nabla_q G(\phi, x) F(\phi, x) \right\rangle_{\mu_\beta} \\ &= \left\langle \mathbf{a}_q \left(F, q\right) \left(G, q\right) \right\rangle_{\mu_\beta} \\ &= \left\langle \mathbf{a}_q \nabla_q F \nabla_q G \right\rangle_{\mu_\beta}. \end{aligned}$$

Additionally, due to the assumption $\mathrm{d}q = 0$, the function $\nabla_q^* \cdot \mathbf{a}_q \nabla_q F$ depends only on the discrete gradient ∇F. Equipped with these definitions, we introduce the Helffer-Sjöstrand operator acting on functions $F : \Omega \times \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}}$

$$\mathcal{L} := -\Delta_\phi - \frac{1}{2\beta}\Delta + \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1} + \sum_{q\in\mathcal{Q}}\nabla_q^*\cdot\mathbf{a}_q\nabla_q. \tag{3.57}$$

The following definition introduce a notion of weak solutions for the Helffer-Sjöstrand operator.

Definition 3.7 (Solution of the Helffer-Sjöstrand equation)**.** Let $f : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ be such that $f \in L^2(\mathbb{Z}^d, \mu_\beta)$. A function $u : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ is called a weak solution of the Helffer-Sjöstrand equation

$$\mathcal{L}u = f \text{ in } \Omega \times \mathbb{Z}^d,$$

if, for any function $F \in C_c^\infty(\mathbb{Z}^d \times \Omega)$, one has the identity

$$\sum_{x\in\mathbb{Z}^d}\langle u(x,\cdot), \mathcal{L}F(x,\cdot)\rangle_{\mu_\beta} = \sum_{x\in\mathbb{Z}^d}\langle f(x,\cdot), F(x,\cdot)\rangle_{\mu_\beta}.$$

The next proposition establishes the solvability of the Helffer-Sjöstrand using the variational approach used by Naddaf-Spencer [81]. We recall the definition of the space $\dot H^1(\mathbb{Z}^d,\mu_\beta)$ introduced in (2.7) of Section 2.

Proposition 3.8 (Variational solvability)**.** *For any $f : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{d\times\binom{d}{2}}$ satisfying $f \in L^2(\mathbb{Z}^d, \mu_\beta)$, there exists a unique weak solution $u : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ in the space $\dot H^1(\mathbb{Z}^d, \mu_\beta)$ of the equation*

$$\mathcal{L}u = \nabla\cdot f \quad \text{in } \mathbb{Z}^d\times\Omega, \tag{3.58}$$

which satisfies, for some $C(d,\beta) < \infty$,

$$\sup_{x\in\mathbb{Z}^d}\|u(x,\cdot)\|^2_{L^2(\mu_\beta)} + \sum_{x\in\mathbb{Z}^d}\|\partial_x u\|^2_{L^2(\mathbb{Z}^d,\mu_\beta)} + \|\nabla u\|^2_{L^2(\mathbb{Z}^d,\mu_\beta)} \leq C\|f\|^2_{L^2(\mathbb{Z}^d,\mu_\beta)}. \tag{3.59}$$

Remark 3.9. We require that the right-hand side (3.58) is in divergence form. This assumption simplifies the proof but is not strictly necessary. Indeed, using the Gagliardo-Nirenberg-Sobolev inequality, one could prove the existence and uniqueness of variational solutions of the Helffer-Sjöstrand equation $\mathcal{L}u = f$ if the function $f : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ satisfies

$$\left(\sum_{x\in\mathbb{Z}^d}|f(x,\cdot)|^{2d/(d-2)}\right)^{\frac{d-2}{2d}} \in L^2(\mu_\beta). \tag{3.60}$$

Proof. Using that the space $C_c^\infty(\mathbb{Z}^d\times\Omega)$ is dense in $\dot H^1(\mathbb{Z}^d,\mu_\beta)$, we see that a function $u \in \dot H^1(\mathbb{Z}^d,\mu_\beta)$ is a solution of (3.58) if and only if

$$\begin{aligned}\sum_{x,y\in\mathbb{Z}^d}\langle(\partial_y u(x,\cdot))(\partial_y w(x,\cdot))\rangle_{\mu_\beta} &+ \frac{1}{2\beta}\sum_{x\in\mathbb{Z}^d}\langle\nabla u(x,\cdot)\nabla w(x,\cdot)\rangle_{\mu_\beta} \\ &+ \frac{1}{2\beta}\sum_{n\geq1}\frac{1}{\beta^{\frac{n}{2}}}\sum_{x\in\mathbb{Z}^d}\left\langle\nabla^{n+1}u(x,\cdot),\nabla^{n+1}w(x,\cdot)\right\rangle_{\mu_\beta} + \sum_{q\in\mathcal{Q}}\langle\nabla_q u\cdot\mathbf{a}_q\nabla_q w\rangle_{\mu_\beta} \\ &= -\sum_{x\in\mathbb{Z}^d}\langle f(x,\cdot)\nabla w(x,\cdot)\rangle_{\mu_\beta}, \quad \forall w\in\dot H^1(\mathbb{Z}^d,\mu_\beta).\end{aligned} \tag{3.61}$$

As in the proof of Proposition 3.4, we may use the estimate $|\mathbf{a}_q| \leq Ce^{-c\beta\|q\|_1}$ to show that, if β is sufficiently large, the bilinear form on the left side of the previous display is coercive with respect to the $\dot H^1(\mathbb{Z}^d,\mu_\beta)$-norm. The Lax-Milgram Theorem therefore yields the existence of a unique solution $u \in \dot H^1(\mathbb{Z}^d,\mu_\beta)$. Applying the Gagliardo-Nirenberg-Sobolev inequality (for a fixed field $\phi\in\Omega$) yields

$$\sup_{x\in\mathbb{Z}^d}|u(x,\phi)|^2 \leq \left(\sum_{y\in\mathbb{Z}^d}|u(y,\phi)|^{2d/(d-2)}\right)^{(d-2)/d} \leq \sum_{y\in\mathbb{Z}^d}|\nabla u(y,\phi)|^2. \tag{3.62}$$

Integrating over the measure μ_β completes the proof of Proposition 3.8. □

As it has been observed in the literature [81, 54], there exists a dynamical representation for the solution u of the Helffer-Sjöstrand PDE $\mathcal{L}u = \nabla \cdot f$. The formula is stated in Proposition 3.12, and we will use it in this article to obtain upper bounds on the solution u. In order to state the result, we introduce a few additional definitions. Given a field $\phi \in \Omega$, we consider the solution of the Langevin dynamics started from ϕ at time $t = 0$, that is,

$$(3.63)\qquad \begin{cases} d\phi_t(x) = \frac{1}{2\beta}\Delta\phi_t(x)dt - \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1}\phi_t(x)dt + \sum_{q\in\mathcal{Q}}\left(\nabla_q^* \cdot \mathbf{a}_q(\phi_t)\nabla_q\phi_t\right)(x)dt + \sqrt{2}dB_t(x), \\ \phi_0(x) = \phi(x). \end{cases}$$

We denote by $\mathbb{P}_\phi$ the law of the dynamics $(\phi_t)_{t\geq 0}$ starting from ϕ and by $\mathbb{E}_\phi$ the expectation with respect to the measure $\mathbb{P}_\phi$. The solvability of the SDE (3.63) is guaranteed for μ_β-almost every $\phi \in \Omega$ by the arguments of [54, Section 2.1.3] or [52, Section 2.2].

For $y \in \mathbb{Z}^d$, we define the Dirac mass $\delta_y : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}\times\binom{d}{2}}$ to be the diagonal matrix

$$\delta_y(x) := \left(\mathbb{1}_{\{x=y\}} \cdot \mathbb{1}_{\{i=j\}}\right)_{1\leq i,j\leq \binom{d}{2}}.$$

For any fixed realization of the dynamics $\left\{\phi_t(x) : t \geq 0, x \in \mathbb{Z}^d\right\}$, we let $P^\phi : [0,\infty] \times \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}} \times \mathbb{R}^{\binom{d}{2}}$ be the fundamental solution (also referred to as heat kernel) of the parabolic system of equations

$$(3.64)\qquad \begin{cases} \partial_t P^{\phi_\cdot}(\cdot,\cdot;y) - \frac{1}{2\beta}\Delta P^{\phi_\cdot}(\cdot,\cdot;y) + \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1}P^{\phi_\cdot}(\cdot,\cdot;y) + \sum_{q\in\mathcal{Q}}\nabla_q^*\cdot\mathbf{a}_q(\phi_t)\nabla_q P^{\phi_\cdot}(\cdot,\cdot;y) = 0 \text{ in } [0,\infty]\times\mathbb{Z}^d, \\ P^{\phi_\cdot}(0,\cdot,y) = \delta_y \text{ in } \mathbb{Z}^d. \end{cases}$$

To be more specific, we follow the standard technique to define the fundamental solution of a system of parabolic equations: For any fixed column in the matrix δ_y, we solve the system (3.64) with this specific column and obtain a function valued in the space $\mathbb{R}^{\binom{d}{2}}$. We then use the $\binom{d}{2}$ solutions obtained this way to define the matrix valued function P^ϕ.

There are two important properties about the fundamental solution $P^{\phi_\cdot}$. First, the bound $|\mathbf{a}_q| \leq Ce^{-c\beta\|q\|_1}$ shows that, for β large enough, the system (3.64) is a small perturbation of the heat equation (or equivalently has a small ellipticity contrast). This observation implies that the system of equations (3.64) is in the range of applicability of the Schauder regularity theory. This is the subject of Section 5, where we adapt the arguments of the Schauder regularity to the system (3.64) and obtain the bounds on the heat kernel $P^{\phi_\cdot}$, its gradient and mixed derivative collected in the following proposition.

Proposition 3.10 (Nash-Aronson estimate and regularity for the heat kernel)**.** *There exists an inverse temperature $\beta_1 := \beta_1(d) < \infty$ such that for any $\beta \geq \beta_1$, there exists a constant $C := C(d,\beta) < \infty$ such that, for any realization of the dynamics $(\phi_t)_{t\geq 0}$, any $(t,x,y) \in [1,\infty] \times \mathbb{Z}^d \times \mathbb{Z}^d$,*

$$\left|P^{\phi_\cdot}(t,x;y)\right| \leq \frac{C}{t^{d/2}}\exp\left(-\frac{|x-y|}{Ct}\right).$$

Additionally, for any regularity exponent $\varepsilon > 0$, there exists an inverse temperature $\beta_1(d,\varepsilon) < \infty$ such that, for any $\beta \geq \beta_1$,

$$\left|\nabla_x P^{\phi_\cdot}(t,x;y)\right| \leq \frac{C}{t^{d/2+1/2-\varepsilon}}\exp\left(-\frac{|x-y|}{Ct}\right) \quad \text{and} \quad \left|\nabla_x\nabla_y P^{\phi_\cdot}(t,x;y)\right| \leq \frac{C}{t^{d/2+1-\varepsilon}}\exp\left(-\frac{|x-y|}{Ct}\right).$$

Remark 3.11. The bounds on the coefficients $\mathbf{a}_q$ show that the ellipticity contrast of the system (3.64) does not depend on the realization of the dynamics $(\phi_t)_{t\geq 0}$. A consequence of this observation is that the Schauder regularity theory applies uniformly in the realization of the dynamics, and thus the upper bounds of Proposition 3.10 are uniform over the dynamics $(\phi_t)_{t\geq 0}$.

The proof of these properties can be found in Proposition 5.7 of Section 5. The second important property of the heat kernel $P^{\phi_\cdot}$ is that, as observed in [81, Section 2.2.2] and [54, Section 3], it is related to the solutions of the Helffer-Sjöstrand equation as explained below.

Proposition 3.12 (Dynamical solvability of the Helffer-Sjöstrand equation [81, 54])**.** *Fix $f : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{d\times\binom{d}{2}}$ such that $f \in L^2(\mathbb{Z}^d, \mu_\beta)$ and let $u \in \dot{H}^1(\mathbb{Z}^d, \mu_\beta)$ be the solution of the Helffer-Sjöstrand equation $\mathcal{L}u = \nabla \cdot f$*

defined in Proposition 3.8. Then, one has the identity

$$u(x,\phi) = \int_0^\infty \sum_{y\in\mathbb{Z}^d} \mathbb{E}_\phi\left[f(y,\phi_t)\nabla_y P^{\phi_\cdot}(t,y;x)\right] dt. \tag{3.65}$$

The rigorous justification of the formula (3.65) requires to use tools from spectral theory. The argument in the case of the dual Villain model is identical to the one presented for the uniformly elliptic $\nabla\phi$ model in the articles of Naddaf and Spencer [81, Section 2.2.2] and Giacomin, Olla and Spohn [54, Section 3].

3.4.3. *The Helffer-Sjöstrand representation formula.* The main reason to introduce the Helffer-Sjöstrand operator is that it can be used to compute the covariance of general functional of the field ϕ through the Helffer-Sjöstrand representation formula. The result was initially introduced by Helffer-Sjöstrand [67] and Naddaf and Spencer [81, (1.10)] and is stated below.

Proposition 3.13 (Helffer-Sjöstrand representation formula [67, 81, 54])**.** *Consider two functions $F, G \in H^1(\mu_\beta)$ and assume that there exist $f, g : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{d\times\binom{d}{2}}$ satisfying $f, g \in L^2\left(\mathbb{Z}^d, \mu_\beta\right)$ and such that*

$$\partial_x F = \nabla\cdot f(x,\cdot) \text{ and } \partial_x G = \nabla\cdot g(x,\cdot). \tag{3.66}$$

Let $u \in \dot H^1(\mathbb{Z}^d,\mu_\beta)$ be the solution of the Helffer-Sjöstrand equation $\mathcal{L}u = \nabla\cdot f$. Then one has the identity

$$\mathrm{cov}_{\mu_\beta}\left[F,G\right] = \sum_{x\in\mathbb{Z}^d} \left\langle g(x,\cdot)\nabla u(x,\cdot)\right\rangle_{\mu_\beta}. \tag{3.67}$$

An example of function $F \in H^1(\mu_\beta)$ satisfying (3.66) is the function $F(\phi) := \phi(0) - \phi(x)$ for any $x \in \mathbb{Z}^d$. For any charge $q \in \mathcal{Q}$, the neutrality condition $\mathrm{d}q = 0$ ensures that the function $F_q(\phi) := (q,\phi)$ satisfies (3.66). In general, any reasonable function which depends only on the discrete gradient of the field satisfy this condition. In the rest of this article, we will apply it to general functional of the field such as the random variables X and Y defined in (3.38) and (3.35), which, still due to the neutrality condition $\mathrm{d}q = 0$, satisfy (3.66).

The proof of this result for the $\nabla\phi$ model can be found in [81, (1.10)] and [54, Proposition 3.1]. The proof for the measure μ_β follows from the same arguments.

3.4.4. *The Green's matrix.* In this section, we introduce the Green's matrix associated with the Helffer-Sjöstrand operator $\mathcal{L}$ and state some of its main properties regarding existence, decay and homogenization.

Proposition 3.14. *For any function $\mathbf{f} : \Omega \to \mathbb{R}$ satisfying $\mathbf{f} \in L^2(\mu_\beta)$ and any $y \in \mathbb{Z}^d$, there exists a unique variational solution $\mathcal{G}_{\mathbf{f}}(\cdot;y) : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}\times\binom{d}{2}}$ of the Helffer-Sjöstrand equation*

$$\mathcal{L}\mathcal{G}_{\mathbf{f}}(\cdot;y) = \mathbf{f}\delta_y \text{ in } \mathbb{Z}^d\times\Omega.$$

The map $\mathcal{G}_{\mathbf{f}}(\cdot;y)$ is the fundamental solution of the operator $\mathcal{L}$. As it was the case for the heat kernel $P^{\phi_\cdot}$, since the operator $\mathcal{L}$ is an elliptic system, the fundamental solution takes its values in the set of matrices of size $\binom{d}{2}\times\binom{d}{2}$. We will refer to it as *the Green's matrix*.

The Green's function can be used to decompose general solutions the Helffer-Sjöstrand equation. If we let u be the solution of the equation (3.58) and assume that $f \in L^2(\mathbb{Z}^d,\mu_\beta)$ takes the specific form $f(y,\phi) = \mathbf{f}(\phi)g(y)$ for some $\mathbf{f} \in L^2(\mu_\beta)$ and $g \in L^2(\mathbb{Z}^d)$, then we have

$$u(x,\phi) = \sum_{y\in\mathbb{Z}^d} \nabla_y \mathcal{G}_{\mathbf{f}}(x,\phi;y)g(y).$$

Proof. Using the Gagliardo-Nirenberg-Sobolev inequality, and specifically the inequality (3.62), we have the upper bound

$$\forall u \in \dot H^1(\mathbb{Z}^d,\mu_\beta), \qquad \|u(y,\cdot)\|^2_{L^2(\mu_\beta)} \le \sum_{y\in\mathbb{Z}^d} \|\nabla u(y,\phi)\|^2.$$

The previous inequality implies that the bilinear form

$$\begin{aligned}\sum_{x,y\in\mathbb{Z}^d} \left\langle(\partial_y u(x,\cdot))(\partial_y w(x,\cdot))\right\rangle_{\mu_\beta} + \frac{1}{2\beta}\sum_{x\in\mathbb{Z}^d}\left\langle\nabla u(x,\cdot)\nabla w(x,\cdot)\right\rangle_{\mu_\beta} \\ + \frac{1}{2\beta}\sum_{n\ge1}\frac{1}{\beta^{\frac{n}{2}}}\sum_{x\in\mathbb{Z}^d}\left\langle\nabla^{n+1}u(x,\cdot),\nabla^{n+1}w(x,\cdot)\right\rangle_{\mu_\beta} + \sum_{q\in\mathcal{Q}}\left\langle\nabla_q u\cdot\mathbf{a}_q\nabla_q w\right\rangle_{\mu_\beta} - \left\langle\mathbf{f}w(y,\cdot)\right\rangle_{\mu_\beta}.\end{aligned}$$

is coercive. The result then follows from the Lax-Milgram Theorem. □

As it was the case for the variational solutions of the Helffer-Sjöstrand equation, the Green's matrix admits a dynamical interpretation relying on the heat kernel $P^{\phi\cdot}$ as stated in the following proposition.

Proposition 3.15. *Fix $\mathbf{f} \in L^2(\mu_\beta)$ and $y \in \mathbb{Z}^d$. The Green's matrix $\mathcal{G}_{\mathbf{f}}(\cdot; y)$ satisfies the identity*

$$\mathcal{G}_{\mathbf{f}}(x, \phi; y) = \int_0^\infty \mathbb{E}_\phi \left[\mathbf{f}(\phi_t) P^{\phi\cdot}(t, y, x) \right] dt. \tag{3.68}$$

Remark 3.16. Using the previous proposition with the bound of Proposition 3.10 on the heat kernel, one can extend the definition of the Green's matrix to functions $\mathbf{f} \in L^1(\mu_\beta)$ (instead of $\mathbf{f} \in L^2(\mu_\beta)$).

Combining Proposition 3.10 and Proposition 3.15, we obtain the following upper bounds on the Green's function $\mathcal{G}_{\mathbf{f}}$, its gradient and mixed derivative.

Proposition 3.17. *For any regularity exponent $\varepsilon > 0$, there exists an inverse temperature $\beta_1(d, \varepsilon) < \infty$ such that for any $\beta > \beta_1$ the following result holds. There exists a constant $C(d, \beta) < \infty$ such that for any $x, y \in \mathbb{Z}^d$,*

$$\| \mathcal{G}_{\mathbf{f}}(x, \cdot; y) \|_{L^2(\mu_\beta)} \leq \frac{C \|\mathbf{f}\|_{L^2(\mu_\beta)}}{|x - y|^{d-2}},$$

and the regularity estimates on the gradient and the mixed derivative

$$\| \nabla_x \mathcal{G}_{\mathbf{f}}(x, \cdot; y) \|_{L^2(\mu_\beta)} \leq \frac{C \|\mathbf{f}\|_{L^2(\mu_\beta)}}{|x - y|^{d-1-\varepsilon}} \qquad \text{and} \qquad \| \nabla_x \nabla_y \mathcal{G}_{\mathbf{f}}(x, \cdot; y) \|_{L^2(\mu_\beta)} \leq \frac{C \|\mathbf{f}\|_{L^2(\mu_\beta)}}{|x - y|^{d-\varepsilon}}.$$

In the second part of this section, we investigate the homogenization properties of the Green's matrix, which are a crucial ingredient in the proof of Theorem 1. The main result we establish is stated in Theorem 2. The proof of this result is the subject of Sections 6 and 7.

From the theory of stochastic homogenization and its application to the Helffer-Sjöstrand equation [58, 8, 81, 9], one expects that there exists a deterministic, positive definite matrix $\bar{\mathbf{a}}_\beta$ (which is a small perturbation of the matrix $\frac{1}{2\beta} I_d$) such that the Green's matrix associated with the Helffer-Sjöstrand operator (3.57), defined by

$$\mathcal{L}\mathcal{G} = \delta_0 \text{ in } \mathbb{Z}^d \times \Omega$$

homogenizes to the Green's matrix $\overline{G}$ associated with the Laplacian operator $\nabla \cdot \bar{\mathbf{a}}_\beta \nabla$

$$-\nabla \cdot \bar{\mathbf{a}}_\beta \nabla \overline{G} = \delta_0 \text{ in } \mathbb{Z}^d, \tag{3.69}$$

in the sense that, as $x \to \infty$,

$$\left\| \mathcal{G}(x, \cdot) - \overline{G}(x) \right\|_{L^2(\mu_\beta)} = o\left(\frac{1}{|x|^{d-2}} \right).$$

When applying to the Villain model (see computations in Section 4) we need more precise result: specifically, we need to prove quantitative homogenization for the mixed derivative associated with the Green's matrix. Results of this nature have been established in the homogenization literature (see e.g., [8, Section 8.6] or [20]). The main contribution of Sections 6 and 7 is to adapt the techniques developed in [9, 8] to the setting of the Helffer-Sjöstrand operator $\mathcal{L}$.

In order to state Theorem 2, we need to introduce an important quantity in stochastic homogenization: the first-order corrector. For $i, j \in \{1, \ldots, d\} \times \{1, \ldots, \binom{d}{2}\}$, we recall the notation l_{ij} for the affine function introduced in Section 2

$$l_{ij} := \begin{cases} \mathbb{R}^d \to \mathbb{R}^{\binom{d}{2}}, \\ x \mapsto (0, \ldots, 0, x \cdot e_i, 0, \ldots, 0), \end{cases}$$

where the term $x \cdot e_i$ appears in the j-th position. We denote by $\nabla \chi_{ij}$ the gradient of the first-order corrector, defined to be the unique stationary solution of the Helffer-Sjöstrand equation

$$\mathcal{L}(l_{ij} + \chi_{ij}) = 0 \text{ in } \mathbb{Z}^d \times \Omega.$$

It is precisely defined in Proposition 6.29. Once equipped with the gradient of the corrector, we can define the exterior derivative $\mathrm{d}^* \chi_{ij}$ by using that the codifferential d^* is a linear functional of the gradient (see (A.23)). The following theorem proves a quantitative homogenization result for a version of the mixed derivative of the Green's function (3.68), the specific form of the function (3.70) is justified by the fact that it is the correct object to consider in order to prove Theorem 1 in Section 4.

Theorem 2 (Homogenization of the mixed derivative of the Green's matrix)**.** *We fix a charge $q_1 \in \mathcal{Q}$ such that 0 belongs to the support of n_{q_1}, let $\mathcal{U}_{q_1}$ be the solution of the Helffer-Sjöstrand equation*

$$\mathcal{L}\mathcal{U}_{q_1} = \cos\left(2\pi\left(\phi, q_1\right)\right) q_1 \text{ in } \mathbb{Z}^d \times \Omega, \tag{3.70}$$

and let $\overline{G}_{q_1} := \left(\overline{G}_{q_1,1}, \ldots, \overline{G}_{q_1,\binom{d}{2}}\right)$ be the map defined by the formula, for each integer $k \in \left\{1, \ldots, \binom{d}{2}\right\}$,

$$\overline{G}_{q_1,k} = \sum_{1 \le i \le d} \sum_{1 \le j \le \binom{d}{2}} \left\langle \cos\left(2\pi\left(\phi, q_1\right)\right) \left(n_{q_1}, \mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{ij}\right)\right\rangle_{\mu_\beta} \nabla_i \overline{G}_{jk}. \tag{3.71}$$

There exist an inverse temperature $\beta_1 := \beta_1(d) < \infty$, an exponent $\gamma := \gamma(d) > 0$ and a constant C_{q_1} which satisfies the estimate $|C_{q_1}| \le C \left\|q_1\right\|_1^k$ for some $C := C(d, \beta) < \infty$ and $k := k(d) < \infty$, such that for each $\beta \ge \beta_0$, and each radius $R \ge 1$, one has the inequality

$$\left\| \nabla \mathcal{U}_{q_1} - \sum_{1 \le i \le d} \sum_{1 \le j \le \binom{d}{2}} \left(e_{ij} + \nabla \chi_{ij}\right) \nabla_i \overline{G}_{q_1,j} \right\|_{\underline{L}^2\left(B_{2R} \setminus B_R, \mu_\beta\right)} \le \frac{C_{q_1}}{R^{d+\gamma}}. \tag{3.72}$$

Remark 3.18. The functions $\nabla \mathcal{U}_{q_1}$ and $\nabla_i \overline{G}_{q_1}$ behave like mixed derivative of Green's matrices, in particular, they should decay like the map $x \to |x|^{-d}$. Theorem 2 states that their difference is quantitatively smaller than the typical size of the two functions: we obtain an algebraic rate of convergence with additional exponent $\gamma > 0$ in the right side of (3.72).

Remark 3.19. For the purposes of Section 4, we record here that the statement of Theorem 2 can be simplified by using the formalism of discrete differential forms and exploiting the symmetries of the system. In particular, we have the following properties:

- The operator $-\nabla \cdot \bar{\mathbf{a}}_\beta \nabla$ can be written

$$-\nabla \cdot \bar{\mathbf{a}}_\beta \nabla = \frac{1}{2\beta}\left(\mathrm{d}^*\mathrm{d} + \left(1 + \overline{\lambda}_\beta\right)\mathrm{d}\mathrm{d}^*\right), \tag{3.73}$$

  where $\overline{\lambda}_\beta$ is a real coefficient which is small and tends to 0 as β tends to infinity. This property is stated in Remark 6.11;
- The gradient of the infinite volume corrector only depends on the value of the codifferential $\mathrm{d}^* l_{e_{ij}}$ (in particular, it is equal to 0 if $\mathrm{d}^* l_{e_{ij}} = 0$) as mentioned in Remarks 6.27 and 6.30. We use the notation of Remark 6.30: given an integer $k \in \{1, \ldots, d\}$, we let select a vector $p := \sum_{1 \le i \le d} \sum_{1 \le j \le \binom{d}{2}} p_{ij} e_{ij}$ such that $\mathrm{d}^* l_p = e_k$ and denote by $\nabla \chi_k := \sum_{1 \le i \le d} \sum_{1 \le j \le \binom{d}{2}} p_{ij} \nabla \chi_{ij}$.

Using these ingredients, we can rewrite the definition of the map $\overline{G}_{q_1,k}$ stated in (3.71): we have

$$\overline{G}_{q_1,k} = \sum_{1 \le i \le d} \left\langle \cos\left(2\pi\left(\phi, q_1\right)\right)\left(n_{q_1}, e_i + \mathrm{d}^* \chi_i\right)\right\rangle_{\mu_\beta} \left(\mathrm{d}^* \overline{G}_{\cdot k}\right) \cdot e_i.$$

We then use that, by definition, the map $\overline{G}_{\cdot,k}$ solves the equation $-\nabla \cdot \bar{\mathbf{a}}_\beta \nabla \overline{G} = \delta_0$, and the identities $-\Delta = \mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}$, $\mathrm{d} \circ \mathrm{d} = 0$, and $\mathrm{d}^* \circ \mathrm{d}^* = 0$ to write

$$\begin{aligned} -\left(1 + \overline{\lambda}_\beta\right) \Delta \mathrm{d}^* \overline{G}_{\cdot,k} = \left(1 + \overline{\lambda}_\beta\right)\left(\mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}\right) \mathrm{d}^* \overline{G}_{\cdot k} &= \left(1 + \overline{\lambda}_\beta\right) \mathrm{d}^* \mathrm{d}\mathrm{d}^* \overline{G}_{\cdot k} \\ &= \mathrm{d}^* \left(\mathrm{d}^* \mathrm{d} \overline{G}_{\cdot k} + \left(1 + \overline{\lambda}_\beta\right) \mathrm{d}\mathrm{d}^*\right) \overline{G}_{\cdot k} \\ &= \mathrm{d}^* \left(-\nabla \cdot \bar{\mathbf{a}}_\beta \nabla \overline{G}_{\cdot,k}\right) \\ &= \mathrm{d}^* \delta_0. \end{aligned}$$

The exterior derivative $\mathrm{d}^* \overline{G}$ can thus be explicitly computed in terms of the gradient of the Green's matrix associated with the operator $-\left(1 + \overline{\lambda}_\beta\right)\Delta$, which is equal to the standard random walk Green's function on the lattice $\mathbb{Z}^d$ multiplied by the value $\left(1 + \overline{\lambda}_\beta\right)^{-1}$.

4. First-order expansion of the two-point function: Overview of the proof

In this section, we show that Theorem 1 can be obtained by combining Theorem 2, which gives a quantitative rate of convergence of the mixed gradients of the Helffer-Sjöstrand Green's matrix, with the regularity theory for the Helffer-Sjöstrand operator established in Section 5.

In order to prove Theorem 1 it is enough, by Proposition 3.6, to prove the expansion stated in the following theorem.

Theorem 3. *There exist constants $\beta_0 := \beta_0(d)$, $c_0 := c_0(\beta, d), c_1(\beta, d)$, and an exponent $\gamma' := \gamma'(d) > 0$ such that for every $\beta > \beta_0$, and every $x \in \mathbb{Z}^d$,*

$$\frac{Z(\sigma_{0x})}{Z(0)} = c_0 + \frac{c_1}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right),$$

and

$$\frac{Z(\overline{\sigma}_{0x})}{Z(0)} = c_0 + \frac{\overline{c}_1}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right).$$

The proof of Theorem 3 requires to use the following statements stated in Section 3 and proved in Sections 5, 6 and 7:

- We use the quantitative homogenization of the mixed derivative of the Green's matrix associated with the Helffer-Sjöstrand operator $\mathcal{L}$. The precise statement we need to use is stated in Theorem 2. The proof of this theorem is the subject of Sections 6 and 7;
- We use the $C^{0,1-\varepsilon}$-regularity theory established in Section 5; more specifically, we need to use the regularity estimates for the Helffer-Sjöstrand Green's matrix stated in Proposition 3.17 and on the Green's matrix associated with the second-order Helffer-Sjöstrand operator stated in Proposition 5.13. We additionally make the assumption that the regularity exponent ε is very small compared to the exponent γ which appears in the statement of Theorem 2 (for instance, we assume that the ratio γ/ε is larger than $100d$). This condition can always be ensured by increasing the inverse temperature β (as the exponent γ depends only on the dimension).

Apart from these three results, the proof of Theorem 3, which is contained in this section (and Section 8 for the technical estimates), is largely independent from Sections 5, 6 and 7.

This section is organized as follows. We first set up the argument and introduce some preliminary notation in Section 4.1. We then simplify the expression (4.1) below in a series of technical lemmas stated in Sections 4.2, 4.3 and 4.4. In particular, in Sections 4.3 and 4.4, we sketch the argument that one can decouple the Helffer-Sjöstrand Green's matrix from the exponential terms arising from the dual model in Section 3. The proofs of these lemmas rely on the $C^{0,1-\varepsilon}$-regularity theory established in Section 5, we give an outline of the arguments and postpone the proofs to Section 8. The core of the proof of Theorem 3 (thus Theorem 1) is given in Section 4.5. This section is decomposed into two subsections. We first write an outline of the argument in Section 4.5.1 and then present the details of the proof in Section 4.5.2.

4.1. Preliminary notation.

We first recall that we have the identity

$$\frac{Z_\beta(\sigma_{0x})}{Z_\beta(0)} = \left\langle \exp\left(\sum_{q \in \mathcal{Q}} z(\beta, q) \sin(2\pi(\phi, q)) \sin(2\pi(\sigma_{0x}, q)) + \sum_{q \in \mathcal{Q}} z(\beta, q) \cos(2\pi(\phi, q)) (\cos(2\pi(\sigma_{0x}, q)) - 1)\right)\right\rangle_{\mu_\beta}. \tag{4.1}$$

We also recall that, by the definition of the function σ_{0x} given in Section 3.1, we have the equality

$$\begin{aligned} d^* \sigma_{0x} = d^* d (-\Delta)^{-1} h_{0x} = h_{0,x} - dd^* (-\Delta)^{-1} h_{0x} &= h_{0,x} - d(-\Delta)^{-1} d^* h_{0x} \\ &= h_{0,x} - d(-\Delta)^{-1} (\mathbb{1}_x - \mathbb{1}_0) \\ &= h_{0,x} + \nabla G - \nabla G_x. \end{aligned}$$

We then use the identity $q = dn_q$, that the maps q, n_q and $h_{0,x}$ take values in $\mathbb{Z}$, and the periodicity of the sine and the cosine to deduce that

$$\sin(2\pi(\sigma_{0x}, q)) = \sin(2\pi(\nabla G - \nabla G_x, n_q)) \quad \text{and} \quad \cos(2\pi(\sigma_{0x}, q)) = \cos(2\pi(\nabla G - \nabla G_x, n_q)).$$

One can then expand the sine and the cosine by using the trigonometric formulae. We obtain the identities

$$\begin{aligned}\sin\left(2\pi(\nabla G - \nabla G_x, n_q)\right) &= \sin\left(2\pi(\nabla G, n_q)\right) - \sin\left(2\pi(\nabla G_x, n_q)\right) \\ &\quad + \left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right)\sin\left(2\pi(\nabla G, n_q)\right) - \left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right)\sin\left(2\pi(\nabla G_x, n_q)\right),\end{aligned} \tag{4.2}$$

and

$$\begin{aligned}\cos\left(2\pi(\nabla G - \nabla G_x, n_q)\right) - 1 &= \left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right)\left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) \\ &\quad + \left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right) + \left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) + \sin\left(2\pi(\nabla G, n_q)\right)\sin\left(2\pi(\nabla G_x, n_q)\right).\end{aligned} \tag{4.3}$$

We then combine the identities (4.2) and (4.3) with the right side of (4.1). To ease the notation, we introduce the following random variables

$$\begin{cases} X_x := \exp\left(- \sum_{q \in \mathcal{Q}} z(\beta, q) \left(\sin\left(2\pi(\phi, q)\right)\sin\left(2\pi(\nabla G_x, n_q)\right) - \frac{1}{2}\cos\left(2\pi(\phi, q)\right)\left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) \right) \right), \\ Y_0 := \exp\left(\sum_{q \in \mathcal{Q}} z(\beta, q) \left(\sin\left(2\pi(\phi, q)\right)\sin\left(2\pi(\nabla G, n_q)\right) + \frac{1}{2}\cos\left(2\pi(\phi, q)\right)\left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right) \right) \right), \\ Y_x := \exp\left(\sum_{q \in \mathcal{Q}} z(\beta, q) \left(\sin\left(2\pi(\phi, q)\right)\sin\left(2\pi(\nabla G_x, n_q)\right) + \frac{1}{2}\cos\left(2\pi(\phi, q)\right)\left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) \right) \right), \\ X_{\sin\cos} := \exp\left(- \sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\phi, q)\right)\sin\left(2\pi(\nabla G_x, n_q)\right)\left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right) \right) \\ \qquad \times \exp\left(\sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\phi, q)\right)\sin\left(2\pi(\nabla G, n_q)\right)\left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) \right), \\ X_{\cos\cos} := \exp\left(\sum_{q \in \mathcal{Q}} z(\beta, q) \cos\left(2\pi(\phi, q)\right)\left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right)\left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) \right), \\ X_{\sin\sin} := \exp\left(\sum_{q \in \mathcal{Q}} z(\beta, q) \cos\left(2\pi(\phi, q)\right)\sin\left(2\pi(\nabla G, n_q)\right)\sin\left(2\pi(\nabla G_x, n_q)\right) \right). \end{cases} \tag{4.4}$$

In this notation we have

$$\frac{Z_\beta(\sigma_{0x})}{Z_\beta(0)} = \left\langle Y_0 X_x X_{\sin\cos} X_{\cos\cos} X_{\sin\sin} \right\rangle_{\mu_\beta}. \tag{4.5}$$

Our aim is to first simplify the identity (4.5) and then to apply Theorem 2.

4.2. **Removing the terms $X_{\sin\cos}$, $X_{\cos\cos}$ and $X_{\sin\sin}$.** We first show that the terms $X_{\sin\cos}$, $X_{\cos\cos}$ and $X_{\sin\sin}$ are lower order terms which can be removed from the analysis. We prove the following lemma.

Lemma 4.1. *There exist constants $\beta_0 := \beta_0(d) < \infty$, $c := c(d, \beta)$, and $C := C(d, \beta)$ such that, for each $\beta > \beta_0$,*

$$\frac{Z_\beta(\sigma_{0x})}{Z_\beta(0)} = \left\langle Y_0 X_x \right\rangle_{\mu_\beta} + \frac{c \left\langle Y_0 X_x \right\rangle_{\mu_\beta}}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-1}} \right). \tag{4.6}$$

A consequence of the identity (4.6) is the equivalence

$$\exists c_1, c_2 \in \mathbb{R}, \frac{Z_\beta(\sigma_{0x})}{Z_\beta(0)} = c_1 + \frac{c_2}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}} \right) \\ \iff \exists c_1, c_2 \in \mathbb{R}, \left\langle Y_0 X_x \right\rangle_{\mu_\beta} = c_1 + \frac{c_2}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}} \right).$$

This lemma is technical and its proof is not the core of the argument; the proof is thus deferred to Section 8. We provide here a sketch of the argument.

Sketch of the proof of Lemma 4.1. To prove the identity (4.6), we first record four standard inequalities, for each $y \in \mathbb{Z}^d$, and each $a \in \mathbb{R}$,

$$(4.7) \qquad |\nabla G(y)| \leq \frac{C}{|y|^{d-1}}, \quad |\nabla G_x(y)| \leq \frac{C}{|y-x|^{d-1}}, \quad |\sin a| \leq |a|, \quad \text{and} \quad |\cos a - 1| \leq \frac{1}{2}|a|^2.$$

Using the estimates (4.7) and the exponential decay of the coefficient $z(\beta, q)$, we prove the following estimates:

(i) The random variables $X_{\sin\cos}$ and $X_{\cos\cos}$ belong to the space $L^\infty(\mu_\beta)$ and satisfy the estimates

$$(4.8) \qquad \begin{cases} \|X_{\sin\cos} - 1\|_{L^\infty} \leq \dfrac{C}{|x|^{d-1}}, \\ \|X_{\cos\cos} - 1\|_{L^\infty} \leq \dfrac{C}{|x|^{d-1}}. \end{cases}$$

(ii) We prove that the random variable $X_{\sin\sin}$ also belongs to the space $L^\infty(\mu_\beta)$ and that its fluctuations around the value 1 are of order $|x|^{2-d}$. This is larger than the fluctuations of the random variables $X_{\sin\cos}$ and $X_{\cos\cos}$ and one needs to be more precise in the analysis: we prove the following estimates on the expectation and the variance of $X_{\sin\sin}$

$$(4.9) \qquad \begin{cases} \operatorname{var}_{\mu_\beta} X_{\sin\sin} \leq \dfrac{C}{|x|^{2d-2}}, \\ \langle X_{\sin\sin} \rangle_{\mu_\beta} = 1 + \dfrac{c}{|x|^{d-2}} + O\left(\dfrac{C}{|x|^{d-1}}\right). \end{cases}$$

The variance is estimated thanks to the Brascamp-Lieb inequality and the expectation is estimated thanks to the estimates (4.7) and a Taylor expansion of the exponential.

A combination of the estimates (4.8) and (4.9) is then sufficient to prove Lemma 4.1. □

Remark 4.2. The same proof also yields

$$(4.10) \qquad \frac{Z_\beta(\overline{\sigma}_{0x})}{Z_\beta(0)} = \langle Y_0 Y_x \rangle_{\mu_\beta} + \frac{\overline{c}\,\langle Y_0 Y_x \rangle_{\mu_\beta}}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-1}}\right).$$

In general, $\overline{c} \neq c$ since the $O(|x|^{d-2})$ term above is contributed by $\langle X_{\sin\sin}^{-1} \rangle_{\mu_\beta}$ instead of $\langle X_{\sin\sin} \rangle_{\mu_\beta}$.

4.3. **Removing the contributions of the cosines.** From Lemma 4.1, we see that to prove Theorem 1, it is sufficient to obtain the following expansion

$$(4.11) \qquad \exists c_1, c_2 \in \mathbb{R}, \quad \langle Y_0 X_x \rangle_{\mu_\beta} = c_1 + \frac{c_2}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right).$$

Let us note that, by the translation invariance of the measure μ_β, the expectation of the random variable X_x does not depend on the point $x \in \mathbb{Z}^d$: we have, for each $x \in \mathbb{Z}^d$, $\langle X_x \rangle_{\mu_\beta} = \langle X_0 \rangle_{\mu_\beta}$. A consequence of this observation is that to prove (4.11), it is sufficient to show

$$(4.12) \qquad \operatorname{cov}[X_x, Y_0] = \frac{c_2}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right).$$

Indeed, the expansion (4.12) implies (4.11) with the value $c_1 = \langle Y_0 \rangle_{\mu_\beta} \langle X_0 \rangle_{\mu_\beta}$. To prove the identity (4.11), we use the Helffer-Sjöstrand representation formula and write the covariance in the following form

$$(4.13) \qquad \operatorname{cov}[X_x, Y_0] = \sum_{y \in \mathbb{Z}^d} \langle (\partial_y X_x) \mathcal{Y}(y, \cdot) \rangle_{\mu_\beta},$$

where $\mathcal{Y} : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ is the solution of the Helffer-Sjöstrand equation, for each $(y, \phi) \in \mathbb{Z}^d \times \Omega$,

$$(4.14) \qquad \mathcal{L}\mathcal{Y}(y, \phi) = \partial_y Y_0(\phi).$$

For each point $x \in \mathbb{Z}^d$, we introduce the notation $Q_x : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ to denote the following function: for each pair $(y, \phi) \in \mathbb{Z}^d \times \Omega$,

$$(4.15) \qquad Q_x(y, \phi) := 2\pi \sum_{q \in \mathcal{Q}} z(\beta, q) \cos(2\pi(\phi, q)) \sin(2\pi(\nabla G_x, n_q))\, q(y).$$

These charges are defined so as to have the identities, for each $y \in \mathbb{Z}^d$,

$$\partial_y Y_0(\phi) = \left(Q_0(y,\phi) - \frac{1}{2} 2\pi \sum_{q \in \mathcal{Q}} z(\beta, q) \sin(2\pi(\phi, q)) \left(\cos(2\pi(\nabla G, n_q)) - 1\right) q(y) \right) Y_0(\phi) \tag{4.16}$$

and

$$\partial_y X_x(\phi) = -\left(Q_x(y,\phi) + \frac{1}{2} 2\pi \sum_{q \in \mathcal{Q}} z(\beta, q) \sin(2\pi(\phi, q)) \left(\cos(2\pi(\nabla G_x, n_q)) - 1\right) q(y) \right) X_x(\phi). \tag{4.17}$$

We also define the random charges $n_{Q_x} : \mathbb{Z}^d \times \Omega \to \mathbb{R}^d$ according to the formula

$$n_{Q_x} := \sum_{q \in \mathcal{Q}} 2\pi z(\beta, q) \left(\cos(2\pi(\phi, q)) \sin(2\pi(\nabla G_x, n_q))\right) n_q \quad \text{so that} \quad \mathrm{d} n_{Q_x} = Q_x. \tag{4.18}$$

We note that, by the exponential decay $|z(\beta, q)| \leq C e^{-c\sqrt{\beta}\|q\|_1}$, the decay of the gradient of the Green's matrix stated in (4.7), and the inequality $|\sin a| \leq |a|$, the random charges Q_x and n_{Q_x} satisfy the $L^\infty(\mu_\beta)$-estimate: for each $y \in \mathbb{Z}^d$,

$$\|Q_x(y, \cdot)\|_{L^\infty(\mu_\beta)} \leq \frac{C}{|y - x|^{d-1}} \text{ and } \|n_{Q_x}(y, \cdot)\|_{L^\infty(\mu_\beta)} \leq \frac{C}{|y - x|^{d-1}}. \tag{4.19}$$

By a similar argument, but using this time the inequality $|\cos a - 1| \leq \frac{1}{2}|a|^2$, one obtains the inequality, for each $y \in \mathbb{Z}^d$,

$$\left| \sum_{q \in \mathcal{Q}} z(\beta, q) \sin(2\pi(\phi, q)) \left(\cos(2\pi(\nabla G_x, n_q)) - 1\right) q(y) \right| \leq \frac{C}{|y - x|^{2d-2}} \tag{4.20}$$

and

$$\left| \sum_{q \in \mathcal{Q}} z(\beta, q) \sin(2\pi(\phi, q)) \left(\cos(2\pi(\nabla G_x, n_q)) - 1\right) n_q(y) \right| \leq \frac{C}{|y - x|^{2d-2}}. \tag{4.21}$$

The reason we record the inequalities (4.20) and (4.21) is that, since $2d - 2 > d - 1$, the function $x \mapsto |x|^{2d-2}$ decays faster than $x \mapsto |x|^{d-1}$. From this observation, we expect that the terms $Q_0(y) Y_0$ and $Q_x(y) X_x$ are the leading order terms in the identities (4.16) and (4.17), and that the terms involving the cosine of the gradient of the Green's functions are lower order terms which can be removed from the analysis. This is what is proved in the following lemma.

Lemma 4.3 (Removing the contributions of the cosines)**.** *One has the identity*

$$\operatorname{cov}[X_x, Y_0] = \sum_{y \in \mathbb{Z}^d} \langle X_x Q_x(y) \mathcal{V}(y, \cdot) \rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1-\varepsilon}} \right), \tag{4.22}$$

where $\mathcal{V} : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ *is the solution of the Helffer-Sjöstrand equation, for each pair* $(y, \phi) \in \mathbb{Z}^d \times \Omega$,

$$\mathcal{L} \mathcal{V}(y, \phi) = Q_0(y, \phi) Y_0(\phi). \tag{4.23}$$

A consequence of the identity (4.22) *is the equivalence*

$$\exists c_2 \in \mathbb{R},\ \operatorname{cov}[X_x, Y_0] = \frac{c_2}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}} \right) \iff \exists c_2 \in \mathbb{R},\ \sum_{y \in \mathbb{Z}^d} \langle X_x Q_x(y) \mathcal{V}(y, \cdot) \rangle_{\mu_\beta} = \frac{c_2}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}} \right).$$

The proof of this result is again technical and does not represent the core of the argument; it is thus deferred to Section 8. The argument relies on two ingredients:

(i) We use the decay estimates for the Helffer-Sjöstrand Green's matrix, its gradient and its mixed derivative stated in Proposition 3.17;
(ii) We use the estimates (4.19) and (4.21) and take advantage of the fact that the function $x \mapsto |x|^{2d-2}$ decays faster than the map $x \mapsto |x|^{d-1}$.

We complete this section by recording that we may also prove

$$\exists c_1, c_2 \in \mathbb{R}, \quad \langle Y_0 Y_x \rangle_{\mu_\beta} = c_1 + \frac{c_2}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}} \right) \tag{4.24}$$

by showing

$$\operatorname{cov}[Y_x, Y_0] = \frac{c_2}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right). \tag{4.25}$$

Indeed, we have the following analogue of (4.22)

$$\operatorname{cov}[Y_x, Y_0] = \sum_{y \in \mathbb{Z}^d} \langle Y_x Q_x(y) \mathcal{V}(y, \cdot) \rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1-\varepsilon}}\right).$$

The proof of this identity is almost the same as (4.22) with only notational changes, and is therefore omitted.

4.4. **Decoupling the exponentials.** The next (and final) technical step consists in removing the exponential terms X_x and Y_0 from the computation. To this end, we prove the decorrelation estimate stated in the following lemma.

Lemma 4.4 (Decoupling the exponential terms)**.** *One has the expansions*

$$\operatorname{cov}[X_x, Y_0] = \langle Y_0 \rangle_{\mu_\beta} \langle X_0 \rangle_{\mu_\beta} \sum_{y \in \mathbb{Z}^d} \langle Q_x(y, \cdot) \mathcal{U}(y, \cdot) \rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1+\varepsilon}}\right), \tag{4.26}$$

and

$$\operatorname{cov}[Y_x, Y_0] = \langle Y_0 \rangle_{\mu_\beta}^2 \sum_{y \in \mathbb{Z}^d} \langle Q_x(y, \cdot) \mathcal{U}(y, \cdot) \rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1+\varepsilon}}\right). \tag{4.27}$$

where the function $\mathcal{U} : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ is the solution of the Helffer-Sjöstrand equation

$$\mathcal{L}\mathcal{U} = Q_0 \ \text{in } \mathbb{Z}^d \times \Omega.$$

The identity (4.26) implies the equivalence

$$\exists c_2 \in \mathbb{R}, \ \sum_{y \in \mathbb{Z}^d} \langle X_x Q_x(y, \cdot) \mathcal{V}(y, \cdot) \rangle_{\mu_\beta} = \frac{c_2}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right)$$
$$\iff \quad \exists c_2 \in \mathbb{R}, \ \sum_{y \in \mathbb{Z}^d} \langle Q_x(y) \mathcal{U}(y, \cdot) \rangle_{\mu_\beta} = \frac{c_2}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right).$$

Remark 4.5. The function $\mathcal{U}$ can be decomposed according to the following procedure: if, for each charge $q \in \mathcal{Q}$, we denote by $\mathcal{U}_q : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ the solution of the Helffer-Sjöstrand equation

$$\mathcal{L}\mathcal{U}_q = \cos\left(2\pi\left(\phi, q\right)\right) q \ \text{in } \mathbb{Z}^d \times \Omega, \tag{4.28}$$

then we have the identity

$$\mathcal{U} = 2\pi \sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\nabla G, n_q)\right) \mathcal{U}_q. \tag{4.29}$$

Remark 4.6. By writing $q = \mathrm{d} n_q$, we can rewrite the equation (4.28) in the following form

$$\mathcal{L}\mathcal{U}_q = \mathrm{d}\left(\cos\left(2\pi\left(\cdot, q_1\right)\right) n_q\right) \ \text{in } \mathbb{Z}^d \times \Omega.$$

As a consequence the function $\mathcal{U}_q$ can be expressed in terms of the Helffer-Sjöstrand Green's matrix according to the formula, for each pair $(y, \phi) \in \mathbb{Z}^d \times \Omega$,

$$\mathcal{U}_q(y, \phi) = \sum_{z \in \operatorname{supp} n_q} \mathrm{d}_z^* \mathcal{G}_{\cos(2\pi(\cdot, q))}(y, \phi; z) n_q(z). \tag{4.30}$$

Using the decay estimate on the gradient and mixed derivative of the Green's matrix stated in Proposition 3.17, we obtain that the map $\mathcal{U}_q$ satisfies the upper bounds, for each $y \in \mathbb{Z}^d$,

$$\|\mathcal{U}_q(y, \cdot)\|_{L^\infty(\mu_\beta)} \leq \frac{C_q}{|y - z|^{d-1-\varepsilon}} \quad \text{and} \quad \|\nabla \mathcal{U}_q(y, \cdot)\|_{L^\infty(\mu_\beta)} \leq \frac{C_q}{|y - z|^{d-\varepsilon}}, \tag{4.31}$$

where z is a point which belongs to the support of the charge n_q (chosen arbitrarily).

Remark 4.7. A consequence of the estimate (4.31) is that by using the exponential decay of the coefficient $z(\beta, q)$ (see (3.14)) and the inequality, for each charge $q \in \mathcal{Q}$,

$$|\sin(2\pi(\nabla G, n_q))| \leq 2\pi |(\nabla G, n_q)| \leq 2\pi \|\nabla G\|_{L^2(\operatorname{supp} n_q)} \|n_q\|_2 \leq \frac{C_q}{|z|^{d-1}},$$

where z is a point in the support of n_q (chosen arbitrarily), we deduce the inequality, for each point $y \in \mathbb{Z}^d$,

$$\begin{aligned}
\|\mathcal{U}(y,\cdot)\|_{L^\infty(\mu_\beta)} &\leq 2\pi \sum_{z\in\mathbb{Z}^d} \sum_{q\in\mathcal{Q}_z} |z(\beta,q)\sin(2\pi(\nabla G, n_q))| \, \|\mathcal{U}_q(y,\cdot)\|_{L^\infty(\mu_\beta)} \\
&\leq \sum_{z\in\mathbb{Z}^d} \sum_{q\in\mathcal{Q}_z} e^{-c\sqrt{\beta}\|q\|_1} \frac{C_q}{|z|^{d-1} \times |y-z|^{d-1-\varepsilon}} \\
&\leq C \sum_{z\in\mathbb{Z}^d} \frac{1}{|z|^{d-1} \times |y-z|^{d-1-\varepsilon}} \\
&\leq \frac{C}{|y|^{d-2-\varepsilon}}.
\end{aligned}$$

where we used the exponential decay of the term $e^{-c\sqrt{\beta}\|q\|_1}$ to absorb the algebraic growth of the term $C_q \leq C\|q\|_1^k$ in the third inequality. The same argument also yields the estimate

$$\|\nabla\mathcal{U}(y,\cdot)\|_{L^\infty(\mu_\beta)} \leq \frac{C}{|y|^{d-1-\varepsilon}}.$$

We now give an heuristic argument explaining why we expect the decoupling estimate (4.26) to hold.

Heuristic of the proof of Lemma 4.4. The strategy of the proof is to first decouple the exponential term X_x and then decouple the exponential term Y_0; to decouple the term X_x, we prove the expansion

$$\sum_{y\in\mathbb{Z}^d} \langle X_x Q_x(y,\cdot)\mathcal{V}(y,\cdot)\rangle_{\mu_\beta} = \langle X_0\rangle_{\mu_\beta} \sum_{y\in\mathbb{Z}^d} \langle Q_x(y,\cdot)\mathcal{V}(y,\cdot)\rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1-\varepsilon}}\right). \tag{4.32}$$

A heuristic reason justifying why one can expect the expansion (4.32) to hold is the following. By the definition of the random variable X_x given in (4.4) and the decay of the gradient of the Green's function ∇G_x stated in (4.7), we expect the random variable X_x to essentially depend on the value of the gradient of the field around the point x. The statement is voluntarily vague; one could give a mathematical meaning to it by arguing that if one considers a large constant C depending only on the dimension d, then the conditional expectation of the random variable X_x with respect to the sigma-algebra generated by the fields $(\nabla\phi(y))_{y\in B(x,C)}$ is a good approximation of the random variable X_x in the space $L^2(\mu_\beta)$.

Additionally, using similar arguments to the one presented in Remarks 4.6 and 4.7, but using the $L^2(\mu_\beta)$-estimate $\|Y_0\|_{L^2(\mu_\beta)} \leq C$ instead of the (stronger) pointwise upper bound $|\cos(2\pi(\phi,q))| \leq 1$, one obtains the $L^2(\mu_\beta)$-estimate, for each $y \in \mathbb{Z}^d$,

$$\|\nabla\mathcal{V}(y,\cdot)\|_{L^2(\mu_\beta)} \leq \frac{C}{|y|^{d-1-\varepsilon}}. \tag{4.33}$$

While we can prove the estimate (4.33) using Proposition 3.17, we expect that its real decay is of order $|y|^{1-d}$, and make this assumption for the rest of the argument. We use an integration by parts to write, for each field $\phi \in \Omega$,

$$\sum_{y\in\mathbb{Z}^d} Q_x(y,\phi)\mathcal{V}(y,\phi) = \sum_{y\in\mathbb{Z}^d} n_{Q_x}(y,\phi)\mathrm{d}^*\mathcal{V}(y,\phi).$$

Since we expect the random charge $n_{Q_x}(y)$ to decay like $|y-x|^{1-d}$ (see the estimate (4.19)), and the random variable $\mathrm{d}^*\mathcal{V}(y,\cdot)$ to decay $|y|^{1-d}$ (since the codifferential d^* is a linear functional of the gradient ∇), we have

$$\sum_{y\in\mathbb{Z}^d} n_{Q_x}(y,\phi)\mathrm{d}^*\mathcal{V}(y,\phi) \simeq \sum_{y\in\mathbb{Z}^d} \frac{1}{|y-x|^{d-1}} \times \frac{1}{|y|^{d-1}} \simeq \frac{1}{|x|^{d-2}}. \tag{4.34}$$

The point of the identity (4.34) is that while we expect the sum $\sum_{y\in\mathbb{Z}^d} n_{Q_x}(y,\phi)\mathrm{d}^*\mathcal{V}(y,\phi)$ to be of order $|x|^{2-d}$, its restriction to the ball $B(x,C)$ is of lower-order since we have

$$\sum_{y\in B(x,C)} n_{Q_x}(y,\phi)\mathrm{d}^*\mathcal{V}(y,\phi) \simeq \sum_{y\in B(x,C)} \frac{1}{|y-x|^{d-1}} \times \frac{1}{|y|^{d-1}} \simeq \frac{1}{|x|^{d-1}}.$$

A consequence of this result is that we expect the main contribution of the sum $\sum_{y\in\mathbb{Z}^d} n_{Q_x}(y,\phi)\mathrm{d}^*\mathcal{V}(y,\phi)$ to come mostly from the points y outside the ball $B(x,C)$.

To summarize the heuristic explanation, one should expect that:

- The random variable X_x depends mostly on the gradient of the field inside a ball $B(x,C)$ for some large but fixed constant C depending only on the dimension;
- The random variable $\sum_{y\in\mathbb{Z}^d} n_{Q_x}(y,\phi)\mathrm{d}^*\mathcal{V}(y,\phi)$ depends mostly on the value of the gradient of the field outside the ball $B(x,C)$.

Since the gradient of the field decorrelates (sufficiently fast in our case), we expect the random variable $\sum_{y\in\mathbb{Z}^d} n_{Q_x}(y,\phi)\mathrm{d}^*\mathcal{V}(y,\phi)$ and X_x to decorrelate; this is what is proved by (4.32).

Once we have proved the identity (4.32), we can prove the expansion (4.22) by applying the same argument, and by using the symmetry of the Helffer-Sjöstrand operator. □

4.5. **First order expansion of the two-point function.** Once the Lemmas 4.1, 4.3 and 4.4 are established, we have showed that, to prove Theorem 3, it is enough to obtain the expansion

$$\exists c\in\mathbb{R},\ \sum_{y\in\mathbb{Z}^d}\langle Q_x(y)\mathcal{U}(y,\cdot)\rangle_{\mu_\beta} = \frac{c}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right). \tag{4.35}$$

This section is devoted to the proof of (4.35). We first give a sketch of the proof in Section 4.5.1 and provide the details of the argument in Section 4.5.2.

4.5.1. *Heuristic argument.* In this section, we present a heuristic argument for the proof of the expansion (4.35). A large part of the proof is concerned with the treatment of the technicalities inherent to the dual Villain model (sum over all the charges $q\in\mathcal{Q}$, presence of a sine etc.). In order to highlight the main ideas of the argument, we make the following simplifications:

- We assume that for β large enough, one may essentially reduce the charges to the collection of dipoles $\left(\mathrm{d}\mathbb{1}_{\{y,y+e_i\}}\right)_{y\in\mathbb{Z}^d,1\leq i\leq d}$. The exponential decay on the coefficient $z(\beta,q)$ constraints the L^1-norm of the charge q to be small. One can thus assume that only the charges $q\in\mathcal{Q}$ which minimize the value $\|q\|_1$ contribute to the sum; this leads us to considering the dipoles $\left(\mathrm{d}\mathbb{1}_{\{x,x+e_i\}}\right)_{x\in\mathbb{Z}^d,1\leq i\leq d}$. We will thus assume that only the dipoles $\left(\mathrm{d}\mathbb{1}_{\{x,x+e_i\}}\right)_{x\in\mathbb{Z}^d,1\leq i\leq d}$ count in the sum, and we will denote by $z(\beta) = z(\beta;\mathrm{d}\mathbb{1}_{\{x,x+e_i\}})$.

  An important, but mostly technical, part of the argument presented in Section 4.5.2 is devoted to proving that this dipole approximation yields the correct picture. Under this assumption, one has the simplifications

$$Q_x = z(\beta)\sum_{i=1}^d\sum_{y\in\mathbb{Z}^d} 2\pi\sin\left(2\pi\nabla_i G(y)\right)\mathrm{d}\mathbb{1}_{\{y,y+e_i\}} \quad\text{and}\quad \mathcal{U} = z(\beta)\sum_{i=1}^d\sum_{y\in\mathbb{Z}^d} 2\pi\sin\left(2\pi\nabla_i G_x(y)\right)\mathcal{U}_{y,i},$$

  where the function $\mathcal{U}_{y,i}$ is the solution of the Helffer-Sjöstrand equation

$$\mathcal{L}\mathcal{U}_{y,i} = \mathrm{d}\left(\cos\left(2\pi\left(\mathrm{d}^*\phi(y)\cdot e_i\right)\right)\mathbb{1}_{\{y,y+e_i\}}\right) \quad\text{in}\quad \mathbb{Z}^d\times\Omega.$$

- Since the gradients of the Green's functions $\nabla_i G(y)$ are usually small, we consider the first-order expansion of the sine and replace the value $\sin\left(2\pi\nabla_i G_x(y)\right)$ by $2\pi\nabla_i G_x(y)$. With this assumption, we have

$$Q_x = z(\beta)(2\pi)^2\sum_{i=1}^d\sum_{y\in\mathbb{Z}^d}\nabla_i G(y)\mathrm{d}\mathbb{1}_{\{y,y+e_i\}} \quad\text{and}\quad \mathcal{U} = z(\beta)(2\pi)^2\sum_{i=1}^d\sum_{y\in\mathbb{Z}^d}\nabla_i G_x(y)\mathcal{U}_{y,i}.$$

Using these simplifications, we compute

$$\sum_{y\in\mathbb{Z}^d}\langle Q_x(y)\mathcal{U}(y,\cdot)\rangle_{\mu_\beta} = z(\beta)^2(4\pi^2)^2\sum_{i,j=1}^d\sum_{y,y_1\in\mathbb{Z}^d}\nabla_i G(y)\nabla_j G_x(y_1)\left\langle\cos(2\pi\mathrm{d}^*\phi(y_1)\cdot e_i)\mathrm{d}^*\mathcal{U}_{y,j}(y_1,\phi)\cdot e_i\right\rangle_{\mu_\beta}. \tag{4.36}$$

Using the translation invariance of the measure μ_β, one has the identity, for each pair of points $y,y_1\in\mathbb{Z}^d$,

$$\left\langle\cos(2\pi\mathrm{d}^*\phi(y_1)\cdot e_i)\mathrm{d}^*\mathcal{U}_{y,j}(y_1,\phi)\cdot e_i\right\rangle_{\mu_\beta} = \left\langle\cos(2\pi\mathrm{d}^*\phi(y_1-y)\cdot e_i)\mathrm{d}^*\mathcal{U}_{0,j}(y_1-y,\phi)\cdot e_i\right\rangle_{\mu_\beta}. \tag{4.37}$$

Putting the identity (4.37) into the equality (4.36) and performing the change of variable $\kappa := y_1 - y$, we obtain (4.38)

$$\sum_{y\in\mathbb{Z}^d} \left\langle Q_x(y)\mathcal{U}(y,\cdot)\right\rangle_{\mu_\beta} = z(\beta)^2(4\pi^2)^2 \sum_{i,j=1}^d \sum_{y,\kappa\in\mathbb{Z}^d} \nabla_i G(y)\nabla_j G_x(\kappa-y)\left\langle \cos(2\pi \mathrm{d}^*\phi(\kappa)\cdot e_i)\mathrm{d}^*\mathcal{U}_{0,j}(\kappa,\phi)\cdot e_i\right\rangle_{\mu_\beta}.$$

The strategy is then to simplify the right side of (4.38) by arguing that the term $\mathrm{d}^*\mathcal{U}_{0,j}$ behaves like the mixed derivative of a deterministic Green's function. Proving a quantitative version of this result is the subject of Theorem 2 which is proved in Sections 7 and 8; in this setting, it can be stated as follows: there exists an exponent $\gamma := \gamma(d) > 0$ and, for each pair of integers $i,j\in\{1,\ldots,d\}$, there exist deterministic constants $c_{i,j} := c_{i,j}(d,\beta)$ such that, for each radius $R\geq 1$,

$$\sum_{\kappa\in B_{2R}\setminus B_R}\left|\left\langle \cos(2\pi \mathrm{d}^*\phi(\kappa)\cdot e_i)\mathrm{d}^*\mathcal{U}_{0,j}(\kappa,\phi)\cdot e_i\right\rangle_{\mu_\beta} - \sum_{i_1,j_1=1}^d c_{i,i_1}c_{j,j_1}\nabla_{i_1}\nabla_{j_1}G(\kappa)\right| \leq \frac{C}{R^\gamma}. \tag{4.39}$$

Once equipped with this estimate, we let $\mathcal{E}_{i,j}:\mathbb{Z}^d\mapsto\mathbb{R}$ be the error term defined according to the formula, for each $\kappa\in\mathbb{Z}^d$,

$$\mathcal{E}_{i,j}(\kappa) := \left\langle \cos(\mathrm{d}^*\phi(\kappa)\cdot e_i)\mathrm{d}^*\mathcal{U}_{0,j}(\kappa,\phi)\cdot e_i\right\rangle_{\mu_\beta} - \sum_{i_1,j_1=1}^d c_{i,i_1}c_{j,j_1}\nabla_{i_1}\nabla_{j_1}G(\kappa).$$

According to the regularity estimate on the gradient of the Helffer-Sjöstrand Green's matrix stated in Proposition 3.17 (via the formula (4.30)) and the homogenization estimate (4.39), this term satisfies the L^1 and pointwise estimates

$$\forall R\geq 1,\ \ \frac{1}{R^d}\sum_{\kappa\in B_{2R}\setminus B_R}\left|\mathcal{E}_{i,j}(\kappa)\right| \leq \frac{C}{R^{d+\gamma}} \quad\text{and}\quad \forall\kappa\in\mathbb{Z}^d,\ \ \left|\mathcal{E}_{i,j}(\kappa)\right|\leq\frac{C}{|\kappa|^{d-\varepsilon}}. \tag{4.40}$$

We can use the definition of the term $\mathcal{E}_{i,j}$ to rewrite the identity (4.38). We obtain

$$\begin{aligned}\sum_{y\in\mathbb{Z}^d} \left\langle Q_x(y)\mathcal{U}(y,\cdot)\right\rangle_{\mu_\beta} &= z(\beta)^2(4\pi^2)^2 \sum_{i,i_1,j,j_1=1}^d c_{i,i_1}c_{j,j_1}\sum_{y,\kappa\in\mathbb{Z}^d}\nabla_i G(y)\nabla_j G_x(\kappa-y)\nabla_{i_1}\nabla_{j_1}G(\kappa) \\ &\quad + z(\beta)^2(4\pi^2)^2\sum_{i,j=1}^d \nabla_i G(y)\nabla_j G_x(\kappa-y)\mathcal{E}_{i,j}(\kappa).\end{aligned} \tag{4.41}$$

The right side of the identity (4.41) can then be refined. First using the estimates (4.40) on the error term $\mathcal{E}_{i,j}$ and Proposition 8.4 proved in Section 8.5, we can show the following expansion: there exists an exponent $\gamma' := \gamma'(d) > 0$ such that

$$\sum_{i,j=1}^d\sum_{y,\kappa\in\mathbb{Z}^d}\nabla_i G(y)\nabla_j G_x(\kappa-y)\mathcal{E}_{i,j}(\kappa) = \sum_{i,j=1}^d K_{i,j}\sum_{y,\kappa\in\mathbb{Z}^d}\nabla_i G(y)\nabla_j G_x(\kappa-y) + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right), \tag{4.42}$$

where the constants $K_{i,j}$ are obtained from the error terms $\mathcal{E}_{i,j}$ according to the formula

$$K_{i,j} := z(\beta)^2\left(4\pi^2\right)^2\sum_{\kappa\in\mathbb{Z}^d}\mathcal{E}_{i,j}(\kappa),$$

which, by the estimate (4.40), is well-defined. A combination of the identity (4.41) with the expansion (4.42) then shows

$$\begin{aligned}\sum_{y\in\mathbb{Z}^d} \left\langle Q_x(y)\mathcal{U}(y,\cdot)\right\rangle_{\mu_\beta} &= z(\beta)^2(4\pi^2)^2 \sum_{i,i_1,j,j_1=1}^d c_{i,i_1}c_{j,j_1}\sum_{y,\kappa\in\mathbb{Z}^d}\nabla_i G(y)\nabla_j G_x(\kappa-y)\nabla_{i_1}\nabla_{j_1}G(\kappa) \\ &\quad + \sum_{i,j=1}^d K_{i,j}\sum_{y,\kappa\in\mathbb{Z}^d}\nabla_i G(y)\nabla_j G_x(\kappa-y) + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right).\end{aligned} \tag{4.43}$$

This expansion does not give the result (4.35) directly and we need to exploit the symmetries of the dual Villain model to conclude. The argument relies on the following observation: since the Villain and dual Villain model are invariant under the action of the group H of the lattice preserving transformations introduced in Section 2, the same property holds for the two-point function, and thus for the map $x\mapsto\sum_{y\in\mathbb{Z}^d}\left\langle Q_x(y)\mathcal{U}(y,\cdot)\right\rangle_{\mu_\beta}$.

One can then use this invariance property together with the expansion (4.43) to prove that this expansion must take the simpler form

$$\sum_{y \in \mathbb{Z}^d} \langle Q_x(y) \mathcal{U}(y, \cdot) \rangle_{\mu_\beta} = \frac{c}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}} \right). \tag{4.44}$$

This is achieved by using the property of the discrete Green's function and relies on tools from Fourier analysis. The proof can be found in Section 8.4. The expansion (4.44) is exactly (4.35); the proof is thus complete.

4.5.2. *Proof of the expansion* (4.35). We first write $Q_x = \mathrm{d} n_{Q_x}$, perform an integration by parts, and use the identities (4.18) and (4.29) to expand the sum $\sum_{y \in \mathbb{Z}^d} \langle Q_x(y) \mathcal{U}(y, \cdot) \rangle_{\mu_\beta}$. We obtain

$$\begin{aligned} &\sum_{y \in \mathbb{Z}^d} \langle Q_x(y) \mathcal{U}(y, \cdot) \rangle_{\mu_\beta} \\ &\quad = \sum_{y \in \mathbb{Z}^d} \langle n_{Q_x}(y) \mathrm{d}^* \mathcal{U}(y, \cdot) \rangle_{\mu_\beta} \\ &\quad = (4\pi^2)^2 \sum_{y \in \mathbb{Z}^d} \sum_{q_1, q_2 \in \mathcal{Q}} z(\beta, q_1) z(\beta, q_2) \sin\left(2\pi (\nabla G, n_{q_2})\right) \sin\left(2\pi (\nabla G_x, n_{q_1})\right) \langle \cos\left(2\pi (\phi, q_1)\right) \mathrm{d}^* \mathcal{U}_{q_2}(y, \phi) \rangle_{\mu_\beta} n_{q_1}(y). \end{aligned} \tag{4.45}$$

To simplify the sum over all the charges q_1, q_2, we introduce an equivalence relation on the set of charges $\mathcal{Q}$: we say that two charges q and q' are equivalent, and denote it by $q \sim q'$, if and only if one is the translation of the other, i.e.,

$$q \sim q' \iff \exists z \in \mathbb{Z}^d, \ \ q(z + \cdot) = q'.$$

This relation gives rise to a quotient space, which we denote by $\mathcal{Q}/\mathbb{Z}^d$. For each charge $q \in \mathcal{Q}$, we denote by $[q]$ its equivalence class. For each equivalence class $[q] \in \mathcal{Q}/\mathbb{Z}^d$, we select a charge $q \in \mathcal{Q}_0$ which belongs to this equivalent class (and break ties among the possible candidates by using an arbitrary criterion). We note that, for each charge $q \in \mathcal{Q}$, by the definition of the charge n_q and of the coefficient $z(\beta, q)$, we have the identities, for each point $z \in \mathbb{Z}^d$,

$$z(\beta, q) = z(\beta, q(\cdot - z)), \qquad n_{q(\cdot - z)} = n_q(\cdot - z) \quad \text{and} \quad \left(n_{q(\cdot - z)}\right) = (n_q). \tag{4.46}$$

We also note that, by using the translation invariance of the measure μ_β and the definition of the function $\mathcal{U}_{q_2}$ given in (4.28), we have the equality, for each pair of points $(y, z) \in \mathbb{Z}^d$,

$$\left\langle \cos\left(2\pi (\phi, q_1)\right) \mathrm{d}^* \mathcal{U}_{q_2(\cdot - z)}(y, \phi) \right\rangle_{\mu_\beta} = \langle \cos\left(2\pi (\phi, q_1(\cdot + z))\right) \mathrm{d}^* \mathcal{U}_{q_2}(y - z, \phi) \rangle_{\mu_\beta}.$$

Additionally, we can decompose the sum over the charges $q \in \mathcal{Q}$ along the equivalence classes, i.e., we can write, for any summable function $F : \mathcal{Q} \to \mathbb{R}$,

$$\sum_{q \in \mathcal{Q}} F(q) = \sum_{[q] \in \mathcal{Q}/\mathbb{Z}^d} \sum_{z \in \mathbb{Z}^d} F(q(\cdot - z)). \tag{4.47}$$

Combining the identities (4.46) and (4.47), we can rewrite the equality (4.45),

$$\begin{aligned} \sum_{y \in \mathbb{Z}^d} \langle Q_x(y) \mathcal{U}(y, \cdot) \rangle_{\mu_\beta} &= (4\pi^2)^2 \sum_{[q_1], [q_2] \in \mathcal{Q}/\mathbb{Z}^d} z(\beta, q_1) z(\beta, q_2) \\ &\times \left[\sum_{z_1, z_2, y \in \mathbb{Z}^d} \sin\left(2\pi (\nabla G, n_{q_2}(\cdot - z_2))\right) \sin\left(2\pi (\nabla G_x, n_{q_1}(\cdot - z_1))\right) \langle \cos\left(2\pi (\phi, q_1(\cdot - z_1 + z_2))\right) \mathrm{d}^* \mathcal{U}_{q_2}(y - z_2, \phi) \rangle_{\mu_\beta} n_{q_1}(y - z_1) \right]. \end{aligned} \tag{4.48}$$

We first rearrange the identity (4.48). We use the identities $(\nabla G_x, n_{q_1}(\cdot - z_1)) = (\nabla G_x(\cdot + z_1), n_{q_1})$, $(\nabla G, n_{q_2}(\cdot - z_2)) = (\nabla G(\cdot + z_2), n_{q_2})$, and perform the change of variable $y := y - z_1$. We obtain

$$\begin{aligned} \sum_{y \in \mathbb{Z}^d} \langle Q_x(y) \mathcal{U}(y, \cdot) \rangle_{\mu_\beta} &= (4\pi^2)^2 \sum_{[q_1], [q_2] \in \mathcal{Q}/\mathbb{Z}^d} z(\beta, q_1) z(\beta, q_2) \\ &\times \underbrace{\left[\sum_{z_1, z_2, y \in \mathbb{Z}^d} \sin\left(2\pi (\nabla G(\cdot + z_2), n_{q_2})\right) \sin\left(2\pi (\nabla G_x(\cdot + z_1), n_{q_1})\right) \langle \cos\left(2\pi (\phi, q_1(\cdot - z_1 + z_2))\right) \mathrm{d}^* \mathcal{U}_{q_2}(y + z_1 - z_2, \phi) \rangle_{\mu_\beta} n_{q_1}(y) \right]}_{(4.49)-(q_1, q_2)}. \end{aligned} \tag{4.49}$$

The rest of the proof is decomposed into two steps:

- In the first step, we use Theorem 2 and the regularity estimates established in Proposition 3.17 to prove the following result: there exists an exponent $\gamma' := \gamma'(d) > 0$ such that for each pair of charges $q_1, q_2 \in \mathcal{Q}$, and each pair of integers $(i,j) \in \{1, \ldots, d\}^2$, there exist constants $K_{q_1,q_2} := K_{q_1,q_2}(q_1, q_2, d, \beta)$, $C_{q_1,q_2} := C_{q_1,q_2}(q_1,q_2,d,\beta)$, $c_{ij}^{q_1} := c_{ij}^{q_1}(i,j,q_1,d,\beta)$ such that the term (4.49) $-(q_1,q_2)$ satisfies the expansion

$$\text{(4.49)} - (q_1, q_2) = \sum_{i,j,k,l=1}^{d} c_{ij}^{q_1} c_{kl}^{q_2} \sum_{z_1, z_2 \in \mathbb{Z}^d} \nabla_i G(z_1) \nabla_j \nabla_k G(z_1 - z_2) \nabla_l G(x - z_2) \\ + K_{q_1,q_2} \sum_{z_1 \in \mathbb{Z}^d} \nabla G(z_1) \cdot (n_{q_2}) \times \nabla G_x(x - z_1) \cdot (n_{q_1}) + O\left(\frac{C_{q_1,q_2}}{|x|^{d-2+\gamma'}} \right). \tag{4.50}$$

We recall that the vectors (n_{q_1}) and (n_{q_2}) belongs to $\mathbb{R}^d$ and are defined by the formulae

$$(n_{q_1}) := \sum_{y \in \mathbb{Z}^d} n_{q_1}(y) \quad \text{and} \quad (n_{q_2}) := \sum_{y \in \mathbb{Z}^d} n_{q_2}(y).$$

We also record that all the constants K_{q_1,q_2}, $c_{ij}^{q_1,q_2}$ and C_{q_1,q_2} grow at most algebraically fast in the values $\|q_1\|_1$ and $\|q_2\|_1$, i.e., there exist an exponent $k := k(d) < \infty$ and a constant $C := C(d,\beta) < \infty$ such that one has the estimates

$$\left| c_{ij}^{q_1} \right| \leq C \|q_1\|_1^k, \quad |K_{q_1,q_2}| \leq C \|q_1\|_1^k \|q_2\|_1^k, \quad \text{and} \quad |C_{q_1,q_2}| \leq C \|q_1\|_1^k \|q_2\|_1^k. \tag{4.51}$$

- In the second step, we use the symmetry and rotation invariance of the dual Villain model to prove that the expansion (4.50) implies the expansion (4.35).

We focus on the proof of (4.50) and first simplify the term (4.49) $-(q_1,q_2)$ by removing the sine. To this end, we use the following ingredients:

- We use the inequality, $|\sin a - a| \leq \frac{1}{6} a^3$, valid for any real number $a \in \mathbb{R}$, and the inequality, for each charge $q \in \mathcal{Q}_0$, and each point $z \in \mathbb{Z}^d$,

$$|(\nabla G, n_q(\cdot - z))| \leq \frac{C_q}{|z|^{d-1}}.$$

We deduce that, for each pair of points $z_1, z_2 \in \mathbb{Z}^d$,

$$\left| \sin\left(2\pi (\nabla G(\cdot + z_2), n_{q_2})\right) - 2\pi (\nabla G(\cdot + z_2), n_{q_2}) \right| \leq \frac{C_{q_2}}{|z_2|^{3d-3}}, \tag{4.52}$$

and

$$\left| \sin\left(2\pi (\nabla G_x(\cdot + z_1), n_{q_1})\right) - 2\pi (\nabla G_x(\cdot + z_1), n_{q_1}) \right| \leq \frac{C_{q_1}}{|z_1 - x|^{3d-3}}; \tag{4.53}$$

- We further simplify the terms $2\pi(\nabla G, n_{q_2}(\cdot - z_2))$ and $2\pi(\nabla G_x, n_{q_1}(\cdot - z_1))$. We use that the double gradient of the Green's function decays like $|z|^{-d}$, and the assumption that the point 0 belongs to the supports of the charges n_{q_1} and n_{q_2}. We obtain

$$|2\pi(\nabla G_x(\cdot + z_1), n_{q_1}) - 2\pi\,(n_{q_1}) \cdot \nabla G_x(z_1)| = |2\pi(\nabla G_x(z_1 + \cdot) - \nabla G_x(z_1), n_{q_1})| \leq \frac{C_{q_2}}{|z_2 - x|^d}.$$

The same argument shows the estimate

$$|2\pi(\nabla G, n_{q_2}(\cdot - z_2)) - 2\pi\,(n_{q_2}) \cdot \nabla G(z_2)| \leq \frac{C_{q_2}}{|z_2|^d}. \tag{4.54}$$

We then combine the inequalities (4.52) and (4.54) on the one hand, (4.53) and (4.54) on the other hand, and use the inequality $3d - 3 > d$. We obtain the two estimates

$$\left| \sin\left(2\pi(\nabla G_x, n_{q_2}(\cdot - z_2))\right) - 2\pi\,(n_{q_2}) \cdot \nabla G_x(z_2) \right| \leq \frac{C_{q_2}}{|x - z_1|^d}, \tag{4.55}$$

and

$$\left| \sin\left(2\pi(\nabla G_x, n_{q_1}(\cdot - z_1))\right) - 2\pi\,(n_{q_1}) \cdot \nabla G(z_1) \right| \leq \frac{C_{q_1}}{|z_1|^d}; \tag{4.56}$$

- We use the estimate (4.31) and deduce that, for each point y in the support of n_{q_1},

$$\left| \left\langle \cos\left(2\pi(\phi, q_1(\cdot - z_1 + z_2))\right) \mathrm{d}^* \mathcal{U}_{q_2}(y + z_1 - z_2, \phi) \right\rangle_{\mu_\beta} \right| \leq \frac{C_{q_1,q_2}}{|z_1 - z_2|^{d-\varepsilon}};$$

- We have the inequalities, for each point $x \in \mathbb{Z}^d$,

$$\sum_{z_1, z_2 \in \mathbb{Z}^d} \frac{1}{|x - z_1|^d} \times \frac{1}{|z_1 - z_2|^{d-\varepsilon}} \times \frac{1}{|z_2|^{d-1}} \leq \frac{C \ln |x|}{|x|^{d-1-\varepsilon}} \text{ and } \sum_{z_1, z_2 \in \mathbb{Z}^d} \frac{1}{|x - z_1|^{d-1}} \times \frac{1}{|z_1 - z_2|^{d-\varepsilon}} \times \frac{1}{|z_2|^d} \leq \frac{C}{|x|^{d-1-\varepsilon}}.$$

A combination of the four items listed above implies the expansion

(4.57)
$$\begin{aligned}(4.49) - (q_1, q_2) &= \underbrace{(4\pi^2)^2 \sum_{z_1, z_2, y \in \mathbb{Z}^d} \nabla G(z_2) \cdot (n_{q_2}) \nabla G_x(z_1) \cdot (n_{q_1}) \left\langle \cos\left(2\pi(\phi, q_1(\cdot - z_1 + z_2))\right) \mathrm{d}^* \mathcal{U}_{q_2}(y + z_1 - z_2, \phi) \right\rangle_{\mu_\beta} n_{q_1}(y)}_{(4.57)-(q_1,q_2)} \\ &\quad + O\left(\frac{C_{q_1,q_2}}{|x|^{d-1-\varepsilon}} \right).\end{aligned}$$

A consequence of the identity (4.57) is that to prove the expansion (4.50), it is enough to prove the following result

(4.58)
$$\begin{aligned}(4.57) - (q_1, q_2) &= \sum_{i,j,k,l=1}^{d} c_{ij}^{q_1} c_{kl}^{q_2} \sum_{z_1, z_2 \in \mathbb{Z}^d} \nabla_i G(z_1) \nabla_j \nabla_k G(z_1 - z_2) \nabla_l G(x - z_2) \\ &\quad + K_{q_1,q_2} \sum_{z_1, y \in \mathbb{Z}^d} \nabla G(z_2) \cdot (n_{q_2}) \nabla G_x(z_2 + \kappa) \cdot (n_{q_1}) + O\left(\frac{C_{q_1,q_2}}{|x|^{d-2+\gamma}} \right).\end{aligned}$$

The rest of the argument is devoted to the proof of (4.58) and relies on the homogenization of the mixed derivative of the Helffer-Sjöstrand Green's matrix stated in Theorem 2.

We first consider the term $(4.57) - (q_1, q_2)$ and perform the change of variable $\kappa := z_1 - z_2$. We obtain

(4.59)
$$\begin{aligned}&(4.57) - (q_1, q_2) \\ &\quad = (4\pi^2)^2 \sum_{z_1, \kappa, y \in \mathbb{Z}^d} \nabla G(z_2) \cdot (n_{q_2}) \nabla G_x(z_2 + \kappa) \cdot (n_{q_1}) \left\langle \cos\left(2\pi(\phi, q_1(\cdot - \kappa))\right) \mathrm{d}^* \mathcal{U}_{q_2}(y + \kappa, \phi) \right\rangle_{\mu_\beta} n_{q_1}(y).\end{aligned}$$

We then post-process the result of Theorem 2 so that it can be used to estimate the term $(4.57) - (q_1, q_2)$; the objective is to prove the estimate (4.63) below. We use that the codifferential d^* is a linear functional of the gradient to deduce from Theorem 2 that, for each radius $R \geq 1$,

(4.60)
$$\left\| \mathrm{d}^* \mathcal{U}_{q_2} - \sum_{1 \leq i \leq d} \sum_{1 \leq j \leq \binom{d}{2}} \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{ij} \right) \nabla_i \overline{G}_{q_2, j} \right\|_{\underline{L}^2(B_{2R} \setminus B_R, \mu_\beta)} \leq \frac{C_{q_2}}{R^{d+\gamma}}.$$

We recall the notation $A_R := B_{2R} \setminus B_R$. Using the arguments and notation introduced in Remark 3.19, we obtain the identity

$$\sum_{1 \leq i \leq d} \sum_{1 \leq j \leq \binom{d}{2}} \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{ij} \right) \nabla_i \overline{G}_{q_2, j} = \sum_{1 \leq i \leq d} \left(e_i + \mathrm{d}^* \chi_i \right) \left(\mathrm{d}^* \overline{G}_{q_2} \cdot e_i \right).$$

The estimate (4.60) then implies, by using the stationarity of the gradient of the infinite-volume corrector and the Cauchy-Schwarz inequality,

(4.61)
$$\sum_{\kappa \in A_R} \left| \left\langle \cos\left(2\pi(\phi, q_1(\cdot - \kappa))\right) \mathrm{d}^* \mathcal{U}_{q_2}(\kappa, \phi) \right\rangle_{\mu_\beta} - \sum_{1 \leq i \leq d} \left\langle \cos\left(2\pi(\phi, q_1)\right) \left(e_i + \mathrm{d}^* \chi_i(0, \phi) \right) \right\rangle_{\mu_\beta} \left(\mathrm{d}^* \overline{G}_{q_2}(\kappa) \cdot e_i \right) \right| \leq \frac{C_{q_2}}{R^\gamma}.$$

The inequality (4.61) can be generalized into the following result: for each point $y \in \mathbb{Z}^d$,

(4.62)
$$\sum_{\kappa \in A_R} \left| \left\langle \cos\left(2\pi(\phi, q_1(\cdot - \kappa))\right) \mathrm{d}^* \mathcal{U}_{q_2}(y+\kappa, \phi) \right\rangle_{\mu_\beta} - \sum_{1 \le i \le d} \left\langle \cos\left(2\pi(\phi, q_1)\right) \left(e_i + \mathrm{d}^*\chi_i(y,\phi)\right) \right\rangle_{\mu_\beta} \left(\mathrm{d}^* \overline{G}_{q_2}(\kappa) \cdot e_i\right) \right| \le \frac{C_{q_2}(1+|y|^{2d+\gamma})}{R^\gamma}.$$

The proof of the estimate (4.62) relies on a technical argument, we omit it here and refer to the long version of the article for the details ([36, Chapter 4, Section 5.2]).

We then consider the estimate (4.62) for a point y in the support of the charge n_{q_1}, take the scalar product with the vector $n_{q_1}(y)$, and sum over all the points y in the support of n_{q_1}. We obtain

(4.63)
$$\sum_{\kappa \in A_R} \left| \left\langle \cos\left(2\pi(\phi, q_1(\cdot - \kappa))\right) \left(n_{q_1}, \mathrm{d}^* \mathcal{U}_{q_2}(\cdot+\kappa, \phi)\right) \right\rangle_{\mu_\beta} - \sum_{1 \le i \le d} \left\langle \cos\left(2\pi(\phi, q_1)\right) \left(n_{q_1}, e_i + \mathrm{d}^*\chi_i(y,\phi)\right) \right\rangle_{\mu_\beta} \left(\mathrm{d}^* \overline{G}_{q_2}(\kappa) \cdot e_i\right) \right| \le \frac{C_{q_1,q_2}}{R^\gamma}.$$

We then focus on the term (4.57) – (q_1,q_2) (and more specifically on the right side of (4.59)) and use the inequality (4.63) to simplify it. To ease the notation, we introduce the following definitions:

- We let $\mathcal{E}_{q_1,q_2}$ be the map from $\mathbb{Z}^d$ to $\mathbb{R}$ defined according to the formula, for each point $\kappa \in \mathbb{Z}^d$,
$$\mathcal{E}_{q_1,q_2}(\kappa) := \left\langle \cos\left(2\pi(\phi, q_1(\cdot - \kappa))\right) \left(n_{q_1}, \mathrm{d}^* \mathcal{U}_{q_2}(\cdot+\kappa, \phi)\right) \right\rangle_{\mu_\beta} - \sum_{1 \le i \le d} \left\langle \cos\left(2\pi(\phi, q_1)\right) \left(n_{q_1}, e_i + \mathrm{d}^*\chi_i\right) \right\rangle_{\mu_\beta} \left(\mathrm{d}^* \overline{G}_{q_2}(\kappa) \cdot e_i\right).$$
It is an error term which is small; in view of the estimate (4.63), Remark 4.6, and the definition of the map $G_{q_2,j}$ stated in (3.71), it satisfies the inequalities
$$\forall R \ge 1, \quad \sum_{\kappa \in A_R} |\mathcal{E}_{q_1,q_2}(\kappa)| \le \frac{C_{q_1,q_2}}{R^\gamma} \quad \text{and} \quad \forall \kappa \in \mathbb{Z}^d, \quad |\mathcal{E}_{q_1,q_2}(\kappa)| \le \frac{C_{q_1,q_2}}{|\kappa|^{d-\varepsilon}}; \tag{4.64}$$
- We recall the definition of the coefficient $\overline{\lambda}_\beta$ stated in Remark 4.6. For each pair of integers $(i,j) \in \{1,\ldots,d\}^2$, we define the coefficient c_{ij}^q according to the formula
$$c_{ij}^q := 4\pi^2 \left(1+\overline{\lambda}_\beta\right)^{-\frac{1}{2}} \left[(n_q)\cdot e_i\right] \times \left\langle \cos\left(2\pi(\phi, q_1)\right) \left(n_{q_1}, e_j + \mathrm{d}^*\chi_j\right) \right\rangle_{\mu_\beta}.$$

Using these notation together with Remark 3.19 and an explicit computation (which we omit here), we obtain the formula

$$\sum_{1 \le i \le d} \left\langle \cos\left(2\pi(\phi, q_1)\right) \left(n_{q_1}, e_i + \mathrm{d}^*\chi_i(y,\phi)\right) \right\rangle_{\mu_\beta} \left(\mathrm{d}^* \overline{G}_{q_2}(\kappa) \cdot e_i\right) = \left(1+\overline{\lambda}_\beta\right)^{-1} \sum_{1 \le i,j \le d} \left\langle \cos\left(2\pi(\phi, q_1)\right) \left(n_{q_1}, e_i + \mathrm{d}^*\chi_i\right) \right\rangle_{\mu_\beta} \left\langle \cos\left(2\pi(\phi, q_2)\right) \left(n_{q_2}, e_j + \mathrm{d}^*\chi_j\right) \right\rangle_{\mu_\beta} \nabla_i \nabla_j G.$$

The term (4.57) – (q_1,q_2) then becomes

$$(4.57) - (q_1,q_2) = \sum_{i,j,k,l=1}^{d} c_{ij}^{q_1} c_{kl}^{q_2} \sum_{z_1,z_2 \in \mathbb{Z}^d} \nabla_i G(z_1) \nabla_j \nabla_k G(z_1 - z_2) \nabla_l G(x - z_2) + 4\pi^2 \sum_{z_2,\kappa \in \mathbb{Z}^d} \nabla G(z_2) \cdot (n_{q_2}) \nabla G(z_2 + \kappa - x) \cdot (n_{q_1}) \mathcal{E}_{q_1,q_2}(\kappa).$$

To prove the estimate (4.58), it is thus sufficient to prove that there exists a constant K_{q_1,q_2} such that

$$4\pi^2 \sum_{z_2,\kappa \in \mathbb{Z}^d} \nabla G(z_2) \cdot (n_{q_2}) \nabla G_x(z_2 + \kappa) \cdot (n_{q_1}) \mathcal{E}_{q_1,q_2}(\kappa) = K_{q_1,q_2} \sum_{z_1 \in \mathbb{Z}^d} \nabla G(z_2) \cdot (n_{q_2}) \nabla G_x(z_2) \cdot (n_{q_1}) + O\left(\frac{C_{q_1,q_2}}{R^{d+\gamma'}}\right).$$

The proof of this result relies on the estimates (4.64); it is the subject of Proposition 8.4 and is deferred to Section 8. We note that the argument gives the following explicit value for the constant K_{q_1,q_2}

$$K_{q_1,q_2} = 4\pi^2 \sum_{\kappa \in \mathbb{Z}^d} \mathcal{E}_{q_1,q_2}(\kappa).$$

By the estimates (4.64), the constant K_{q_1,q_2} is well-defined and grows at most algebraically fast in the parameters $\|q_1\|_1$ and $\|q_2\|_1$ as required. The proof of the expansion (4.50) is complete.

We complete the proof of Theorem 1 by showing that (4.50) implies the result. We first sum the expansion (4.50) over all the equivalence classes $[q_1],[q_2] \in \mathcal{Q}/\mathbb{Z}^d$, and use the exponential decay of the coefficients $z(\beta,q_1)$ and $z(\beta,q_2)$ to absorb the algebraic growth of the constants $c_{ij}^{q_1}$, $c_{ij}^{q_2}$ and C_{q_1,q_2}. We obtain

(4.65)

$$\begin{aligned}
\sum_{y\in\mathbb{Z}^d} \left\langle Q_x(y)\mathcal{U}(y,\cdot)\right\rangle_{\mu_\beta} &= 4\pi^2 \sum_{[q_1],[q_2]\in\mathcal{Q}/\mathbb{Z}^d} z(\beta,q_1)\, z(\beta,q_2) \times (4.49) - (q_1,q_2) \\
&= 4\pi^2 \sum_{[q_1],[q_2]\in\mathcal{Q}/\mathbb{Z}^d} \sum_{i,j,k,l=1}^d z(\beta,q_1)\, z(\beta,q_2)\, c_{ij}^{q_1} c_{kl}^{q_1} \sum_{z_1,z_2\in\mathbb{Z}^d} \nabla_i G(z_1)\nabla_j\nabla_k G(z_1-z_2)\nabla_l G(x-z_2) \\
&\quad + 4\pi^2 \sum_{[q_1],[q_2]\in\mathcal{Q}/\mathbb{Z}^d} z(\beta,q_1)\, z(\beta,q_2)\, K_{q_1,q_2} \sum_{z_1\in\mathbb{Z}^d} \nabla G(z_1)\cdot(n_{q_2}) \times \nabla G(x-z_1)\cdot(n_{q_1}) \\
&\quad + 4\pi^2 \sum_{[q_1],[q_2]\in\mathcal{Q}/\mathbb{Z}^d} z(\beta,q_1)\, z(\beta,q_2)\, O\left(\frac{C_{q_1,q_2}}{|x|^{d-2+\gamma'}}\right) \\
&= \sum_{i,j,k,l=1}^d c_{ij}c_{kl} \sum_{z_1,z_2\in\mathbb{Z}^d} \nabla_i G(z_1)\nabla_j\nabla_k G(z_1-z_2)\nabla_l G(x-z_2) \\
&\quad + \sum_{i,j=1}^d K_{i,j} \sum_{z_1\in\mathbb{Z}^d} \nabla_i G(z_1)\nabla_j G(x-z_1) + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right),
\end{aligned}$$

where we have set

$$c_{ij} := 4\pi^2 \sum_{[q]\in\mathcal{Q}/\mathbb{Z}^d} z(\beta,q)\, c_{ij}^q \quad \text{and} \quad K_{i,j} := 4\pi^2 \sum_{[q_1],[q_2]\in\mathcal{Q}/\mathbb{Z}^d} z(\beta,q_1)\, z(\beta,q_2)\, K_{q_1,q_2}\left[(n_{q_1})\cdot e_i\right]\times\left[(n_{q_1})\cdot e_j\right],$$

which are well-defined by the exponential decay of the coefficient $z(\beta,q)$.

We then simplify the expansion (4.65) by noting that, since the measure μ_β is invariant under the symmetries and rotations of the lattice $\mathbb{Z}^d$, the function $x \mapsto \sum_{y\in\mathbb{Z}^d} \left\langle Q_x(y)\mathcal{U}(y,\cdot)\right\rangle_{\mu_\beta}$ satisfies the same invariance property. It is proved in Proposition 8.4 in Section 8 that this invariance property combined with the expansion (4.65) implies that there exists a constant $c := c(d,\beta)$ such that

$$\sum_{y\in\mathbb{Z}^d} \left\langle Q_x(y)\mathcal{U}(y,\cdot)\right\rangle_{\mu_\beta} = \frac{c}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\gamma'}}\right).$$

This is precisely the expansion (4.35). The proof of Theorem 1 is complete.

5. Regularity theory for low temperature dual Villain model

In this section, we study the regularity properties of the solutions of the Helffer-Sjöstrand operator

$$\mathcal{L} := -\Delta_\phi - \frac{1}{2\beta}\Delta + \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1} + \sum_{q\in\mathcal{Q}}\nabla_q^* \cdot \mathbf{a}_q \nabla_q, \tag{5.1}$$

where we recall the notation, for each charge $q \in \mathcal{Q}$,

$$\nabla_q^* \cdot \mathbf{a}_q \nabla_q u = z(\beta, q)\cos(2\pi(\phi, q))(u,q)\,q. \tag{5.2}$$

We decompose this operator into two terms: the Witten Laplacian $-\Delta_\phi$ which acts on the field ϕ and the spatial term $\mathcal{L}_{\mathrm{spat}}$ defined by the formula

$$\mathcal{L}_{\mathrm{spat}} := -\frac{1}{2\beta}\Delta + \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1} + \sum_{q\in\mathcal{Q}}\nabla_q^* \cdot \mathbf{a}_q \nabla_q.$$

The operator $\mathcal{L}_{\mathrm{spat}}$ is uniformly elliptic. The purpose of this section is to apply the techniques from the theory of elliptic regularity to understand the large-scale behavior of the solutions of the equation $\mathcal{L}u = 0$. We study three types of objects:

- In Sections 5.1 and 5.2, we study the solutions of the equation $\mathcal{L}u = 0$. We establish a Caccioppoli inequality (Proposition 5.1) and $C^{0,1-\varepsilon}$-regularity estimates (Proposition 5.6);
- In Section 5.3, we study the Helffer-Sjöstrand Green's matrix and heat kernel. We prove Gaussian bounds on the heat kernel, decay estimates on the Green's matrix, and $C^{0,1-\varepsilon}$-regularity estimates for both functions;
- In Section 5.4, we introduce the last important tool in the proof of Theorem 1: the second-order Helffer-Sjöstrand equation. This equation is used to understand how a solution of the Helffer-Sjöstrand equation depends on the underlying field ϕ, and is used to compute convariances of the form $\mathrm{cov}[u, X]$, where X is an explicit random variable (depending on ϕ) and u is a solution of the Helffer-Sjöstrand equation;

Let us give a few comments and heuristic of the proofs presented in this section. The demonstrations rely on two main ingredients:

- If we decompose the Helffer-Sjöstrand operator $\mathcal{L}$ as follows

$$\mathcal{L} = \underbrace{-\Delta_\phi - \frac{1}{2\beta}\Delta}_{\mathcal{L}_0} + \underbrace{\frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1} + \sum_{q\in\mathcal{Q}}\nabla_q^* \cdot \mathbf{a}_q \nabla_q}_{\mathcal{L}_{\mathrm{pert}}}, \tag{5.3}$$

  then the operator $\mathcal{L}_{\mathrm{pert}}$ is a perturbation of $\mathcal{L}_0$ if the inverse temperature β is large enough. The operator $\mathcal{L}_0$ has properties similar to the ones of the Laplacian and a complete regularity theory is available. The strategy to obtain regularity estimates relies on *Schauder theory* (see [66, Section 3]): since the operator $\mathcal{L}_{\mathrm{pert}}$ is a perturbation of the operator $\mathcal{L}_0$, one can prove that each solution of the equation $\mathcal{L}u = 0$ is well-approximated on every scale by a solution $\overline{u}$ of the equation $\mathcal{L}_0\overline{u} = 0$. One can then borrow the regularity of the function $\overline{u}$ and transfer it to the function u. This process causes a deterioration of the regularity for the function u: one obtains a $C^{0,1-\varepsilon}$-regularity theory for the solutions of the system $\mathcal{L}u = 0$, for some strictly positive exponent ε. The size of the exponent ε depends on the size of the perturbative term and thus on the inverse temperature β; it tends to 0 as β tends to infinity.
- The second ingredient is the dynamical solvability of Proposition 3.12 which allows to express the Helffer-Sjöstrand Green's matrix as the integral over time of a heat-kernel associated to a parabolic, time-dependent, uniformly elliptic system of equations. The system we obtain is a small perturbation of the standard discrete heat equation, and we can apply the Schauder regularity theory described in the previous item to prove regularity properties on the heat-kernel (e.g., Nash-Aronson estimate, $C^{0,1-\varepsilon}$-regularity estimates). We then transfer these properties to the Helffer-Sjöstrand Green's matrix by an integration over time.

We complete the introduction of this section by mentioning that we need to keep track of the dependence of the constants on the inverse temperature β, since one of our objectives is to prove that the regularity exponent ε tends to 0 as the inverse temperature β tends to infinity. The constants are thus only allowed to depend on the dimension.

5.1. **Caccioppoli inequality for the solutions of the Helffer-Sjöstrand equation.** In this section, we prove a Caccioppoli inequality for the operator $\mathcal{L}$, the proof follows the standard technique but some technical difficulties have to be taken into account due the infinite range of the operator $\mathcal{L}$. In particular, the result obtained is slightly different from the one of the standard Caccioppoli inequality: there is a long range term in the right sides of (5.5) and (5.6). Since the coefficients of the operator $\mathcal{L}$ decay exponentially fast, the long range terms in the right sides of (5.5) and (5.6) exhibit the same decay. Before stating the result, we recall the notation for the average (over the space variable) of a function over a ball: for $u : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$,

$$(u)_{B_R} : \phi \mapsto \frac{1}{|B_R|} \sum_{x \in B_R} u(x, \phi).$$

Proposition 5.1 (Caccioppoli inequality)**.** *Fix a radius $R \geq 1$ and let $u : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ be a solution of the Helffer-Sjöstrand equation*

$$\mathcal{L}u = 0 \text{ in } B_{2R} \times \Omega. \tag{5.4}$$

Then there exist a constant $C := C(d) < \infty$ and an exponent $c := c(d) > 0$ such that the following estimates hold

$$\beta \sum_{y \in \mathbb{Z}^d} \|\partial_y u\|_{\underline{L}^2(B_R, \mu_\beta)} + \|\nabla u\|_{\underline{L}^2(B_R, \mu_\beta)} \leq \frac{C}{R} \|u\|_{\underline{L}^2(B_{2R}, \mu_\beta)} + \sum_{x \in \mathbb{Z}^d \setminus B_{2R}} e^{-c(\ln \beta)|x|} \|u(x, \cdot)\|_{L^2(\mu_\beta)}, \tag{5.5}$$

and

$$\|\nabla u\|_{\underline{L}^2(B_R, \mu_\beta)} \leq \frac{C}{R} \left\| u - (u)_{B_{2R}} \right\|_{\underline{L}^2(B_{2R}, \mu_\beta)} + \sum_{x \in \mathbb{Z}^d} e^{-c(\ln \beta)(R \vee |x|)} \|u(x, \cdot)\|_{L^2(\mu_\beta)}. \tag{5.6}$$

Remark 5.2. The two long range terms in the right sides of (5.5) and (5.6) are error terms which are small and are caused by the infinite range of the operator $\mathcal{L}_{\mathrm{spat}}$. They decay exponentially fast and are typically of order e^{-R}.

Proof. The argument follows the standard outline of the proof of the Caccioppoli inequality; a number of technical details, pertaining to the interation of the Laplacian and the discrete differential forms need to be taken into account in the analysis. Since the argument does not contain any new idea regarding the method, we omit it here and refer to the long version of this article for the details ([36, Chapter 5, Proposition 1.1]). □

5.2. **Regularity theory for the Helffer-Sjöstrand operator.** The purpose of this section is to prove the $C^{0,1-\varepsilon}$-regularity of the solutions of the Helffer-Sjöstrand equation (5.1). The result is stated in Proposition 5.6.

The proof relies on Schauder theory; as is explained in (5.3), the strategy is to decompose the Helffer-Sjöstrand operator $\mathcal{L}$ into two terms: the operators $\mathcal{L}_0$ and $\mathcal{L}_{\mathrm{pert}}$. The operator $\mathcal{L}_0$ is the leading order term. For this operator a $C^{0,1}$-regularity theory is available. This result is stated in Proposition 5.3 and the proof is essentially equivalent to the standard proof of the regularity for harmonic functions.

The second operator $\mathcal{L}_{\mathrm{pert}}$ is a perturbative term; it is small when the inverse temperature β is large. The strategy is to argue that any solution u of the Helffer-Sjöstrand equation is well-approximated on every scale by a solution $\overline{u}$ of the equation $-\Delta_\phi \overline{u} + \frac{1}{2\beta} \Delta \overline{u} = 0$ and to transfer the regularity of the function $\overline{u}$ to the solution u. This section can be decomposed into three propositions:

- Proposition 5.3 establishes a regularity theory for the solutions $\overline{u}$ of the equation $\mathcal{L}_0 \overline{u} = 0$;
- Proposition 5.4 states that if a function u is well-approximated, in the sense of the estimate (5.9) below, by a solution of the equation $-\Delta_\phi \overline{u} - \frac{1}{2\beta} \Delta \overline{u} = 0$, then a $C^{0,1-\varepsilon}$-regularity estimate holds for the function u;
- Proposition 5.6 establishes the regularity for the solutions of the Helffer-Sjöstrand equation. We prove that any solution u of the equation $\mathcal{L}u = 0$ is well-approximated by a solution $\overline{u}$ of the equation $\mathcal{L}_0 \overline{u} = 0$ and apply Proposition 5.6 to conclude.

5.2.1. *Regularity theory for the operator $-\Delta_\phi - \frac{1}{2\beta}\Delta$.* In this section, we establish a regularity theory for the operator $-\Delta_\phi - \frac{1}{2\beta}\Delta$.

Proposition 5.3 (Regularity theory for the operator $-\Delta_\phi - \frac{1}{2\beta}\Delta$)**.** *Fix a radius $R > 0$, and let $\overline{u} : B_{2R} \times \Omega$ be a solution of the equation*

$$-\Delta_\phi \overline{u} - \frac{1}{2\beta} \Delta \overline{u} = 0 \text{ in } B_{2R} \times \Omega.$$

Then, for any integer $k \in \mathbb{N}$, there exists a constant $C_k < \infty$ depending on the dimension d and the integer k such that the following estimate holds

$$\sup_{x \in B_R} \left\| \nabla^k \overline{u}(x, \cdot) \right\|_{L^2(\mu_\beta)} \leq \frac{C_k}{R^{k+\frac{d}{2}}} \left\| \overline{u} - (\overline{u})_{B_{2R}} \right\|_{L^2(B_{2R}, \mu_\beta)}. \tag{5.7}$$

Proof. The proof is standard and relies on two ingredients: the Caccioppoli inequality and the observation that the spatial gradient ∇ commutes with the two Laplacians $-\Delta_\phi$ and Δ. First by the Caccioppoli inequality, one has

$$\left\| \nabla \overline{u} \right\|_{\underline{L}^2(B_R, \mu_\beta)} \leq \frac{C}{R} \left\| \overline{u} - (\overline{u})_{B_{2R}} \right\|_{\underline{L}^2(B_{2R}, \mu_\beta)}.$$

We then note that, since $\overline{u}$ is a solution of the equation $\mathcal{L}_0 \overline{u} = 0$, the gradient of u is also a solution of the equation $\mathcal{L}_0 \nabla \overline{u} = 0$. One can thus apply the Caccioppoli inequality to the gradient of $\overline{u}$ and deduce

$$\left\| \nabla^2 \overline{u} \right\|_{\underline{L}^2(B_R, \mu_\beta)} \leq \frac{C}{R} \left\| \nabla \overline{u} \right\|_{\underline{L}^2(B_{2R}, \mu_\beta)}.$$

An iteration of this argument shows that, for any integer $k \geq 1$, the $\underline{L}^2(B_R, \mu_\beta)$-norm of the iterated gradient $\nabla^k \overline{u}$ is controlled by the $\underline{L}^2(B_{2R}, \mu_\beta)$-norm of the function $\overline{u}$ with the appropriate scaling. By an application of the Sobolev embedding theorem (see [1, Section 4]), we obtain the regularity estimate (5.7). □

5.2.2. *Regularity theory for the Helffer-Sjöstrand operator.* The next proposition states that if a map u is well-approximated on every scale by a solution $\overline{u}$ of the equation $\mathcal{L}_0 \overline{u} = 0$, then the function u satisfies a $C^{0,1-\varepsilon}$-regularity estimate for some exponent ε depending only on the dimension d and the precision of the approximation. The proof follows a well-known strategy of Campanato (see e.g. [55]). The proof written below is an adaptation of the one of Hofmann and Kim [68].

Proposition 5.4. *Fix $X \geq 1$, a regularity exponent $\varepsilon > 0$, and a constant $K > 0$. There exists two constants $\delta_\varepsilon > 0$ and $C := C(d, \varepsilon) < \infty$, depending on the parameters d and ε such that the following statement holds. For any $R \geq 2X$, if a function $u \in L^2(B_R, \mu_\beta)$ satisfies the property that, for any $r \in [X, \frac{1}{2}R]$, there exists a solution $\overline{u} \in L^2(B_{2r}, \mu_\beta)$ of the equation*

$$-\Delta_\phi \overline{u} - \frac{1}{2\beta} \Delta \overline{u} = 0 \text{ in } B_{2r} \times \Omega, \tag{5.8}$$

such that

$$\left\| \nabla (u - \overline{u}) \right\|_{L^2(B_r, \mu_\beta)} \leq \delta_\varepsilon \left\| \nabla u \right\|_{L^2(B_{2r}, \mu_\beta)} + K, \tag{5.9}$$

then for every $r \in [X, R]$,

$$\left\| \nabla u \right\|_{\underline{L}^2(B_r, \mu_\beta)} \leq C \left(\frac{R}{r} \right)^\varepsilon \left\| \nabla u \right\|_{\underline{L}^2(B_R, \mu_\beta)} + CK.$$

Before starting the proof, we record the following lemma, which is a consequence of Giaquinta [55, Lemma 2.1].

Lemma 5.5. *Fix two non-negative real numbers X, R such that $R \geq 2X \geq 2$ and two non-negative constants C_0, K. For any regularity exponent $\varepsilon > 0$, there exist two constants $\delta_\varepsilon := \delta_\varepsilon(C, \varepsilon, d)$ and $C_1 := C_1(C, \varepsilon, d)$ such that the following statement holds. If $\phi : \mathbb{R}^+ \to \mathbb{R}$ is a non-negative and non-decreasing function which satisfies the estimate, for each pair of real numbers $\rho, r \in [X, R]$ satisfying $\rho \leq r$,*

$$\phi(\rho) \leq C_0 \left(\left(\frac{\rho}{r} \right)^{\frac{d}{2}} + \delta_\varepsilon \right) \phi(r) + K, \tag{5.10}$$

then one has the estimate, for any $\rho, r \in [X, R]$ satisfying $\rho \leq r$,

$$\phi(\rho) \leq C_1 \left(\left(\frac{\rho}{r} \right)^{\frac{d}{2} - \varepsilon} \phi(r) + K \rho^{\frac{d}{2}} \right). \tag{5.11}$$

Proof. This lemma can be extracted from [55, Lemma 2.1 p86] by setting $\alpha = \frac{d}{2}$, $\beta = \frac{d}{2} - \varepsilon$ and by using that the radii R, r are larger than 1. □

Proof of Proposition 5.4. We fix a regularity exponent $\varepsilon > 0$, let $\delta_\varepsilon > 0$ be the constant provided by Lemma 5.5, and fix two radii $\rho, r \in \left[X, \frac{1}{2}R\right]$ with $\rho \leq r$. We let $\overline{u}$ be the solution of the equation (5.8) in the set $B_r \times \Omega$ such that the estimate (5.9) holds. We note that the estimate (5.9) implies the inequality $\|\nabla \overline{u}\|_{L^2(B_r, \mu_\beta)} \leq C \|\nabla u\|_{L^2(B_{2r}, \mu_\beta)} + K$. By the regularity theory for the map $\overline{u}$ established in Proposition 5.3, we have

$$\|\nabla \overline{u}\|_{L^2(B_\rho, \mu_\beta)} \leq C \left(\frac{\rho}{r}\right)^{\frac{d}{2}} \|\nabla \overline{u}\|_{L^2(B_r, \mu_\beta)}. \tag{5.12}$$

By combining the estimates (5.9) and (5.12) and the estimate on the L^2-norm of the gradient of $\overline{u}$ mentioned above, we compute

$$\begin{aligned}
\|\nabla u\|_{L^2(B_\rho, \mu_\beta)} &\leq \|\nabla (u - \overline{u})\|_{L^2(B_\rho, \mu_\beta)} + \|\nabla \overline{u}\|_{L^2(B_\rho, \mu_\beta)} \\
&\leq \|\nabla (u - \overline{u})\|_{L^2(B_r, \mu_\beta)} + \left(\frac{\rho}{r}\right)^{\frac{d}{2}} \|\nabla \overline{u}\|_{L^2(B_r, \mu_\beta)} \\
&\leq \delta_\varepsilon \|\nabla u\|_{L^2(B_{2r}, \mu_\beta)} + K + \left(\frac{\rho}{r}\right)^{\frac{d}{2}} \left(C \|\nabla u\|_{L^2(B_{2r})} + K \right) \\
&\leq C \left(\left(\frac{\rho}{r}\right)^{\frac{d}{2}} + \delta_\varepsilon \right) \|\nabla u\|_{L^2(B_{2r}, \mu_\beta)} + 2K.
\end{aligned}$$

We apply Lemma 5.5 with the function $\phi(\rho) = \|\nabla u\|_{L^2(B_\rho)}$. The inequality (5.11) with the choice $r = R$ gives, for any radius $\rho \in [X, R]$,

$$\|\nabla u\|_{L^2(B_\rho, \mu_\beta)} \leq C_1 \left(\left(\frac{\rho}{R}\right)^{\frac{d}{2} - \varepsilon} \|\nabla u\|_{L^2(B_R, \mu_\beta)} + 2K \rho^{\frac{d}{2}} \right).$$

Dividing both side of the estimate by $\rho^{\frac{d}{2}}$ completes the proof. □

We now use Propositions 5.3 and 5.4 to obtain $C^{0,1-\varepsilon}$-regularity for the solutions of the Helffer-Sjöstrand equation.

Proposition 5.6 ($C^{0,1-\varepsilon}$-regularity theory)**.** *For any regularity exponent $\varepsilon > 0$, there exists an inverse temperature $\beta_0 := \beta_0(d, \varepsilon) < \infty$ such that the following statement holds. There exist two constants $C := C(d, \varepsilon) < \infty$ and $c := c(d) > 0$ such that for any radius $R \geq 1$, any inverse temperature $\beta \geq \beta_0$, and any function $u : \mathbb{Z}^d \times \Omega \to \mathbb{R}$ solution of the equation*

$$\mathcal{L} u = 0 \text{ in } B_R \times \Omega,$$

one has the estimate

$$\|\nabla u(0, \cdot)\|_{L^2(\mu_\beta)} \leq \frac{C}{R^{1-\varepsilon}} \left\| u - (u)_{B_R} \right\|_{\underline{L}^2(B_R, \mu_\beta)} + \sum_{x \in \mathbb{Z}^d} e^{-c(\ln \beta)(R \vee |x|)} \|u(x, \cdot)\|_{L^2(\mu_\beta)}. \tag{5.13}$$

Proof. The strategy of the proof is to apply Proposition 5.4 to the function u and then to apply the Caccioppoli inequality. We fix a regularity exponent $\varepsilon > 0$, a radius $R \geq 1$, and split the argument into two steps:

- In Step 1, we prove that the map u satisfies the following property: there exist an inverse temperature $\beta_0(\varepsilon, d) < \infty$, and a constant $C := C(d) < \infty$ such that for every $\beta > \beta_0$, and every radius $r \geq (\ln R)^2$, the following estimate holds

$$\|\nabla u\|_{\underline{L}^2(B_r, \mu_\beta)} \leq C \left(\frac{R}{r}\right)^{\frac{\varepsilon}{2}} \|\nabla u\|_{\underline{L}^2(B_R, \mu_\beta)} + \sum_{x \in \mathbb{Z}^d \setminus B_R} e^{-c(\ln \beta)|x|} \|\nabla u(x, \cdot)\|_{L^2(\mu_\beta)}. \tag{5.14}$$

- In Step 2, we deduce from (5.14) and the Caccioppoli inequality stated in Proposition 5.1, the pointwise estimate (5.13).

Step 1. To prove the estimate (5.14), the strategy is to apply Proposition 5.4. To this end, we set $X := (\ln R)^2$, and fix a radius $r \in \left[X, \frac{1}{2}R\right]$. We then define the function $\overline{u}$ to be the solution of the boundary value problem

$$\begin{cases} -\Delta_\phi \overline{u} - \dfrac{1}{2\beta} \Delta \overline{u} = 0 \text{ in } B_r \times \Omega, \\ \qquad\qquad \overline{u} = u \text{ on } \partial B_r \times \Omega. \end{cases} \tag{5.15}$$

We first prove that the map $\overline{u}$ is a good approximation of the map u. Specifically, we prove that there exist two constants $C := C(d) < \infty$ and $c := c(d) > 0$ such that

$$\text{(5.16)} \quad \left\| \nabla (u - \overline{u}) \right\|_{L^2(B_r, \mu_\beta)} \leq \frac{C}{\beta^{\frac{1}{2}}} \left\| \nabla u \right\|_{L^2(B_{2r}, \mu_\beta)} + C e^{-c(\ln \beta)(\ln R)^2} \left\| \nabla u \right\|_{L^2(B_R, \mu_\beta)} + C \sum_{x \in \mathbb{Z}^d \setminus B_R} e^{-c(\ln \beta)|x|} \left\| \nabla u(x, \cdot) \right\|_{L^2(\mu_\beta)}.$$

To prove the estimate (5.16), we note that the map $\overline{u} - u$ is a solution of the following system of equations

$$\text{(5.17)} \quad \begin{cases} -\Delta_\phi (\overline{u} - u) - \dfrac{1}{2\beta} \Delta (\overline{u} - u) = -\dfrac{1}{2\beta} \displaystyle\sum_{n \geq 1} \dfrac{1}{\beta^{\frac{n}{2}}} (-\Delta)^{n+1} u - \sum_{q \in \mathcal{Q}} \nabla_q^* \cdot \mathbf{a}_q \nabla_q u & \text{in } B_r \times \Omega, \\ \overline{u} - u = 0 & \text{on } \partial B_r \times \Omega. \end{cases}$$

We extend the function $(\overline{u} - u)$ by 0 outside the ball B_R so that it is defined on the entire space $\mathbb{Z}^d$ and use it as a test function in the system (5.17). We obtain

$$\text{(5.18)} \quad \sum_{y \in \mathbb{Z}^d} \left\| \partial_y (\overline{u} - u) \right\|_{L^2(B_r, \mu_\beta)}^2 + \frac{1}{2\beta} \left\| \nabla (\overline{u} - u) \right\|_{L^2(B_r, \mu_\beta)}^2 = -\underbrace{\frac{1}{2\beta} \sum_{n \geq 1} \frac{1}{\beta^{\frac{n}{2}}} \sum_{x \in \mathbb{Z}^d} \left\langle \nabla^{n+1} u(x, \cdot) \cdot \nabla^{n+1} (\overline{u} - u)(x, \cdot) \right\rangle_{\mu_\beta}}_{(5.18)-(i)} - \underbrace{\sum_{q \in \mathcal{Q}} \left\langle \nabla_q u \cdot \mathbf{a}_q \nabla_q (\overline{u} - u) \right\rangle_{\mu_\beta}}_{(5.18)-(ii)}.$$

The terms (5.18)-(i) and (5.18)-(ii) are perturbative terms which can be proved to be small by using the two following ingredients:

- The discrete gradient is a bounded operator and the inverse temperature β is chosen large, this is used to estimate the term (5.18)-(i);
- The coefficient $\mathbf{a}_q$ satisfy the upper bound $|\mathbf{a}_q| \leq e^{-c\sqrt{\beta}\|q\|_1}$, this is used to estimate the term (5.18)-(ii).

We omit the technical details here which can be found in the long version of this article ([36, Chapter 5, Proposition 2.4]); the result we obtain is the one stated in (5.16).

We complete Step 1 by proving that the estimate (5.16) implies the estimate (5.14). We consider the regularity exponent ε fixed at the beginning of the proof and the parameter $\delta_{\frac{\varepsilon}{2}}$ provided by Proposition 5.4 (associated with the exponent $\frac{\varepsilon}{2}$). We let $C := C(d) < \infty$ and $c := c(d) > 0$ be the constants which appear in the inequality (5.16) and set

$$X := (\ln R)^2 \text{ and } K := C e^{-c \ln \beta (\ln R)^2} \left\| \nabla u \right\|_{L^2(B_R, \mu_\beta)} + C \sum_{x \in \mathbb{Z}^d \setminus B_R} e^{-c(\ln \beta)|x|} \left\| \nabla u(x, \cdot) \right\|_{L^2(\mu_\beta)}.$$

An application of Proposition 5.4 shows the inequality: for any radius $r \in [X, R]$,

$$\text{(5.19)} \quad \left\| \nabla u \right\|_{\underline{L}^2(B_r, \mu_\beta)} \leq C \left(\frac{R}{r} \right)^{\frac{\varepsilon}{2}} \left\| \nabla u \right\|_{\underline{L}^2(B_R, \mu_\beta)} + C e^{-c \ln \beta (\ln R)^2} \left\| \nabla u \right\|_{L^2(B_R, \mu_\beta)} + C \sum_{x \in \mathbb{Z}^d \setminus B_R} e^{-c(\ln \beta)|x|} \left\| \nabla u(x, \cdot) \right\|_{L^2(\mu_\beta)}.$$

We then note that the exponential term $e^{-c(\ln \beta)(\ln R)^2}$ decays faster than any power of R, so the second term on the right side of (5.19) can be bounded from above by the first term on the right side. This completes the proof of the inequality (5.14).

Step 2. We select $r = (\ln R)^2$, apply the Caccioppoli inequality to estimate the right side of the inequality (5.14), and use that the discrete gradient is a bounded operator to replace the term $\left\| \nabla u(x, \cdot) \right\|_{L^2(\mu_\beta)}$ by $\sum_{y \sim x} \left\| u(y, \cdot) \right\|_{L^2(\mu_\beta)}$. We obtain

$$\left\| \nabla u \right\|_{\underline{L}^2\left(B_{(\ln R)^2}, \mu_\beta\right)} \leq C \left(\frac{R}{(\ln R)^2} \right)^{\frac{\varepsilon}{2}} \frac{1}{R} \left\| u - (u)_{B_R} \right\|_{\underline{L}^2(B_R, \mu_\beta)} + C \sum_{x \in \mathbb{Z}^d} e^{-c(\ln \beta)(R \vee |x|)} \left\| u(x, \cdot) \right\|_{L^2(\mu_\beta)}.$$

We apply the discrete $L^\infty - L^2$ -estimate

$$\left\| \nabla u(0) \right\|_{L^2(\mu_\beta)} \leq \left\| \nabla u \right\|_{L^2\left(B_{(\ln R)^2}, \mu_\beta\right)} \leq (\ln R)^d \left\| \nabla u \right\|_{\underline{L}^2\left(B_{(\ln R)^2}, \mu_\beta\right)}.$$

We then combine the two previous displays and the estimate $(\ln R)^d \leq CR^{\frac{\varepsilon}{2}}$ to obtain the inequality (5.13). The proof of Proposition 5.6 is complete. □

5.3. **Nash-Aronson estimate and regularity theory for heat kernels.** In this section, we study the dynamical solvability of the Helffer-Sjöstrand equation and prove the bounds stated in Section 3 (and specifically Proposition 3.12).

We first recall the definition of the Langevin dynamics associated wiht the Gibbs measure μ_β. Given a field $\phi \in \Omega$, we let $(\phi_t)_{t\geq 0}$ be the diffusion process evolving according to the Langevin dynamics

$$\begin{cases} d\phi_t(x) = \frac{1}{2\beta}\Delta\phi_t(x)dt - \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1}\phi_t(x)dt + \sum_{q\in\mathcal{Q}}\left(\nabla_q^*\cdot\mathbf{a}_q(\phi_t)\nabla_q\phi_t\right)(x)dt + \sqrt{2}dB_t(x), \\ \phi_0(x) = \phi(x), \end{cases} \tag{5.20}$$

where $\left\{B_t(x) : t\geq 0,\, x\in\mathbb{Z}^d\right\}$ is a collection of independent normalized $\mathbb{R}^{\binom{d}{2}}$-valued independent Brownian motions. We denote by $\mathbb{P}_\phi$ the law of the dynamics $(\phi_t)_{t\geq 0}$ starting from ϕ and by $\mathbb{E}_\phi$ the expectation with respect to the measure $\mathbb{P}_\phi$. Given a realization of the dynamics, we let P^ϕ be the solution of the parabolic system

$$\begin{cases} \partial_t P^{\phi_\cdot}(\cdot,\cdot;y) - \frac{1}{2\beta}\Delta P^{\phi_\cdot}(\cdot,\cdot;y) + \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1}P^{\phi_\cdot}(\cdot,\cdot) + \sum_{q\in\mathcal{Q}}\left(\nabla_q^*\cdot\mathbf{a}_q(\phi_t)\nabla_q P^{\phi_\cdot}(\cdot,\cdot;y)\right) = 0 \text{ in } [0,\infty]\times\mathbb{Z}^d, \\ P^{\phi_\cdot}(0,\cdot;y) = \delta_y \text{ in } \mathbb{Z}^d. \end{cases}$$

The main purpose of this section is to prove upper bounds on the heat kernel $P^{\phi_\cdot}$ and on its spatial derivatives. We introduce the following definition. For each constant $C > 0$, we let Φ_C be the function defined from $(0,\infty)\times\mathbb{Z}^d$ to $\mathbb{R}$ by the formula, for each pair $(t,x)\in(0,\infty)\times\mathbb{Z}^d$,

$$\Phi_C(t,x) = \begin{cases} t^{-\frac{d}{2}}\exp\left(-\frac{|x|^2}{Ct}\right) \text{ if } |x|\leq t, \\ \exp\left(-\frac{|x|}{C}\right) \text{ if } |x|\geq t. \end{cases} \tag{5.21}$$

The next proposition is the main result of this section.

Proposition 5.7 (Gaussian bounds and $C^{0,1-\varepsilon}$-regularity for the heat kernel)**.** *For any regularity exponent $\varepsilon > 0$, there exists an inverse temperature $\beta_1(d,\varepsilon) < \infty$, and a constant $C := C(d,\varepsilon) < \infty$ such that for every $\beta > \beta_1$, for any realization of the dynamics $(\phi_t)_{t\geq 0}$, any $(t,x,y)\in[1,\infty]\times\mathbb{Z}^d\times\mathbb{Z}^d$,*

$$\left|P^{\phi_\cdot}(t,x;y)\right| \leq C\Phi_C\left(\frac{t}{\beta}, x-y\right). \tag{5.22}$$

Moreover, one has the following $C^{0,1-\varepsilon}$-regularity estimate on the gradient of the heat kernel

$$\left|\nabla_x P^{\phi_\cdot}(t,x;y)\right| \leq C\left(\frac{\beta}{t}\right)^{\frac{1}{2}-\varepsilon}\Phi_C\left(\frac{t}{\beta}, x-y\right), \tag{5.23}$$

and on the mixed derivative of the heat kernel

$$\left|\nabla_x\nabla_y P^{\phi_\cdot}(t,x;y)\right| \leq C\left(\frac{\beta}{t}\right)^{1-\varepsilon}\Phi_C\left(\frac{t}{\beta}, x-y\right). \tag{5.24}$$

Remark 5.8. Due to the discrete setting of the problem and the infinite range of the operator $\mathcal{L}$, the heat kernel does not have Gaussian decay when the value $|x|$ tends to infinity. Instead it decays exponentially fast; this justifies the introduction of the function Φ_C.

Remark 5.9. For later use, we need to keep track of the dependence of the constants in the inverse temperature β.

In the rest of Section 5.3, we give an outline of the proof of Proposition 5.7; the details of the argument can be found in the long version of this article ([36, Chapter 5, Sections 3.1 and 3.2]).

Outline of the proof of Proposition 5.7. Gaussian bounds on the heat kernel are usually a consequence of the Nash-Aronson estimate (see [11, 43]) for uniformly elliptic operators. This result cannot be applied here since the operator $\partial_t + \mathcal{L}_{\mathrm{spat}}^{\phi_t}$ is a parabolic system of equations, and we refer to the counter-example of De Giorgi [40] disproving the Liouville property and the $C^{0,\alpha}$-regularity theory for systems of elliptic equations. To prove Gaussian bounds and regularity on the heat kernel, we use a different strategy and proceed according to the following outline:

(1) We use that the elliptic operator $\mathcal{L}_{\mathrm{spat}}^{\phi_t}$ is a perturbation of the Laplacian to establish $C^{0,1-\varepsilon}$-regularity for the solutions of the system

$$\partial_t u + \mathcal{L}_{\mathrm{spat}}^{\phi_t} u = 0; \tag{5.25}$$

(2) We use the $C^{0,1-\varepsilon}$-regularity and an interpolation argument to obtain L^∞-bounds on the solutions of the equation (5.25). More precisely, we prove that every solution of the system (5.25) in the parabolic cylinder Q_{2r} satisfies the pointwise estimate

$$\|u\|_{L^\infty(Q_r)} \le C \|u\|_{\underline{L}^2(Q_{2r})} + \int_{-r^2}^0 \sum_{x\in\mathbb{Z}^d\setminus B_r} e^{-c(\ln\beta)|x|} |u(t,x)|^2 \, dt;$$

(3) We prove that the solutions of the adjoint of the parabolic operator $\partial_t + \mathcal{L}_{\mathrm{spat}}^{\phi_t}$ satisfies the same pointwise estimate;

(4) We use the pointwise regularity estimates and the technique of Fabes and Stroock [43], which is based on the technique of Davies [37, 38] (see also the article of Hofmann and Kim [68]) to establish the Gaussian bounds on the heat kernel stated in (5.22);

(5) We combine the Gaussian bounds on the heat kernel with the $C^{1-\varepsilon}$-regularity theory for the solutions of (5.25) to obtain the upper bounds on the gradient and mixed derivative of the heat kernel stated in (5.23) and (5.24).

□

5.4. **Definition and regularity for the second-order Helffer-Sjöstrand operator.** We introduce and study the second-order Helffer-Sjöstrand operator. We mention that this operator was initially introduced in the article of Conlon and Spencer [29], and the general underlying philosophy is closely related to the one developed in stochastic homogenization in [60, 61, 58, 59].

Let us fix a function $G \in C_c^\infty(\mathbb{Z}^d \times \Omega)$ and let u be the solution of the Helffer-Sjöstrand equation

$$\mathcal{L}u = G \text{ in } \Omega \times \mathbb{Z}^d.$$

As mentioned above, in Section 8.3, we will have to estimate covariances of the form $\mathrm{cov}[u, X]$, where X is an explicit functional of the field ϕ. By the Helffer-Sjöstrand representation formula (Proposition 3.13), it is sufficient to understand the properties of the functions $\partial_x u$, for $x \in \mathbb{Z}^d$.

The strategy is then to find an equation satisfied by the map $\partial_x u$. In this direction, we may apply (formally) the operator ∂_x to both the left and right hand sides of the identity $\mathcal{L}u = G$. We obtain the identity

$$\partial_x - \Delta_\phi u - \frac{1}{2\beta}\Delta\partial_x u + \frac{1}{2\beta}\sum_{n\ge1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1}\partial_x u + \partial_x\left(\sum_{q\in\mathcal{Q}} \nabla_q^* \cdot \mathbf{a}_q \nabla_q u\right) = \partial_x G. \tag{5.26}$$

To go further in the computation, let us introduce the following notation:

- We define the function $v, h : \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega \mapsto \mathbb{R}^{\binom{d}{2}\times\binom{d}{2}}$ by the formulae, for each for $(x,y,\phi) \in \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega$, $v(x,y,\phi) = \partial_x u(y,\phi)$ and $h(x,y,\phi) = \partial_x G(y,\phi)$.
- Given a map $h : \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega \mapsto \mathbb{R}^{\binom{d}{2}\times\binom{d}{2}}$, we denote by Δ_x the spatial Laplacian in the first variable and by Δ_y the Laplacian in the second variable. We also denote by $\sum_{q_x\in\mathcal{Q}} \nabla_{q_x}^* \cdot \mathbf{a}_{q_x} \nabla_{q_x} h$ and by $\sum_{q_y\in\mathcal{Q}} \nabla_{q_y}^* \cdot \mathbf{a}_{q_y} \nabla_{q_y} h$ the operators

$$\sum_{q_x\in\mathcal{Q}} \nabla_{q_x}^* \cdot \mathbf{a}_{q_x} \nabla_{q_x} h : (x,y,\phi) \mapsto \sum_{q\in\mathcal{Q}} \mathbf{a}_q(\phi)\left(h(\cdot,y,\phi), q\right) q(x)$$

and

$$\sum_{q_y\in\mathcal{Q}} \nabla_{q_y}^* \cdot \mathbf{a}_{q_y} \nabla_{q_y} h : (x,y,\phi) \mapsto \sum_{q\in\mathcal{Q}} \mathbf{a}_q(\phi)\left(h(x,\cdot,\phi), q\right) q(y).$$

- Finally, we denote by $\mathcal{L}_{\mathrm{spat},x}$ and $\mathcal{L}_{\mathrm{spat},y}$ the operators

$$\mathcal{L}_{\mathrm{spat},x} := -\frac{1}{2\beta}\Delta_x u + \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta_x)^{n+1}u + \sum_{q_x\in\mathcal{Q}}\nabla_{q_x}^* \cdot \mathbf{a}_{q_x}\nabla_{q_x}u,$$

and

$$\mathcal{L}_{\mathrm{spat},y} := -\frac{1}{2\beta}\Delta_y u + \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta_y)^{n+1}u + \sum_{q_y\in\mathcal{Q}}\nabla_{q_y}^* \cdot \mathbf{a}_{q_y}\nabla_{q_y}u.$$

The term $\partial_\cdot - \Delta_\phi u$ can be computed by using the same strategy as the one used to derive the Helffer-Sjöstrand equation in Section 3.4, and we obtain, for each $(x,y,\phi)\in\mathbb{Z}^d\times\mathbb{Z}^d\times\Omega$,

(5.27)
$$\begin{aligned}\partial_x - \Delta_\phi u(y,\phi) &= -\Delta_\phi v(x,y,\phi) - \frac{1}{2\beta}\Delta_x v(x,y,\phi) + \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta_x)^{n+1}v(x,y,\phi) + \sum_{q_x\in\mathcal{Q}}\nabla_{q_x}^*\cdot\mathbf{a}_{q_x}\nabla_{q_x}v(x,y,\phi)\\ &= -\Delta_\phi v(x,y,\phi) + \mathcal{L}_{\mathrm{spat},x}v(x,y,\phi)\end{aligned}$$

The term $\partial_x\left(\sum_{q\in\mathcal{Q}}\nabla_q^*\cdot\mathbf{a}_q\nabla_q u\right)$ can be computed by using the exact formula stated in (5.2):

$$\begin{aligned}(5.28)\quad \partial_x\left(\sum_{q\in\mathcal{Q}}\nabla_q^*\cdot\mathbf{a}_q\nabla_q u(y,\phi)\right) &= \partial_x\left(\sum_{q\in\mathcal{Q}}\mathbf{a}_q\,(u,q)\,q(y)\right)\\ &= \sum_{q\in\mathcal{Q}}\partial_x\mathbf{a}_q\,(u,q)\,q(y) + \sum_{q\in\mathcal{Q}}\mathbf{a}_q\,(v(x,\cdot,\phi),q)\,q\\ &= \sum_{q\in\mathcal{Q}}2\pi z\,(\beta,q)\cos\left(2\pi\,(\phi,q)\right)(u,q)\,q(x)\otimes q(y) + \sum_{q_y\in\mathcal{Q}}\nabla_{q_y}^*\mathbf{a}_{q_y}\nabla_{q_y}v(x,y,\phi).\end{aligned}$$

Combining the identities (5.26), (5.27), and (5.28), we obtain that the map v solves the equation

$$\begin{aligned}(5.29)\quad -\Delta_\phi v(x,y,\phi) + \mathcal{L}_{\mathrm{spat},x}v(x,y,\phi) + \mathcal{L}_{\mathrm{spat},y}v(x,y,\phi)&\\ = -\sum_{q\in\mathcal{Q}}2\pi z\,(\beta,q)\cos\left(2\pi\,(\phi,q)\right)(u,q)\,q(x)\otimes q(y) + \partial_x G(y,\phi).&\end{aligned}$$

This equality can be rigorously justified using the arguments of Section 3.4 and of [81, 54]. The identity (5.29) motivates the definition of the second-order Helffer-Sjöstrand operator acting on functions defined on $\Omega\times\mathbb{Z}^d\times\mathbb{Z}^d$ and valued in $\mathbb{R}^{\binom{d}{2}\times\binom{d}{2}}$

$$\mathcal{L}_{\mathrm{sec}} := -\Delta_\phi + \mathcal{L}_{\mathrm{spat},x} + \mathcal{L}_{\mathrm{spat},y}. \tag{5.30}$$

This operator has the same properties than the one satisfied by the Helffer-Sjöstrand operator and listed in Section 3. In particular, it can be solved using the variational techniques of Proposition 3.8 or the dynamical interpretation of Proposition 3.12.

As it was the case for the Helffer-Sjöstrand operator, it is natural to consider the Green's function associated with the second-order operator It is introduced in the following definition.

Definition 5.10 (Green matrix for the second-order Helffer-Sjöstrand equation)**.** For any $(y,y_1)\in\mathbb{Z}^d\times\mathbb{Z}^d$, we let $\delta_{(y,y_1)}:\mathbb{Z}^d\times\mathbb{Z}^d\to\mathbb{R}^{\binom{d}{2}^2\times\binom{d}{2}^2}$ be the Dirac mass defined by the formula

$$\delta_{(y,y_1)}((x,x_1)) := \left(\mathbb{1}_{\{(x,x_1)=(y,y_1)\}}\cdot\mathbb{1}_{\{i=j\}}\right)_{1\leq i,j\leq\binom{d}{2}^2}.$$

For any function $\mathbf{f}:\Omega\to\mathbb{R}$ satisfying $\mathbf{f}\in L^2(\mu_\beta)$, we define the Green's function associated with the second-order equation $\mathcal{G}_{\mathrm{sec},\mathbf{f}}:\Omega\times\mathbb{Z}^d\times\mathbb{Z}^d\to\mathbb{R}^{\binom{d}{2}^2\times\binom{d}{2}^2}$ according to the formula

$$\mathcal{L}_{\mathrm{sec}}\mathcal{G}_{\mathrm{sec},\mathbf{f}} = \mathbf{f}\delta_{(y,y_1)} \text{ in } \Omega\times\mathbb{Z}^d\times\mathbb{Z}^d. \tag{5.31}$$

As in Proposition 3.14 of Section 3.4.4, the existence of the Green's function can be established variationally, by applying the Gagliardo-Nirenberg-Sobolev inequality.

As in Section 3.4.4, one can solve the second-order Helffer-Sjöstrand equation dynamically as stated in the following proposition.

Proposition 5.11. *Fix* $\mathbf{f} \in L^2(\mu_\beta)$ *and* $(y, y_1) \in \mathbb{Z}^d \times \mathbb{Z}^d$. *The Green's matrix* $\mathcal{G}_{\mathrm{sec},\mathbf{f}}(\cdot; y, y_1)$ *satisfies the identity*

$$\mathcal{G}_{\mathrm{sec},\mathbf{f}}\left(x, x_1, \phi; y, y_1\right) := \int_0^\infty \mathbb{E}_\phi\left[\mathbf{f}(\phi_t) P^{\phi_\cdot}_{\mathrm{sec}}(t, y, y_1; x, x_1)\right] dt,$$

where $P^{\phi_\cdot}_{\mathrm{sec}}(\cdot, \cdot; x, x_1)$ *is the solution of the system of equations,*

$$\begin{cases} \partial_t P^{\phi_\cdot}_{\mathrm{sec}}\left(\cdot, \cdot, \cdot; x, x_1\right) + \left(\mathcal{L}^{\phi_t}_{\mathrm{spat},x} + \mathcal{L}^{\phi_t}_{\mathrm{spat},y}\right) P^{\phi_\cdot}_{\mathrm{sec}}\left(\cdot, \cdot, \cdot; x, x_1\right) = 0 & \text{in } (0, \infty) \times \mathbb{Z}^d \times \mathbb{Z}^d, \\ P^{\phi_\cdot}_{\mathrm{sec}}\left(0, \cdot, \cdot; x, x_1\right) = \delta_{(x, x_1)} & \text{in } \mathbb{Z}^d \times \mathbb{Z}^d. \end{cases}$$

5.4.1. *Gaussian bounds and regularity estimates for the Green's matrix.* In this section, we study the decay properties of the Green's matrix associated with the second-order Helffer-Sjöstrand operator.

The operator $\mathcal{L}^{\phi_\cdot}_{\mathrm{spat},x} + \mathcal{L}^{\phi_\cdot}_{\mathrm{spat},y}$ is a uniformly elliptic operator on the $2d$-dimensional space $\mathbb{Z}^d \times \mathbb{Z}^d$. If the inverse temperature β is chosen large enough, then this operator is a perturbation of the $2d$-dimensional Laplacian $\Delta_x + \Delta_y$. Hence the same arguments as in Section 5.3 can be used to prove Gaussian bounds and $C^{0,1-\varepsilon}$-regularity estimates on the heat kernel $P^{\phi_\cdot}_{\mathrm{sec}}$; the only difference is that the underlying space is $2d$-dimensional.

The result stated in Proposition 5.12 is strictly stronger than the one obtained by the previous argument since we obtain estimates on the triple and quadruple gradients of the heat kernel. These properties are obtained by making use of the specific structure of the problem and relies on the observation that the elliptic operators $\mathcal{L}_{\mathrm{spat},x}$ and $\mathcal{L}_{\mathrm{spat},y}$ only act on the x and y variables respectively. This remark implies that these operators commute and thus the heat kernel $P^{\phi_\cdot}_{\mathrm{sec}}$ can be factorised as follows.

If we let $\delta_y : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}^2 \times \binom{d}{2}^2}$ be the Dirac mass defined by the formula

$$\delta_y(x) := \left(\mathbb{1}_{\{x=y\}} \cdot \mathbb{1}_{\{i=j\}}\right)_{1 \le i, j \le \binom{d}{2}^2}.$$

and consider the solution $P^{\phi_\cdot}_{\mathrm{sec},x} : \mathbb{Z}^d \mapsto \mathbb{R}^{\binom{d}{2}^2 \times \binom{d}{2}^2}$ of the system of equations,

$$\begin{cases} \partial_t P^{\phi_\cdot}_{\mathrm{sec},x}\left(\cdot, \cdot; x\right) + \mathcal{L}^{\phi_t}_{\mathrm{spat},x} P^{\phi_\cdot}_{\mathrm{sec},x}\left(\cdot, \cdot; x\right) = 0 & \text{in } (0, \infty) \times \mathbb{Z}^d, \\ P^{\phi_\cdot}_{\mathrm{sec}}\left(0, \cdot, \cdot; x\right) = \delta_x & \text{in } \mathbb{Z}^d. \end{cases}$$

and define similarly the solution $P^{\phi_\cdot}_{\mathrm{sec},y}$. Then we have the identity

$$P^{\phi_\cdot}_{\mathrm{sec}}\left(t, y, y_1; x, x_1\right) = P^{\phi_\cdot}\left(t, y; x\right) P^{\phi_\cdot}\left(t, y_1; x_1\right), \tag{5.32}$$

where the product in the right-hand side refers to the product of matrices of size $\binom{d}{2}^2 \times \binom{d}{2}^2$. Thanks to this property, one can obtain additional regularity estimates on the map $P^{\phi_\cdot}_{\mathrm{sec}}$; for instance, if we denote by ∇_x, ∇_y, ∇_{x_1}, ∇_{y_1} the gradient with respect to the first, second, third, and fourth spatial variable, then we have

$$\nabla_x \nabla_y \nabla_{x_1} \nabla_{y_1} P^{\phi_\cdot}_{\mathrm{sec}}(y, y_1; x, x_1) = \nabla_x \nabla_y P^{\phi_\cdot}\left(t, y; x\right) \nabla_{x_1} \nabla_{y_1} P^{\phi_\cdot}\left(t, y_1; x_1\right). \tag{5.33}$$

The strategy is then to combine the regularity estimates proved in Proposition 5.7 with the factorization formula (5.32) to obtain additional regularity properties on the heat kernel associated with the second-order equation. The results are collected in the following proposition.

Proposition 5.12. *For any regularity exponent* $\varepsilon > 0$, *there exists an inverse temperature* $\beta_0(d, \varepsilon) < \infty$ *such that the following statement holds. For any inverse temperature* $\beta > \beta_0$ *and any realization of the dynamics* $(\phi_t)_{t \ge 0}$, *there exists a constant* $C(d, \varepsilon) < \infty$ *such that for each* $(x, y, x_1, y_1) \in \left(\mathbb{Z}^d\right)^4$, *one has the estimate*

$$\left|P^{\phi_\cdot}_{\mathrm{sec}}\left(t, x, x_1; y, y_1\right)\right| \le C \Phi_C\left(\frac{t}{\beta}, x - x_1\right) \Phi_C\left(\frac{t}{\beta}, y - y_1\right),$$

and the $C^{0,1-\varepsilon}$*-regularity estimates: if we let* $\nabla_1, \nabla_2, \nabla_3$ *and* ∇_4 *be any permutation of the set of gradients* $\nabla_x, \nabla_{x_1}, \nabla_y$ *and* ∇_{y_1}, *then one has the four inequalities:*

(i) *On the gradient of the heat kernel*

$$\left|\nabla_1 P^{\phi_\cdot}_{\mathrm{sec}}\left(t, x, x_1; y, y_1\right)\right| \le C \left(\frac{\beta}{t}\right)^{\frac{1}{2} - \varepsilon} \Phi_C\left(\frac{t}{\beta}, x - x_1\right) \Phi_C\left(\frac{t}{\beta}, y - y_1\right);$$

(ii) *On the double gradient of the heat kernel*

$$\left|\nabla_1 \nabla_2 P^{\phi.}_{\mathrm{sec}}(t, x, x_1; y, y_1)\right| \leq C \left(\frac{\beta}{t}\right)^{1-\varepsilon} \Phi_C\left(\frac{t}{\beta}, x - x_1\right) \Phi_C\left(\frac{t}{\beta}, y - y_1\right);$$

(iii) *On the triple gradient of the heat kernel*

$$\left|\nabla_1 \nabla_2 \nabla_3 P^{\phi.}_{\mathrm{sec}}(t, x, x_1; y, y_1)\right| \leq C \left(\frac{\beta}{t}\right)^{\frac{3}{2}-\varepsilon} \Phi_C\left(\frac{t}{\beta}, x - x_1\right) \Phi_C\left(\frac{t}{\beta}, y - y_1\right);$$

(iv) *On the quadruple gradient of the heat kernel*

$$\left|\nabla_1 \nabla_2 \nabla_3 \nabla_4 P^{\phi.}_{\mathrm{sec}}(t, x, x_1; y, y_1)\right| \leq C \left(\frac{\beta}{t}\right)^{2-\varepsilon} \Phi_C\left(\frac{t}{\beta}, x - x_1\right) \Phi_C\left(\frac{t}{\beta}, y - y_1\right).$$

Proposition 5.12 is obtained by combining Proposition 5.7 with the factorization identity (5.33).

From these estimates, we deduce the bounds on the elliptic Green's matrix and its gradients stated in the following proposition.

Proposition 5.13. *For any regularity exponent $\varepsilon > 0$, there exists an inverse temperature $\beta_0(d, \varepsilon) < \infty$ such that the following statement holds. For any inverse temperature $\beta > \beta_0$, there exists a constant $C(d, \varepsilon) < \infty$ such that for each $(x, y, x_1, y_1) \in (\mathbb{Z}^d)^4$, one has the estimate*

$$\left\| \mathcal{G}_{\mathrm{sec}, \mathbf{f}}(x, y, \cdot; x_1, y_1) \right\|_{L^2(\mu_\beta)} \leq \frac{C\beta \left\| \mathbf{f} \right\|_{L^2(\mu_\beta)}}{|x - x_1|^{2d-2} + |y - y_1|^{2d-2}}.$$

Additionally, for any permutation $\nabla_1, \nabla_2, \nabla_3$ and ∇_4 of the set of gradients $\nabla_x, \nabla_{x_1}, \nabla_y$ and ∇_{y_1}, one has the estimates:

(i) *On the gradient of the Green's matrix*

$$\left\| \nabla_1 \mathcal{G}_{\mathrm{sec}, \mathbf{f}}(x, y, \cdot; x_1, y_1) \right\|_{L^2(\mu_\beta)} \leq \frac{C\beta \left\| \mathbf{f} \right\|_{L^2(\mu_\beta)}}{|x - x_1|^{2d-1-\varepsilon} + |y - y_1|^{2d-1-\varepsilon}};$$

(ii) *On the double gradient of the Green's matrix*

$$\left\| \nabla_1 \nabla_2 \mathcal{G}_{\mathrm{sec}, \mathbf{f}}(x, y, \cdot; x_1, y_1) \right\|_{L^2(\mu_\beta)} \leq \frac{C\beta \left\| \mathbf{f} \right\|_{L^2(\mu_\beta)}}{|x - x_1|^{2d-\varepsilon} + |y - y_1|^{2d-\varepsilon}};$$

(iii) *On the triple gradient of the Green's matrix*

$$\left\| \nabla_1 \nabla_2 \nabla_3 \mathcal{G}_{\mathrm{sec}, \mathbf{f}}(x, y, \cdot; x_1, y_1) \right\|_{L^2(\mu_\beta)} \leq \frac{C\beta \left\| \mathbf{f} \right\|_{L^2(\mu_\beta)}}{|x - x_1|^{2d+1-\varepsilon} + |y - y_1|^{2d+1-\varepsilon}};$$

(iv) *On the quadruple gradient of the Green's matrix*

$$\left\| \nabla_1 \nabla_2 \nabla_3 \nabla_4 \mathcal{G}_{\mathrm{sec}, \mathbf{f}}(x, y, \cdot; x_1, y_1) \right\|_{L^2(\mu_\beta)} \leq \frac{C\beta \left\| \mathbf{f} \right\|_{L^2(\mu_\beta)}}{|x - x_1|^{2d+2-\varepsilon} + |y - y_1|^{2d+2-\varepsilon}}.$$

The estimates on the elliptic Green's matrix are obtained by integrating the inequalities of Proposition 5.12 over the times t in $[0, \infty)$ and applying the Cauchy-Schwarz inequality.

6. Quantitative convergence of the subadditive quantities

The objective of this section and of Section 7 is to prove Theorem 2. The strategy adopted follows the one of [8], and relies on the introduction of two subadditive energy quantities related to the variational formulation associated with the Helffer-Sjöstrand operator. The first one, denoted by $\nu(\square, p)$, represents the energy of the minimizer associated with the Dirichlet problem in a cube $\square$ with affine boundary condition $l_p(x) := p \cdot x$. The second one, denoted by $\nu^*(\square, q)$, represents the energy of the minimizer associated with the Neumann problem with boundary flux ∇l_q. These two quantities satisfy a subadditivity property with respect to the domain of integration and converge as the sidelength of the cube tends to infinity. Moreover, the quantities ν and ν^* are convex with respect to the slopes of the boundary conditions, and are approximately convex dual to each other. The main focus of this section is to prove by a multiscale argument that, as the size of the domains tends to infinity, these quantities converge to a pair of dual convex conjugate functions, and to extract from the proof a quantification of the rate of convergence.

While the general strategy comes from the theory of quantitative stochastic homogenization presented in [8], the adaptation of the techniques presented in this monograph requires to overcome three types of difficulties:

- One needs to take into account the Laplacian with respect to the ϕ-variable;
- One needs to take into account the infinite range of the operator $\mathcal{L}$;
- We need to homogenize an elliptic system instead of an elliptic PDE.

While the first point has been successfully treated in [9] to study the $\nabla\phi$ model, the last two points are intrinsic to the Coulomb gas representation of the Villain model and will be treated in this section.

This section is organized as follows. In Sections 6.1 and 6.2, we define the subadditive energy quantities ν and ν^*, and collect some of their basic properties. In Section 6.3, we obtain a quantitative rate of convergence for these quantities. In Section 6.4, we introduce a finite-volume version of first-order corrector associated with the Helffer-Sjöstrand operator $\mathcal{L}$. We use the quantitative rate of convergence of the energy ν to establish quantitative sublinearity of the corrector and to prove a quantitative estimate on the weak norm of its flux. This function and its properties are crucial to prove the quantitative homogenization of the mixed derivative of the Green's matrix in Section 7.

Throughout this entire section, we fix a regularity exponent ε which is small compared to 1 and depends only on the dimension d. We assume that the inverse temperature β is large enough so that all the results presented in Section 5 hold with the regularity exponent ε.

We complete this introduction by mentioning that in this section, the constants are only allowed to depend in the dimension d as we need to keep track of their dependence on the inverse temperature β. The objective is to prove that the quantitative rate of convergence α obtained in Proposition 6.10 and 6.28 remains bounded away from 0 as β tends to infinity.

6.1. Definition of the subadditive quantities and basic properties.

6.1.1. *Definition of the energy quantities.* Let $\square$ be a cube of $\mathbb{Z}^d$, we define the energy functional $\mathbf{E}_\square$ according to the formula, for each function $u \in H^1\left(\mathbb{Z}^d, \mu_\beta\right)$,

$$\mathbf{E}_\square[u] := \beta \sum_{y \in \mathbb{Z}^d} \left\| \partial_y u \right\|^2_{L^2(\square, \mu_\beta)} + \frac{1}{2} \left\| \nabla u \right\|^2_{L^2(\square, \mu_\beta)} + \frac{1}{2} \sum_{n \geq 1} \frac{1}{\beta^{\frac{n}{2}}} \left\| \nabla^{n+1} u \right\|^2_{L^2(\mathbb{Z}^d, \mu_\beta)} - \beta \sum_{\operatorname{supp} q \cap \square \neq \emptyset} \left\langle \nabla_q u \cdot \mathbf{a}_q \nabla_q u \right\rangle_{\mu_\beta}.$$

We introduce the bilinear form associated with the energy $\mathbf{E}_\square$: for each function $u \in H^1\left(\mathbb{Z}^d, \mu_\beta\right)$,

$$\begin{aligned}\mathbf{B}_\square[u, v] := \beta \sum_{x \in \square} \sum_{y \in \mathbb{Z}^d} \left\langle \partial_y u(x, \cdot), \partial_y v(x, \cdot) \right\rangle_{\mu_\beta} + \frac{1}{2} \sum_{x \in \square} \left\langle \nabla u(x, \cdot), \nabla v(x, \cdot) \right\rangle_{\mu_\beta} \\ + \frac{1}{2} \sum_{n \geq 1} \sum_{x \in \mathbb{Z}^d} \frac{1}{\beta^{\frac{n}{2}}} \left\langle \nabla^{n+1} u(x, \cdot), \nabla^{n+1} v(x, \cdot) \right\rangle_{\mu_\beta} - \sum_{\operatorname{supp} q \cap \square \neq \emptyset} \beta \left\langle \nabla_q u \cdot \mathbf{a}_q \nabla_q v \right\rangle_{\mu_\beta}.\end{aligned}$$

This energy and bilinear form are useful to define the energy quantity ν. To define the dual energy ν^*, we need to introduce an alternative definition of the mappings $\mathbf{E}_\square$ and $\mathbf{B}_\square$. The technical difficulty encountered is the following: one cannot consider the energy $\mathbf{E}_\square$ of a function v only defined in the cube $\square$ since the infinite range of the operator $\mathcal{L}$ requires to know the values of the function in the entire space $\mathbb{Z}^d$. To fix this issue, we restrict the summation over the set of charges $q \in \mathcal{Q}$ whose support is included in the cube $\square$, and over the sets of integers n and points x such that the value $\Delta^n v(x)$ can be computed by only knowing the values of the function v in the cube $\square$. As it will be useful later in the proofs, we also remove a boundary layer term, and we recall the definition of trimmed cube stated in (A.2) of Appendix A. We define the energy $\mathbf{E}^*_\square$ by the formula

$$\mathbf{E}^*_\square[u] := \beta \sum_{y \in \mathbb{Z}^d} \|\partial_y u\|^2_{L^2(\square, \mu_\beta)} + \frac{1}{2} \sum_{n \geq 0} \sum_{x \in \square, \mathrm{dist}(x, \partial\square) \geq n} \frac{1}{\beta^{\frac{n}{2}}} \left\| \nabla^{n+1} u(x, \cdot) \right\|^2_{L^2(\mu_\beta)} - \frac{1}{\beta^{\frac{1}{4}}} \|\nabla u\|^2_{L^2(\square \setminus \square^-, \mu_\beta)} - \beta \sum_{\mathrm{supp}\, q \subseteq \square} \langle \nabla_q u \cdot \mathbf{a}_q \nabla_q u \rangle_{\mu_\beta},$$

as well as the corresponding bilinear form $\mathbf{B}^*_\square$, for each $u, v \in H^1(\square, \mu_\beta)$,

$$\mathbf{B}^*_\square[u, v] := \beta \sum_{x \in \square} \sum_{y \in \mathbb{Z}^d} \langle \partial_y u(x, \cdot), \partial_y v(x, \cdot) \rangle_{\mu_\beta} + \frac{1}{2} \sum_{n \geq 1} \sum_{x \in \square, \mathrm{dist}(x, \partial\square) \geq n} \frac{1}{\beta^{\frac{n}{2}}} \left\langle \nabla^{n+1} u(x, \cdot), \nabla^{n+1} v(x, \cdot) \right\rangle_{\mu_\beta} - \frac{1}{\beta^{\frac{1}{4}}} \sum_{x \in \square \setminus \square^-} \langle \nabla u(x, \cdot), \nabla v(x, \cdot) \rangle_{\mu_\beta} - \beta \sum_{\mathrm{supp}\, q \subseteq \square} \langle \nabla_q u \cdot \mathbf{a}_q \nabla_q v \rangle_{\mu_\beta}.$$

Let us make a few remarks about the definition of the energy $\mathbf{E}^*_\square$.

Remark 6.1. The iterated Laplacian Δ^n has range $2n$; given a point $x \in \square$, we only consider the iteration of the Laplacian until the integer $n := \mathrm{dist}(x, \partial\square)$. This ensures that for any function $v \in H^1(\square, \mu_\beta)$, the quantity $\Delta^n v$ is well-defined.

Remark 6.2. We only consider the charges q whose support is included in the cube $\square$, this ensures that for any function $v \in H^1(\square, \mu_\beta)$, the quantity $\nabla_q \cdot \mathbf{a}_q \nabla_q v$ is well-defined.

Remark 6.3. We subtract an additional term in the boundary layer $\left\{x \in \square : \mathrm{dist}(x, \partial\square) \leq \sqrt{R}/10\right\}$, where R denotes the sidelength of the cube $\square$. This term is a perturbative terms for two reasons: (i) we are only summing on a small boundary layer of size $\sqrt{R}/10$ of the cube $\square$, and (ii) the multiplicative factor $\beta^{-\frac{1}{4}}$ is much smaller than the leading order term of the energy $\mathbf{E}^*_\square$, which is of order 1. The reason justifying the presence of this term is that it is useful to deal with the infinite range of the operator $\mathcal{L}$; in particular, it is useful to prove the subadditivity of the energy functional ν^* in Proposition 6.17. The specific choice for the exponent $1/4$ for the power of β is arbitrary; we only need an exponent which is strictly between 0 and $1/2$.

By choosing the inverse temperature β sufficiently large, one can prove that the energy $\mathbf{E}_\square$ satisfies the following coercivity and boundedness properties: there exist constants $c(d) > 0$ and $C(d) < \infty$ such that, for each map $u \in H^1_0(\mathbb{Z}^d, \mu_\beta)$,

$$c \left[\!\left[u \right]\!\right]_{H^1(\square, \mu_\beta)} \leq \mathbf{E}_\square[u] \leq C \left[\!\left[u \right]\!\right]_{H^1(\square, \mu_\beta)}, \tag{6.1}$$

where we recall the notation $\left[\!\left[u \right]\!\right]_{H^1(\square, \mu_\beta)}$ introduced in Section 2.1.3. The same estimate holds for the energy functional $\mathbf{E}^*_\square$: for each $u \in H^1(\square, \mu_\beta)$,

$$c \left[\!\left[u \right]\!\right]_{H^1(\square, \mu_\beta)} \leq \mathbf{E}^*_\square[u] \leq C \left[\!\left[u \right]\!\right]_{H^1(\square, \mu_\beta)}. \tag{6.2}$$

We now proceed by giving the definitions of the subadditive quantities ν and ν^*.

Definition 6.4 (Subadditive quantities)**.** For each cube $\square$ of $\mathbb{Z}^d$, and each pair of vectors $p, p^* \in \mathbb{R}^{d \times \binom{d}{2}}$, we define the energies

$$\nu(\square, p) := \inf_{u \in l_p + H^1_0(\square, \mu_\beta)} \frac{1}{2|\square|} \mathbf{E}_\square[u], \tag{6.3}$$

and

$$\nu^*(\square, p^*) := \sup_{v \in H^1(\square, \mu_\beta)} -\frac{1}{2|\square|} \mathbf{E}^*_\square[v] + \frac{1}{|\square|} \sum_{x \in \square} p^* \cdot \langle \nabla v(x) \rangle_{\mu_\beta}. \tag{6.4}$$

Remark 6.5. We recall the definition of the affine function l_p stated in (A.4). We implicitly extend the functions of the space $l_p + H^1_0(\square, \mu_\beta)$ by the affine function l_p outside the cube $\square$.

It is clear from the estimate (6.1) that the energy quantities ν and ν^* are well-defined, quadratic in the variables p and p^* respectively, and that they satisfy the upper and lower bounds, for each cube $\square \subseteq \mathbb{Z}^d$ and each pair of vectors $p, p^* \in \mathbb{R}^{d \times \binom{d}{2}}$,

$$c|p|^2 \leq \nu(\square, p) \leq C|p|^2 \text{ and } c|p^*|^2 \leq \nu^*(\square, p^*) \leq C|p^*|^2. \tag{6.5}$$

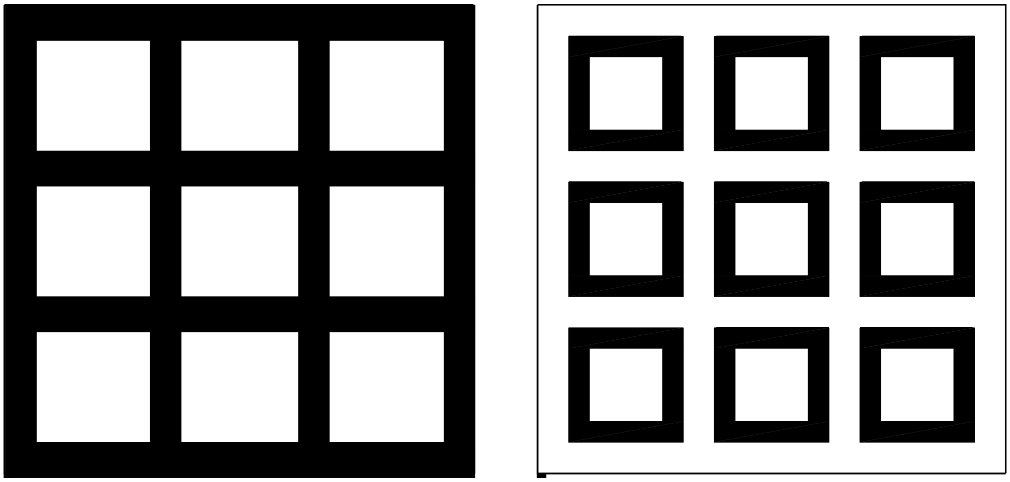

FIGURE 1. The picture on the left represents the cube $\square_{n+1}$, the white interior cubes are the cubes $(z+\square_n)_{z\in\mathcal{Z}_n}$ and the set in black is the boundary layer BL_n.

It follows from the standard argument of the calculus of variations that the minimizer in the variational definition (6.3) exists and is unique; we denote it by $u(\cdot,\square,p)$. The maximizer in the variational formulation (6.4) exists and is unique up to additive constant. This property is not a direct consequence of the standard arguments; it requires to use the properties of the Helffer-Sjöstrand equation and the regularity estimates established in Section 5. We omit the details of the argument and refer to the long version of this article ([36, Appendix B], first version of the arXiv submission). We denote by $v(\cdot,\square,p^*)$ the unique maximizer which satisfies $\sum_{x\in\square}\langle v(x,\cdot,\square,p^*)\rangle_{\mu_\beta}=0$. Additionally, we record that this maximizer satisfies the interior variance estimate

$$\sup_{x\in\frac{1}{3}\square}\operatorname{var}\left[v(x,\cdot,\square_n,p^*)\right]\leq C\left|p^*\right|^2. \tag{6.6}$$

The maps $p\mapsto u(\cdot,\cdot,\square,p)$ and $p^*\mapsto v(\cdot,\cdot,\square,p^*)$ are linear, and they satisfy the estimates

$$\left\|\nabla u(\cdot,\cdot,\square,p)\right\|_{\underline{L}^2(\square,\mu_\beta)}\leq C\left|p\right| \text{ and } \left\|\nabla v(\cdot,\cdot,\square,p^*)\right\|_{\underline{L}^2(\square,\mu_\beta)}\leq C\left|p^*\right|. \tag{6.7}$$

The goal of this section is to prove that, as the size of the cube $\square$ tends to infinity, the two quantities ν and ν^* converge and to obtain an algebraic rate of convergence. We obtain a result along a specific sequence of cubes defined below.

Definition 6.6 (Triadic cube and $\mathcal{Z}_n$)**.** We define the sequence l_n of non-negative real numbers according to the induction formula

$$l_0=1 \text{ and for each } n\in\mathbb{N},\ l_{n+1}=3l_n+\sqrt{l_n}.$$

For each $n\in\mathbb{N}$, we define the cube $\square_n:=\left(-\frac{l_n}{2},\frac{l_n}{2}\right)^d\cap\mathbb{Z}^d$. We denote by $\mathcal{Z}_{m,n}:=l_n 3^{m-n}\mathbb{Z}^d\cap\square_n$ and by $BL_{m,n}$ the mesoscopic boundary layer defined by the formula $BL_{m,n}:=\square_n\smallsetminus\bigcup_{z\in\mathcal{Z}_{m,n}}(z+\square_m)$. The cube $\square_n$ can be partitioned according to the formula

$$\square_n:=\bigcup_{z\in\mathcal{Z}_{m,n}}(z+\square_m)\cup BL_{m,n}.$$

We also introduce the notation $\mathcal{Z}_n:=\mathcal{Z}_{n,m}$, $BL_n:=BL_{n+1,n}$. We refer to Figure 1 for an illustration of these definitions. The set $BL_{m,n}$ is introduced to treat the infinite range of the operator $\mathcal{L}$.

In the following remarks, we record without proof some properties pertaining the Definition 6.6.

Remark 6.7. There exists a universal constant C such that, for each integer $n\in\mathbb{N}$, $3^n\leq l_n\leq C3^n$.

Remark 6.8. The cardinality of the set $\mathcal{Z}_{m,n}$ is equal to $3^{d(n-m)}$.

Remark 6.9. One has the volume estimate $|BL_{m,n}|\leq C3^{-\frac{m}{2}}|\square_n|$.

6.1.2. *Statement of the main result.* The main result obtained in this section is a quantitative rate of convergence for the two energy quantities ν and ν^*; it is stated below.

Proposition 6.10. *There exists an inverse temperature $\beta_0 := \beta_0(d) < \infty$ such that the following statement holds. There exist constants $c := c(d) > 0$, $C := C(d) < \infty$ and an exponent $\alpha := \alpha(d) > 0$ such that for each inverse temperature $\beta \geq \beta_0$, there exists a symmetric positive definite matrix $\bar{\mathbf{a}} \in \mathbb{R}^{d\binom{d}{2} \times d\binom{d}{2}}$ such that for each integer $n \in \mathbb{N}$, and each pair of vectors $p, p^* \in \mathbb{R}^{d \times \binom{d}{2}}$, one has the estimates*

$$\left| \nu\left(\square_n^-, p\right) - \frac{1}{2} p \cdot \bar{\mathbf{a}} p \right| \leq C 3^{-\alpha n} |p|^2 \quad \text{and} \quad \left| \nu^*\left(\square_n, p^*\right) - \frac{1}{2} p^* \cdot \bar{\mathbf{a}}^{-1} p^* \right| \leq C 3^{-\alpha n} |p^*|^2.$$

Remark 6.11. Using the symmetries of the model, we can prove the following properties. If we let L_{2,d^*} be the linear map introduced in Section 2.1.5, then there exists a coefficient $\overline{\lambda}_\beta := \overline{\lambda}_\beta(d, \beta)$ which tends to 0 as β tends to infinity such that

$$\begin{cases} \bar{\mathbf{a}} = \dfrac{1}{2} I_d \text{ in the space } \operatorname{Ker} L_{2,\mathrm{d}^*}, \\ \bar{\mathbf{a}} = \dfrac{(1 + \overline{\lambda}_\beta)}{2} I_d \text{ in the space } \left(\operatorname{Ker} L_{2,\mathrm{d}^*}\right)^\perp. \end{cases} \tag{6.8}$$

A direct consequence of (6.8) is the identity between the elliptic systems

$$-\nabla \cdot \bar{\mathbf{a}} \nabla = \frac{1}{2} \left(\mathrm{d}^* \mathrm{d} + \left(1 + \overline{\lambda}_\beta\right) \mathrm{d} \mathrm{d}^* \right).$$

These properties are a consequence of Proposition 6.10 and of Property (3) of Proposition 6.12.

6.1.3. *Outline of the argument.* The proof of Proposition 6.10 relies on ideas which were initially developed in [10], and follows the presentation given in [8]. The argument relies on the definition of the quantity

$$J\left(\square, p, p^*\right) := \nu\left(\square^-, p\right) + \nu^*\left(\square, p^*\right) - p \cdot p^*. \tag{6.9}$$

By the estimate (6.20) below, we know that the quadratic form J is almost positive, in the sense that it satisfies the inequality, for each cube $\square$ of size R, and each pair of vectors $p, p^* \in \mathbb{R}^{d \times \binom{d}{2}}$,

$$J\left(\square, p, p^*\right) \geq -C R^{-\frac{1}{2}} \left(|p|^2 + |p^*|^2 \right).$$

To prove Proposition 6.10, we argue that the map $J\left(\square, p, p^*\right)$ can be bounded from above in the following sense: for each vector $p \in \mathbb{R}^d$, there exists a vector $p^* \in \mathbb{R}^d$ such that

$$J\left(\square, p, p^*\right) \leq C 3^{-\alpha n} |p|^2. \tag{6.10}$$

Additionally, we prove that the vector p^* is close to $\bar{\mathbf{a}} p$. The quantitative rate of convergence stated in Proposition 6.10 is then a relatively straightforward consequence of the estimate (6.10). The proof of (6.10) is the core of the argument, it relies on a hierarchical decomposition of space and requires to introduce the subadditivity defect at scale l_n,

$$\begin{aligned} \tau_n &:= \sup_{p, p^* \in B_1} \left(\nu\left(\square_n^-, p\right) - \nu\left(\square_{n+1}^-, p\right) \right) + \left(\nu^*\left(\square_n, p^*\right) - \nu^*\left(\square_{n+1}, p^*\right) \right) \\ &= \sup_{p, p^* \in B_1} J\left(\square_n, p, p^*\right) - J\left(\square_{n+1}, p, p^*\right). \end{aligned} \tag{6.11}$$

We then prove a series of propositions and lemmas (Propositions 6.13 and 6.17, Lemmas 6.19, 6.20, 6.22 and 6.23), where various quantities are estimated in terms if the subadditivity defect τ_n. From these results we deduce an inequality of the form: for each integer $n \in \mathbb{N}$, and each vector $p \in \mathbb{R}^{d \times \binom{d}{2}}$, there exists a vector $p^* \in \mathbb{R}^{d \times \binom{d}{2}}$ such that

$$J\left(\square_{n+1}, p, p^*\right) \leq C \tau_n,$$

which can be rewritten

$$J\left(\square_{n+1}, p, p^*\right) \leq \frac{C}{C+1} J\left(\square_n, p, p^*\right). \tag{6.12}$$

The estimate (6.12) shows that, by passing from one scale to another, the energy quantity J has to contract by a multiplicative factor strictly less than 1. An iteration of the inequality (6.12) yields the algebraic rate of convergence stated in the inequality (6.10).

6.1.4. *Basic properties.* We first record some basic properties of the energy quantities ν and ν^*; they are analogous to [8, Lemma 2.2].

Proposition 6.12 (Basic properties of ν and ν^*)**.** *Fix a cube $\square \subseteq \mathbb{Z}^d$, and two vectors $p, p^* \in \mathbb{R}^{d\times\binom{d}{2}}$. The energy quantity $\nu(\square, p)$ (resp. $\nu^*(\square, p^*)$) and the minimizer $u(\cdot, \square, p)$ (resp. maximizer $v(\cdot, \square, p^*)$) satisfy the properties:*

(1) First variation. *The optimizing functions satisfy the following identities:*

$$\mathbf{B}_\square[u(\cdot,\square,p), w] = 0, \ \ \forall w \in H_0^1(\square, \mu_\beta),$$

and

$$\mathbf{B}^*_\square[v(\cdot,\square,p^*), w] = \frac{1}{|\square|}\sum_{x\in\square} p^* \cdot \langle \nabla w(x,\cdot)\rangle_{\mu_\beta}, \ \ \forall w \in H^1(\square, \mu_\beta).$$

(2) Second variation. *For each function $w \in l_p + H_0^1(\square, \mu_\beta)$,*

$$\frac{1}{2|\square|}\mathbf{E}_\square[w] - \nu(\square, p) = \frac{1}{2|\square|}\mathbf{E}_\square[u(\cdot,\square,p) - w]. \tag{6.13}$$

For each $w \in H^1(\square, \mu_\beta)$,

$$\nu^*(\square, p^*) + \frac{1}{2}\mathbf{E}_\square[w] - \frac{1}{|\square|}\sum_{x\in\square} p^* \cdot \langle \nabla w(x)\rangle_{\mu_\beta} = \frac{1}{2|\square|}\mathbf{E}^*_\square[v(\cdot,\square,p^*) - w]. \tag{6.14}$$

(3) Quadratic representation. *There exist two symmetric positive definite matrices $\mathbf{a}(\square), \mathbf{a}_*(\square) \in \mathbb{R}^{d\binom{d}{2}\times d\binom{d}{2}}$ such that*

$$\nu(\square, p) = \frac{1}{2}p\cdot\mathbf{a}(\square)p \quad \text{and} \quad \nu^*(\square, p^*) = \frac{1}{2}p^*\cdot\mathbf{a}_*(\square)^{-1}p^*. \tag{6.15}$$

Additionally, there exist two real coefficients $\lambda_{\beta,\square}$ and $\lambda^*_{\beta,\square}$, which tend to 0 as β tends to infinity, such that

$$\begin{cases} \mathbf{a}(\square) = I_d \text{ in the space } \mathrm{Ker} L_{2,\mathrm{d}^*}, \\ \mathbf{a}(\square) = (1+\lambda_{\beta,\square}) I_d \text{ in the space } (\mathrm{Ker} L_{2,\mathrm{d}^*})^\perp. \end{cases} \tag{6.16}$$

and

$$\begin{cases} \mathbf{a}_*(\square) = I_d \text{ in the space } \mathrm{Ker} L_{2,\mathrm{d}^*}, \\ \mathbf{a}_*(\square) = (1+\lambda^*_{\beta,\square}) I_d \text{ in the space } (\mathrm{Ker} L_{2,\mathrm{d}^*})^\perp. \end{cases} \tag{6.17}$$

We denote by $L^t_{2,\mathrm{d}^*} : \mathbb{R}^d \to \mathbb{R}^{d\times\binom{d}{2}}$ its adjoint of the map L_{2,d^*}. By differentiating the identities (6.15) with respect to the parameters p and p^*, we obtain the equalities

$$\frac{1}{|\square|}\sum_{x\in\square}\left(\frac{1}{2}\langle\nabla u(x,\cdot,\square,p)\rangle_{\mu_\beta} - \beta\sum_{\mathrm{supp}\, q\cap\square\neq\emptyset}\langle \mathbf{a}_q \nabla_q u(\cdot,\cdot,\square,p)\rangle_{\mu_\beta} L^t_{2,\mathrm{d}^*}(n_q(x))\right) = \mathbf{a}(\square)p \tag{6.18}$$

and

$$\frac{1}{|\square|}\sum_{x\in\square}\langle\nabla v(x,\cdot,\square,p^*)\rangle_{\mu_\beta} = \mathbf{a}_*(\square)^{-1}p^*. \tag{6.19}$$

(4) One-sided convex duality. *For each discrete cube $\square$ of sidelength R, we have the estimate*

$$J(\square, p, p^*) = \frac{1}{2|\square|}\mathbf{E}^*_\square[v(\cdot,\square,p^*) - u(\cdot,\square^-,p)] + O\left(C|p|^2 R^{-\frac{1}{2}}\right). \tag{6.20}$$

Proof. The proof of the properties (1) and (2) are straightforward and we refer to [8, Lemma 2.2]. For the identity (6.15), the arguments of [8] give the following results: for each cube $\square \subseteq \mathbb{Z}^d$, there exist two positive definite matrices $\mathbf{a}(\square), \mathbf{a}_*(\square) \in \mathbb{R}^{d\binom{d}{2}\times d\binom{d}{2}}$, such that, for each $p, p^* \in \mathbb{R}^{d\times\binom{d}{2}}$,

$$\nu(\square, p) = \frac{1}{2}p\cdot\mathbf{a}(\square)p \quad \text{and} \quad \nu^*(\square, p^*) = \frac{1}{2}p^*\cdot\mathbf{a}_*(\square)^{-1}p^*.$$

To prove the estimate (6.16), we use that any $p \in \mathrm{Ker}\, L_{2,\mathrm{d}^*}$, one has the identity $\mathrm{d}l_p = 0$. This implies that the minimizer in the energy $\nu(\square, p)$ is attained by the map l_p, from which one obtains that the linear map $\mathbf{a}$ is equal to the identity on the space $\mathrm{Ker}\, L_{2,\mathrm{d}^*}$. The proof of the result on the orthogonal complement of the space $\operatorname{Ker} L_{2,\mathrm{d}^*}$ is a consequence of the rotation and symmetry invariance of the dual Villain model. The proof of (6.17) is identical.

To prove the identity (6.18), we differentiate the equality (6.13) with respect to the variable p. We obtain, for each $p, p' \in \mathbb{R}^{d \times \binom{d}{2}}$,

$$
\begin{aligned}
\text{(6.21)} \qquad \mathbf{a}(\square) p \cdot p' &= \frac{1}{|\square|} \sum_{x \in \square} \frac{1}{2} \left\langle \nabla u(x, \cdot, \square, p) \cdot p' \right\rangle_{\mu_\beta} - \beta \sum_{\operatorname{supp} q \cap \square \neq \emptyset} \left\langle \mathbf{a}_q \nabla_q u(\cdot, \cdot, \square, p) \right\rangle_{\mu_\beta} (n_q, \mathrm{d}^* l_{p'}) \\
&= \frac{1}{|\square|} \sum_{x \in \square} \frac{1}{2} \left\langle \nabla u(x, \cdot, \square, p) \cdot p' \right\rangle_{\mu_\beta} - \beta \sum_{\operatorname{supp} q \cap \square \neq \emptyset} \left\langle \mathbf{a}_q \nabla_q u(\cdot, \cdot, \square, p) \right\rangle_{\mu_\beta} (n_q, L_{2,\mathrm{d}^*} (\nabla l_{p'})) \\
&= \frac{1}{|\square|} \sum_{x \in \square} \frac{1}{2} \left\langle \nabla u(x, \cdot, \square, p) \cdot p' \right\rangle_{\mu_\beta} - \beta \sum_{\operatorname{supp} q \cap \square \neq \emptyset} \left\langle \mathbf{a}_q \nabla_q u(\cdot, \cdot, \square, p) \right\rangle_{\mu_\beta} (n_q, L_{2,\mathrm{d}^*} (p')) \\
&= \frac{1}{|\square|} \sum_{x \in \square} \left(\frac{1}{2} \left\langle \nabla u(x, \cdot, \square, p) \cdot p' \right\rangle_{\mu_\beta} - \beta \sum_{\operatorname{supp} q \cap \square \neq \emptyset} \left\langle \mathbf{a}_q \nabla_q u(\cdot, \cdot, \square, p) \right\rangle_{\mu_\beta} L^t_{2,\mathrm{d}^*} (n_q(x)) \cdot p' \right).
\end{aligned}
$$

Using that the identity (6.21) is valid for every vector $p' \in \mathbb{R}^{d \times \binom{d}{2}}$, we obtain the identity (6.18).

There only remains to prove the one-sided convex duality property stated in (6.20). We apply the second variation formula (6.14), with the function $u = u(\cdot, \square^-, p)$, and use the identity

$$
\frac{1}{|\square|} \sum_{x \in \square} p^* \cdot \left\langle \nabla u(x, \cdot, \square^-, p) \right\rangle_{\mu_\beta} = p \cdot p^*,
$$

which is a consequence of the inclusion $\square^- \subseteq \square$ and the fact that the map u belongs to the space $l_p + H^1_0(\square, \mu_\beta)$. We obtain

$$
\nu^*(\square, p^*) + \frac{1}{2|\square|} \mathbf{E}^*_\square [u] - p^* \cdot p = \frac{1}{2|\square|} \mathbf{E}^*_\square \left[v(\cdot, \square, p^*) - u(\cdot, \square^-, p^*) \right].
$$

By definition of the function u, we have the equality $\nu(\square^-, p) = \frac{1}{2|\square^-|} \mathbf{E}_{\square^-}[u]$. To prove the inequality (6.20), it is thus sufficient to prove

$$
\text{(6.22)} \qquad \left| \frac{1}{|\square^-|} \mathbf{E}_{\square^-}[u] - \frac{1}{|\square|} \mathbf{E}^*_\square [u] \right| \leq C R^{-\frac{1}{2}}.
$$

The proof of the inequality (6.22) relies on the definitions of the two energies $\mathbf{E}_{\square^-}$ and $\mathbf{E}^*_\square$, and the fact that the function u is equal to the affine function l_p outside the cube $\square^-$. We omit the details here and refer to the long version of the article ([36, Chapter 6, Proposition 1.12]). □

6.2. **Subadditivity for the energy quantities.** In this section, we prove a spatial subadditivity property for the two energies ν and ν^*. The result is quantified, and we estimate the $\underline{H}^1(\square, \mu_\beta)$-norm of the difference of the minimizer u (resp. maximizer v) over two different scales in terms of the difference $\nu(\square_m, p) - \nu(\square_n, p)$ (resp. $\nu^*(\square_m, p^*) - \nu^*(\square_n, p^*)$).

6.2.1. *Subadditivity for the energy* ν. In this section, we prove that the energy quantity ν satisfies a subadditivity property with respect to the domain of integration and deduce from it the existence of the homogenized matrix $\bar{\mathbf{a}}$. The statement of Proposition 6.13 is quantified; we prove that the H^1-norm of the difference of the minimizer u over two different scales in terms of the subadditivity defect for the energy ν.

Proposition 6.13 (Subadditivity for ν)**.** *There exists an inverse temperature $\beta_0 := \beta_0(d) < \infty$ such that, for each $\beta \geq \beta_0$, the following statement is valid. There exists a constant $C := C(d) < \infty$ such that for each pair of integers $(m, n) \in \mathbb{N}$ satisfying $n > m$, and each vector $p \in \mathbb{R}^{d \times \binom{d}{2}}$,*

$$
\text{(6.23)} \qquad \frac{1}{|\mathcal{Z}_{m,n}|} \sum_{z \in \mathcal{Z}_{m,n}} \left[\!\left[u(\cdot, \square_n, p) - u(\cdot, z + \square_m, p) \right]\!\right]^2_{\underline{H}^1(\square_{n+1}, \mu_\beta)} \leq C \left(\nu(\square_m, p) - \nu(\square_n, p) + C 3^{-\frac{m}{2}} |p|^2 \right).
$$

Remark 6.14. Since it is useful in the rest of the proof, we note that the demonstration of Proposition 6.13 can be adapted to the case of trimmed cubes so as to obtain the estimate, for each pair of integers $m, n \in \mathbb{N}$ such that $m \leq n$,

$$
\frac{1}{|\mathcal{Z}_{m,n}|} \sum_{z \in \mathcal{Z}_{m,n}} \left[\!\left[u(\cdot, \square^-_n, p) - u(\cdot, z + \square^-_m, p) \right]\!\right]^2_{\underline{H}^1(\square_{n+1}, \mu_L)} \leq C \left(\nu(\square^-_m, p) - \nu(\square^-_n, p) + C 3^{-\frac{m}{2}} |p|^2 \right).
$$

Since the proof is essentially the same as the proof of Proposition 6.13; we omit the details.

Before proving Proposition 6.13, we record an immediate corollary of the subadditivity property for the energy ν.

Corollary 6.15. *There exists an inverse temperature $\beta_0 := \beta_0(d) < \infty$ such that, for each $\beta \geq \beta_0$, there exists a symmetric positive definite matrix $\bar{\mathbf{a}}$ such that, for each vector $p \in \mathbb{R}^{d\times\binom{d}{2}}$, one has*

$$\nu(\square_n, p) \underset{n\to\infty}{\longrightarrow} p \cdot \bar{\mathbf{a}} p.$$

By Property (3) of Proposition 6.12, this statement can be rewritten equivalently as

$$\mathbf{a}(\square_n) \underset{n\to\infty}{\longrightarrow} \bar{\mathbf{a}}.$$

Additionally, one deduces from (6.23) the lower bound estimate in the sense of symmetric positive definite matrices

$$\mathbf{a}(\square_n) \geq \bar{\mathbf{a}} - C3^{-\frac{n}{2}} I_{d\times\binom{d}{2}}, \tag{6.24}$$

where $I_{d\times\binom{d}{2}}$ denotes the identity matrix of the space $\mathbb{R}^{d\times\binom{d}{2}}$.

Remark 6.16. By Remark 6.14, the convergence also holds with the trimmed triadic cubes and we have, for each vector $p \in \mathbb{R}^{d\times\binom{d}{2}}$,

$$\nu(\square_n^-, p) \underset{n\to\infty}{\longrightarrow} p \cdot \bar{\mathbf{a}} p, \quad \mathbf{a}(\square_n^-) \underset{n\to\infty}{\longrightarrow} \bar{\mathbf{a}}, \quad \text{and} \quad \forall n \in \mathbb{N},\ \mathbf{a}(\square_n^-) \geq \bar{\mathbf{a}} - C3^{-\frac{n}{2}} I_{d\binom{d}{2}}.$$

Proof. Since the left side of (6.23) is non-negative, we have the inequality, for each pair of integers $m, n \in \mathbb{N}$ such that $n > m$,

$$\nu(\square_n, p) \leq \nu(\square_m, p) + C3^{-\frac{m}{2}} |p|^2. \tag{6.25}$$

Combining the inequality (6.25) with the fact that the sequence $(\nu(\square_n, p))_{n\in\mathbb{N}}$ is non-negative implies that it converges with the estimate (6.24). □

We now focus on the proof of Proposition 6.13.

Proof of Proposition 6.13. For the sake of simplicity, we only write the proof in the case when the difference between the integers m and n is equal to 1: we consider the specific case of the pair $(n, n+1)$. We assume without loss of generality that $|p| = 1$.

We let w be the function of $l_p + H_0^1(\square_{n+1}, \mu_\beta)$ defined by the following construction:

- For each point $z \in \mathcal{Z}_{n+1}$, we set $w := u(\cdot, z + \square_n, p)$;
- On the mesoscopic boundary layer BL_n, we set $w := l_p$.

Applying the second variation formula (6.13) and the coercivity of the energy functional $\mathbf{E}$ stated in (6.1) gives the inequality

$$[\![u(\cdot, \square_{n+1}, p) - w]\!]^2_{\underline{H}^1(\square_{n+1}, \mu_\beta)} \leq C\left(\frac{1}{2|\square_{n+1}|}\mathbf{E}_{\square_{n+1}}[w] - \nu(\square, p)\right). \tag{6.26}$$

Using that, for each point $z \in \mathcal{Z}_n$, the function w is equal to the minimizer $u(\cdot, z + \square_n, p)$ in the cube $(z + \square_n)$, we have the inequality

$$\sum_{z\in\mathcal{Z}_{n+1}} [\![u(\cdot, \square_{n+1}, p) - u(\cdot, z + \square_n, p)]\!]^2_{\underline{H}^1(\square_{n+1}, \mu_\beta)} \leq [\![u(\cdot, \square_{n+1}, p) - w]\!]^2_{\underline{H}^1(\square_{n+1}, \mu_\beta)}. \tag{6.27}$$

By the estimates (6.26) and (6.27), we see that, to prove the inequality (6.23), it is thus sufficient to prove

$$\frac{1}{2|\square|}\mathbf{E}_{\square_{n+1}}[w] \leq \nu(\square_n, p) + C3^{-\frac{n}{2}}. \tag{6.28}$$

We now prove the inequality (6.28). By definition of the energy $\mathbf{E}$, we have

$$\mathbf{E}_{\square_{n+1}}[w] := \underbrace{\beta \sum_{y \in \mathbb{Z}^d} \|\partial_y w\|^2_{L^2(\square_{n+1}, \mu_\beta)}}_{(6.29)-(i)} + \underbrace{\frac{1}{2}\|\nabla w\|^2_{L^2(\square_{n+1},\mu_\beta)}}_{(6.29)-(ii)} + \underbrace{\frac{1}{2}\sum_{k\geq 1}\frac{1}{\beta^{\frac{k}{2}}}\left\|\nabla^{k+1} w\right\|^2_{L^2(\mathbb{Z}^d,\mu_\beta)}}_{(6.29)-(iii)} - \underbrace{\beta \sum_{\mathrm{supp}\, q \cap \square_{n+1} \neq \emptyset} \langle \nabla_q w \cdot \mathbf{a}_q \nabla_q u\rangle_{\mu_\beta}}_{(6.29)-(iv)}. \tag{6.29}$$

We estimate the four terms on the right side separately. The term (6.29)-(i) involving the derivative with respect to the field ϕ can be estimated by the following argument. Since the map w is equal to the deterministic affine function l_p in the boundary layer BL_n, we have the identity $\partial_y w(x,\cdot) = 0$ for any point $x \in BL_n$ and any point $y \in \mathbb{Z}^d$. This implies the equality

$$\sum_{y \in \mathbb{Z}^d} \|\partial_y w\|^2_{L^2(\square_{n+1}\mu_\beta)} = \sum_{z \in \mathcal{Z}_n}\sum_{y\in\mathbb{Z}^d} \|\partial_y u(\cdot, z+\square_n)\|^2_{L^2(z+\square_n \mu_\beta)}. \tag{6.30}$$

This completes the estimate of the term (6.29)-(i). For the term (6.29)-(ii), we use the same argument and note that $\nabla w(x,\cdot) = p$ for any point $x \in BL_n$. We obtain

$$\begin{aligned}\frac{1}{|\square_{n+1}|}\|\nabla w\|^2_{L^2(\square_{n+1}\mu_\beta)} &= \frac{1}{|\square_{n+1}|}\sum_{z\in\mathcal{Z}_n}\|\nabla u(\cdot, z+\square_n, p)\|^2_{L^2(z+\square_n,\mu_\beta)} + \frac{|BL_n|}{|\square_{n+1}|} \\ &\leq \frac{1}{|\square_{n+1}|}\sum_{z\in\mathcal{Z}_n}\|\nabla u(\cdot, z+\square_n, p)\|^2_{L^2(z+\square_n,\mu_\beta)} + C3^{-\frac{n}{2}}.\end{aligned} \tag{6.31}$$

The term (6.29)-(iii) can be estimated with a similar strategy, but some additional technicalities need to be treated along the way to deal with the iterations of the Laplacian and the sum over the charges. We omit the details here and only give the results

$$(6.29)-(iii) \leq \sum_{z\in\mathcal{Z}_n}\sum_{k\geq 1}\frac{1}{\beta^{\frac{k}{2}}}\left\|\nabla^{k+1}u(\cdot, z+\square_n, p)\right\|^2_{L^2(\mathbb{Z}^d,\mu_\beta)} + Ce^{-c(\ln\beta)3^{\frac{n}{2}}}, \tag{6.32}$$

and

$$(6.29)-(iv) \leq \sum_{z\in\mathcal{Z}_n}\sum_{\mathrm{supp}\, q\cap(z+\square_n)} \langle \nabla_q u(\cdot, z+\square_n, p)\cdot \mathbf{a}_q \nabla_q u(\cdot, z+\square_n, p)\rangle_{\mu_\beta} + C3^{-\frac{n}{2}}|\square_{n+1}|. \tag{6.33}$$

We finally combine the equality (6.29), the estimates (6.30), (6.31), (6.32), (6.33) to obtain the inequality (6.28). The proof of Proposition 6.13 is complete. □

6.2.2. *Subadditivity for the energy ν^*.* In this section, we prove a similar statement for the energy ν^*.

Proposition 6.17 (Subadditivity for ν^*). *There exists a constant $C := C(d) < \infty$ such that for each pair of integers $(n,m) \in \mathbb{N}$ such that $n > m$ and each vector $p^* \in \mathbb{R}^{d\times\binom{d}{2}}$,*

$$\frac{1}{|\mathcal{Z}_{m,n}|}\sum_{z\in\mathcal{Z}_{m,n}} [\![v(\cdot,\cdot,\square_n,p^*) - v(\cdot,\cdot,z+\square_m,p^*)]\!]^2_{H^1(z+\square_m,\mu_\beta)} \leq C\left(\nu^*(\square_m,p) - \nu^*(\square_n,p) + 3^{-\frac{n}{2}}|p^*|^2\right). \tag{6.34}$$

As it was the case for the energy quantity ν, we deduce from Proposition 6.17 that the sequence $(\nu^*(\square_n,p^*))_{n\in\mathbb{N}}$ converges as n tends to infinity.

Corollary 6.18. *There exists an inverse temperature $\beta_0 := \beta_0(d) < \infty$ such that for each $\beta \geq \beta_0$ the following statement is valid. There exists a symmetric definite positive $\bar{\mathbf{a}}_*$ such that for each vector $p^* \in \mathbb{R}^{d\times\binom{d}{2}}$,*

$$\nu^*(\square_n,p^*) \underset{n\to\infty}{\longrightarrow} \bar{\mathbf{a}}_*^{-1}|p^*|^2.$$

By the Property (3) of Proposition 6.12, this statement can be rewritten equivalently

$$\mathbf{a}_*(\square_n)^{-1} \underset{n\to\infty}{\longrightarrow} \bar{\mathbf{a}}_*^{-1}.$$

We also have the lower bound, for each integer $n \in \mathbb{N}$,

$$\mathbf{a}_*(\square_n)^{-1} \geq \bar{\mathbf{a}}_*^{-1} - C3^{-\frac{n}{2}} I_{d\binom{d}{2}}.$$

Proof of Proposition 6.17. For the sake of simplicity, we only write the proof in the specific case of the pair $(m, n) = (n+1, n)$. We assume without loss of generality that $|p^*| = 1$.

We consider the function $v := v(\cdot, \square_{n+1}, p^*)$ and, for $z \in \mathcal{Z}_n$, we restrict it to the cubes $(z + \square_n)$. We apply the second variation formula (6.14) and the coercivity of the energy functional $\mathbf{E}^*_{z+\square_n}$. We obtain, for each point $z \in \mathcal{Z}_n$,

$$[\![v(\cdot, \square_{n+1}, p^*) - v(\cdot, z + \square_n, p^*)]\!]^2_{\underline{H}^1(z+\square_n, \mu_\beta)} \leq C \left(\nu^*(z + \square_n, p^*) + \frac{1}{2|\square_n|} \mathbf{E}^*_{z+\square_n}[v] + \frac{1}{|\square_n|} \sum_{x \in z + \square} p^* \cdot \langle \nabla v(x) \rangle_{\mu_\beta} \right).$$

Summing over the points $z \in \mathcal{Z}_n$, and dividing by the cardinality of $\mathcal{Z}_n$ shows

$$\frac{1}{|\mathcal{Z}_n|} \sum_{z \in \mathcal{Z}_n} [\![v(\cdot, \square_{n+1}, p^*) - v(\cdot, z + \square_n, p^*)]\!]^2_{\underline{H}^1(z+\square_n, \mu_\beta)}$$
$$\leq C \left(\nu^*(\square_n, p^*) + \sum_{z \in \mathcal{Z}_n} \frac{1}{2|\mathcal{Z}_n| \cdot |\square_n|} \mathbf{E}^*_{z+\square_n}[v] + \frac{1}{|\mathcal{Z}_n| \cdot |\square_n|} \sum_{x \in \square_n} p^* \cdot \langle \nabla v(x) \rangle_{\mu_\beta} \right).$$

The factor $|\mathcal{Z}_n| = 3^d$ on the left side depends only on the dimension d, and can thus be incorporated in the constant C in the right side. We deduce that, to prove the inequality (6.34), it is sufficient to prove

$$\sum_{z \in \mathcal{Z}_n} \frac{1}{|\mathcal{Z}_n| \cdot |\square_n|} \left(\frac{1}{2} \mathbf{E}^*_{z+\square_n}[v] + \sum_{x \in z + \square_n} p^* \cdot \langle \nabla v(x) \rangle_{\mu_\beta} \right)$$
$$\leq \underbrace{\frac{1}{2|\square_{n+1}|} \mathbf{E}^*_{\square_{n+1}}[v]}_{(6.35)-(i)} + \underbrace{\frac{1}{|\square_{n+1}|} \sum_{x \in \square_{n+1}} p^* \cdot \langle \nabla v(x) \rangle_{\mu_\beta}}_{(6.35)-(ii)} + C 3^{-\frac{n}{2}}. \tag{6.35}$$

We first estimate the term (6.35)-(ii). We use the estimate (6.7) on the L^2-norm of the gradient of the function v, the Cauchy-Schwarz inequality, and the volume estimate

$$|\square_{n+1}| - |\mathcal{Z}_n| \cdot |\square_n| = \left| \square_{n+1} \setminus \bigcup_{z \in \mathcal{Z}_n} (z + \square_n) \right| = |BL_n| \leq C 3^{-\frac{n}{2}} |\square_{n+1}|.$$

We obtain

(6.36)

$$\sum_{z \in \mathcal{Z}_n} \frac{1}{|\mathcal{Z}_n| \cdot |\square_n|} \sum_{x \in z + \square_n} p^* \cdot \langle \nabla v(x, \cdot) \rangle_{\mu_\beta} \leq \frac{1}{|\square_{n+1}|} \sum_{z \in \mathcal{Z}_n} \sum_{x \in z + \square_n} p^* \cdot \langle \nabla v(x) \rangle_{\mu_\beta} + \left(\frac{|BL_n|}{|\square_{n+1}|} \right)^{\frac{1}{2}} \| \nabla v \|_{\underline{L}^2(\square_{n+1}, \mu_\beta)}$$
$$\leq \frac{1}{|\square_{n+1}|} \sum_{x \in \square_{n+1}} p^* \cdot \langle \nabla v(x, \cdot) \rangle_{\mu_\beta} + \frac{1}{|\square_{n+1}|} \sum_{x \in BL_n} p^* \cdot \langle \nabla v(x, \cdot) \rangle_{\mu_\beta} + C 3^{-\frac{n}{4}}$$
$$\leq \frac{1}{|\square_{n+1}|} \sum_{x \in \square_{n+1}} p^* \cdot \langle \nabla v(x, \cdot) \rangle_{\mu_\beta} + \left(\frac{|BL_n|}{|\square_{n+1}|} \right)^{\frac{1}{2}} \| \nabla v \|_{\underline{L}^2(\square_{n+1}, \mu_\beta)} + C 3^{-\frac{n}{4}}$$
$$\leq \frac{1}{|\square_{n+1}|} \sum_{x \in \square_{n+1}} p^* \cdot \langle \nabla v(x, \cdot) \rangle_{\mu_\beta} + C 3^{-\frac{n}{4}}.$$

To estimate the term (6.35)-(i), we compare the two energies $\sum_{z \in \mathcal{Z}_n} \mathbf{E}^*_{z+\square_n}[v]$ and $\mathbf{E}^*_{\square_{n+1}}[v]$, and estimate the terms which differ in the two quantities. These terms are either boundary layer terms or terms coming from the sum over the charges and the iterations of the Laplacian. In both cases, we can prove that they are small; we omit the technical details and only write the result

$$\sum_{z \in \mathcal{Z}_n} \frac{1}{2|\mathcal{Z}_n| \cdot |\square_n|} \mathbf{E}^*_{z+\square_n}[v] \leq \frac{1}{2|\square_{n+1}|} \mathbf{E}^*_{\square_{n+1}}[v] + C e^{-c 3^{\frac{n}{2}}}. \tag{6.37}$$

Combining the estimates (6.37) and (6.36) shows the inequality (6.35) and completes the proof of Proposition 6.17. □

6.3. **Quantitative convergence of the subadditive quantities.** In this section, we prove an algebraic rate of convergence for the quantity J defined in (6.9). We recall the definition of the subadditivity defect τ_n given in (6.11), and we introduce the following notation: for each integer $n \in \mathbb{N}$,

$$\bar{\mathbf{a}}_n := \mathbf{a}_* (\square_n) , \tag{6.38}$$

and call the matrix $\bar{\mathbf{a}}_n$ the *approximate homogenized matrix*. We first prove a series of lemmas, estimating various quantities in terms of the subadditivity defect τ_n.

Before starting the proofs, let us make the following remark. By Corollaries 6.15 and 6.18, the subadditivity defect τ_n converges to 0 as n tends to infinity. In particular all the quantities which are bounded from above by the subadditivity defect τ_n tend to 0 as n tends to infinity.

6.3.1. *Control over the approximate homogenized coefficient.* The first lemma we prove establishes that the difference between the matrices $\bar{\mathbf{a}}_n$ over two different scales can be estimated in terms of the subadditivity defect τ_n.

Lemma 6.19. *There exists a constant $C := C(d) < \infty$ such that, for any pair of integers $(m,n) \in \mathbb{N}^2$ with $m \leq n$,*

$$\left| \bar{\mathbf{a}}_n^{-1} - \bar{\mathbf{a}}_m^{-1} \right|^2 \leq \sum_{k=m}^{n} \tau_k + C 3^{-\frac{m}{2}} .$$

Proof. Before starting the proof, we collect a few ingredients and notation used in the argument:

- By the formula (6.19), we have the identity $\sum_{x \in \square_n} \langle \nabla v(x, \cdot, \square_n, p^*) \rangle_{\mu_\beta} = \mathbf{a}_n^{-1} p^*$;
- By definition of the subadditivity defect τ_k, we have the identity, for each $p \in \mathbb{R}^{d\binom{d}{2}}$,

$$\nu^* (\square_m, p) - \nu^* (\square_n, p) \leq |p|^2 \sum_{k=m}^{n} \tau_k .$$

We fix a vector $p^* \in \mathbb{R}^{d\binom{d}{2}}$ such that $|p^*| = 1$, and use the formula (6.19) to write

$$\begin{aligned} \bar{\mathbf{a}}_n^{-1} p^* &= \frac{1}{|\square_n|} \sum_{x \in \square_n} \langle \nabla v(x, \cdot, \square_n, p^*) \rangle_{\mu_\beta} \\ &= \underbrace{\frac{1}{|\square_n|} \sum_{z \in \mathcal{Z}_{m,n}} \sum_{x \in z + \square_m} \langle \nabla v(x, \cdot, \square_n, p^*) \rangle_{\mu_\beta}}_{(6.39)-(i)} + \underbrace{\frac{1}{|\square_n|} \sum_{x \in BL_{m,n}} \langle \nabla v(x, \cdot, \square_n, p^*) \rangle_{\mu_\beta}}_{(6.39)-(ii)} . \end{aligned} \tag{6.39}$$

The term (6.39)-(ii) is the simplest one, we estimate it by the Cauchy-Schwarz inequality, the estimate on the L^2-norm of the gradient of v stated in (6.7), and the volume estimate $|BL_{m,n}| \leq C 3^{-\frac{m}{2}} |\square_n|$. We obtain

$$\left| \frac{1}{|\square_n|} \sum_{x \in BL_{m,n}} \langle \nabla v(x, \cdot, \square_n, p^*) \rangle_{\mu_\beta} \right| \leq C 3^{-\frac{m}{2}} . \tag{6.40}$$

To estimate the term (6.39)-(i), we use the estimate (6.7), the identity $BL_{m,n} = \square_n \setminus \bigcup_{z \in \mathcal{Z}_{m,n}} (z + \square_n)$, and the volume estimate $|BL_{m,n}| \leq C 3^{-\frac{m}{2}} |\square_n|$. We obtain

$$\frac{1}{|\square_n|} \sum_{z \in \mathcal{Z}_n} \sum_{x \in z + \square_m} \langle \nabla v(x, \cdot, \square_n, p^*) \rangle_{\mu_\beta} = \frac{1}{|\mathcal{Z}_{m,n}|} \frac{1}{|z + \square_m|} \sum_{z \in \mathcal{Z}_n} \sum_{x \in z + \square_m} \langle \nabla v(x, \cdot, \square_n, p^*) \rangle_{\mu_\beta} + O\left(C 3^{-\frac{m}{2}} \right) .$$

Applying the subadditivity estimate stated in Proposition 6.13, we find that

$$\begin{aligned}(6.41)\qquad &\frac{1}{|\mathcal{Z}_{m,n}|}\frac{1}{|z+\square_m|}\sum_{z\in\mathcal{Z}_n}\sum_{x\in z+\square_m}\left|\left\langle\nabla v(x,\cdot,\square_n,p^*)-\nabla v(x,\cdot,z+\square_m,p^*)\right\rangle_{\mu_\beta}\right| \\ &\leq \frac{1}{|\mathcal{Z}_{m,n}|}\sum_{z\in\mathcal{Z}_{m,n}}\left\|\nabla v\left(\cdot,\square_n,p^*\right)-\nabla v\left(\cdot,z+\square_m,p^*\right)\right\|_{\underline{L}^2(z+\square_m,\mu_\beta)} \\ &\leq \left(\frac{1}{|\mathcal{Z}_{m,n}|}\sum_{z\in\mathcal{Z}_{m,n}}\left\|\nabla v\left(\cdot,\square_n,p^*\right)-\nabla v\left(\cdot,z+\square_m,p^*\right)\right\|^2_{\underline{L}^2(z+\square_m,\mu_\beta)}\right)^{\frac12} \\ &\leq \left(\frac{1}{|\mathcal{Z}_{m,n}|}\sum_{z\in\mathcal{Z}_{m,n}}\left[\!\left[\nabla v\left(\cdot,\square_n,p^*\right)-\nabla v\left(\cdot,z+\square_m,p^*\right)\right]\!\right]^2_{\underline{H}^1(z+\square_m,\mu_\beta)}\right)^{\frac12} \\ &\leq C\left(\sum_{k=m}^n\tau_k\right)^{\frac12}+C3^{-\frac{m}{2}}.\end{aligned}$$

We then use the inequality (6.41), the translation invariance of the measure μ_β, and the identity $\sum_{x\in\square_n}\left\langle\nabla v(x,\cdot,\square_m,p^*)\right\rangle_{\mu_\beta}=\bar{\mathbf{a}}_m^{-1}p^*$. We obtain

$$\begin{aligned}(6.42)\qquad &\frac{1}{|\mathcal{Z}_{m,n}|}\frac{1}{|z+\square_m|}\sum_{z\in\mathcal{Z}_n}\sum_{x\in z+\square_m}\left\langle\nabla v(x,\cdot,\square_n,p^*)\right\rangle_{\mu_\beta} \\ &=\frac{1}{|\mathcal{Z}_{m,n}|}\sum_{z\in\mathcal{Z}_{m,n}}\frac{1}{|z+\square_m|}\sum_{x\in z+\square_m}\left\langle\nabla v(x,\cdot,z+\square_m,p^*)\right\rangle_{\mu_\beta}+O\left(C\left(\sum_{k=m}^n\tau_k\right)^{\frac12}+C3^{-\frac{m}{2}}\right) \\ &=\frac{1}{|\square_m|}\sum_{x\in\square_m}\left\langle\nabla v(x,\cdot,\square_m,p^*)\right\rangle_{\mu_\beta}+O\left(C\left(\sum_{k=m}^n\tau_k\right)^{\frac12}+C3^{-\frac{m}{2}}\right) \\ &=\bar{\mathbf{a}}_m^{-1}p^*+O\left(C\left(\sum_{k=m}^n\tau_k\right)^{\frac12}+C3^{-\frac{m}{2}}\right).\end{aligned}$$

We then combine the identity (6.39) with the estimates (6.40) and (6.42) to complete the proof of Lemma 6.19. □

6.3.2. *Control over the variance of the spatial average of the maximizer v.* The next step in the argument is to control the variance of the spatial average of the maximiser v. We prove that its variance contracts and obtain an algebraic rate of convergence. The proof relies on an explicit computation and makes use of the second-order Helffer-Sjöstrand equation introduced in Section 3.4 to estimate the correlation between the random variables $\phi\mapsto v(x,\phi,\square_{n+1},p)$ and $\phi\mapsto v(x',\phi,\square_{n+1},p)$ for a pair of points $x,x'\in\square_{n+1}$ distant from one another.

Lemma 6.20 (Variance estimate)**.** *There exists a constant $C:=C(d)<\infty$ such that, for each $n\in\mathbb{N}$, and each $p^*\in\mathbb{R}^{d\times\binom{d}{2}}$,*

$$(6.43)\qquad \mathrm{var}_{\mu_\beta}\left[\frac{1}{|\square_n|}\sum_{x\in\square_n}\nabla v(x,\cdot,\square_{n+1},p^*)\right]\leq C3^{-\left(d-\frac52\right)n}|p^*|^2.$$

For later purposes, we also record that the variance of the flux contracts

$$(6.44)\qquad \mathrm{var}\left[\frac{1}{|\square_n|}\sum_{x\in\square_n}\left(\frac12\nabla v\left(x,\cdot,\square_{n+1},p^*\right)+\beta\sum_{q\in\mathcal{Q}}\mathbf{a}_q\nabla_q v\left(\cdot,\cdot,\square_{n+1},p^*\right)n_q(x)\right)\right]\leq C3^{-\left(d-\frac52\right)n}|p^*|^2.$$

Remark 6.21. The value of the coefficient $d-\frac52$ is arbitrary; we can prove the result for any fixed number strictly smaller than $d-2$ by choosing β large enough.

Proof. We fix an inverse temperature β large enough so that all the regularity results of Section 5 hold with the regularity exponent $\varepsilon=\frac14$. We decompose the argument into two steps.

Step 1. To ease the notation, we denote by $v := v(\cdot, \cdot, \square_{n+1}, p^*)$. We assume without loss of generality that $|p^*| = 1$. We first decompose the variance

$$\operatorname{var}_{\mu_\beta}\left[\frac{1}{|\square_n|}\sum_{x\in\square_n}\nabla v(x,\cdot)\right] = \frac{1}{|\square_n|^2}\sum_{x,x'\in\square_n}\operatorname{cov}_{\mu_\beta}\left[\nabla v(x,\cdot), \nabla v(x',\cdot)\right]. \tag{6.45}$$

We then prove the estimate, for each pair of points $x, x' \in \square_n$,

$$\left|\operatorname{cov}_{\mu_\beta}\left[\nabla v(x,\cdot), \nabla v(x',\cdot)\right]\right| \leq \frac{C3^{\frac{n}{2}}}{|x-x'|^{d-2}}. \tag{6.46}$$

The estimate (6.43) can then be deduced from (6.46) and (6.45); indeed we have

$$\begin{aligned}\operatorname{var}_{\mu_\beta}\left[\frac{1}{|\square_n|}\sum_{x\in\square_n}\nabla v(x,\cdot)\right] &\leq \frac{1}{|\square_n|^2}\sum_{x,x'\in\square_n}\operatorname{cov}_{\mu_\beta}\left[\nabla v(x,\cdot), \nabla v(x',\cdot)\right] \\ &\leq \frac{C3^{\frac{n}{2}}}{|\square_n|^2}\sum_{x,x'\in\square_n}\frac{1}{|x-x'|^{d-2}} \\ &\leq C3^{-\left(d-\frac{5}{2}\right)n}.\end{aligned}$$

We now fix two points $x, x' \in \square_n$, and focus on the proof of (6.46). By applying the Helffer-Sjöstrand formula, we write

$$\operatorname{cov}_{\mu_\beta}\left[\nabla v(x,\cdot), \nabla v(x',\cdot)\right] = \sum_{y\in\mathbb{Z}^d}\left\langle \partial_y \nabla v(x,\cdot)\,\mathcal{H}_{x'}(y,\cdot)\right\rangle_{\mu_\beta}, \tag{6.47}$$

where $\mathcal{H}_{z'}$ is the solution of the Helffer-Sjöstrand equation, for each pair $(y,\phi) \in \mathbb{Z}^d \times \Omega$,

$$\mathcal{L}\mathcal{H}_{x'}(y,\phi) = \partial_y \nabla v(x',\phi).$$

We then decompose the function $\mathcal{H}_{x'}$ according to the collection of Green's matrices $\left(\mathcal{G}_{\partial_y \nabla v(x',\cdot)}\right)_{y\in\mathbb{Z}^d}$. We obtain

$$\mathcal{H}_{x'}(y,\phi) = \sum_{y'\in\mathbb{Z}^d}\mathcal{G}_{\partial_{y'}\nabla v(x',\cdot)}(y,\phi;y').$$

Using Proposition 3.17, we can estimate the $L^2(\mu_\beta)$-norm of the function $\mathcal{H}_{x'}$, for each point $y \in \mathbb{Z}^d$,

$$\left\|\mathcal{H}_{x'}(y,\cdot)\right\|_{L^2(\mu_\beta)} \leq C\sum_{y'\in\mathbb{Z}^d}\frac{\left\|\partial_{y'}\nabla v(x',\cdot)\right\|_{L^2(\mu_\beta)}}{|y-y'|^{d-2}}. \tag{6.48}$$

We then claim that we have the estimates, for each pair of points $y, y' \in \mathbb{Z}^d$,

$$\left\|\partial_y \nabla v(x,\cdot)\right\|_{L^2(\mu_\beta)} \leq \frac{C3^{\frac{n}{4}}}{|y-x|^{d+\frac{3}{4}}} \text{ and } \left\|\partial_{y'}\nabla v(x',\cdot)\right\|_{L^2(\mu_\beta)} \leq \frac{C3^{\frac{n}{4}}}{|y'-x'|^{d+\frac{3}{4}}}. \tag{6.49}$$

The estimate (6.49) is proved in Step 2 below. Combining the inequalities (6.48), (6.49), and the formula (6.47), we obtain

$$\operatorname{cov}_{\mu_\beta}\left[\nabla v(x,\cdot), \nabla v(x',\cdot)\right] \leq C3^{\frac{n}{2}}\sum_{y,y'\in\mathbb{Z}^d}\frac{1}{|y'-x'|^{d+\frac{3}{4}}}\times\frac{1}{|y-x|^{d+\frac{3}{4}}}\times\frac{1}{|y-y'|^{d-2}}. \tag{6.50}$$

The sum in the right side of the inequality (6.50) can be explicitly computed and we obtain the inequality (6.46).

Step 2. Proof of (6.49). The argument relies on the second-order Helffer-Sjöstrand equation introduced in Section 5.4 and on the reflection principle to solve the Neumann problem (6.53) below. Given a cube $Q \subseteq \mathbb{Z}^d$ of sidelength R, we recall the notation $\frac{1}{2}Q$ to denote the cube which has the same center as Q and sidelength $R/2$. We consider the specific cube $\square := (0, l_{n+1})^d$ and the function $v(\cdot,\cdot,\square,p^*)$. Since the cube $\square_{n+1}$ can be obtained from the cube $\square$ by a translation, and since the measure μ_β is translation invariant, we see that to prove the estimate (6.49), it is sufficient to prove the inequality, for each point $y \in \frac{1}{3}\square$, and each point $z \in \mathbb{Z}^d$,

$$\left\|\partial_z \nabla v(y,\cdot,\square,p^*)\right\|_{L^2(\mu_\beta)} \leq \frac{C3^{\frac{n}{4}}}{|y-z|^{d+\frac{3}{4}}}. \tag{6.51}$$

The reason justifying this specific choice for the cube $\square$ will become clear later in the proof. Using the definition of the map $v := v(\cdot,\cdot,\square,p^*)$ as a minimizer in the variational formulation of $\nu^*(\square,p^*)$ stated in (6.4), we see that it is a solution of the Neumann problem

$$\begin{cases} -\Delta_\phi v + \mathcal{L}_\square v = 0 & \text{in } \square\times\Omega, \\ \mathbf{n}\cdot\nabla v = \mathbf{n}\cdot p^* & \text{on } \partial\square\times\Omega, \end{cases} \tag{6.52}$$

where the operator $\mathcal{L}_\square$ is the uniformly elliptic operator defined by the formula

$$\mathcal{L}_\square := -\frac{1}{2\beta}\Delta + \frac{1}{2\beta}\sum_{k\geq 1}\frac{(-1)^{k+1}}{\beta^{\frac{k}{2}}}\nabla^{k+1}\cdot\left(\mathbb{1}_{\square^k}\nabla^{k+1}\right) + \frac{1}{\beta^{\frac{5}{4}}}\nabla\cdot\left(\mathbb{1}_{\square\setminus\square^-}\nabla\right) + \sum_{\operatorname{supp} q\subseteq\square}\nabla_q\cdot\mathbf{a}_q\nabla_q,$$

where we recall the notation $\square^k := \{x\in\square : \operatorname{dist}(x,\partial\square)\geq k\}$. The specific, technical formula of the operator $\mathcal{L}_\square$ is not relevant in the proof; the important point of the argument is that the operator $\mathcal{L}_\square$ is well-defined for functions which are only defined in the interior of the triadic cube $\square$, and that, as it is the case for elliptic operator $\mathcal{L}_{\mathrm{spat}}$, it is uniformly elliptic and is a perturbation of the Laplacian $-\frac{1}{2\beta}\Delta$. As a consequence, all the results stated in Section 5 for the Helffer-Sjöstrand operator $\mathcal{L}$ are also valid for the operator $-\Delta_\phi+\mathcal{L}_\square$. In particular, all the arguments stated in Section 5.4 about the second-order Helffer-Sjöstrand equation apply in this setting. By applying the partial derivative ∂ to the system (6.52), we obtain that, if we denote by $w(y,z,\phi) = \partial_z v(y,\phi)$, then the function w solves the system

$$\begin{cases} -\Delta_\phi w + \mathcal{L}_{\square,y} w + \mathcal{L}_{\mathrm{spat},z} w = \sum_{\operatorname{supp} q\subseteq\square} z(\beta,q)(2\pi)\sin(2\pi(\phi,q))(v,q)\, q_y\otimes q_z & \text{in } \square\times\mathbb{Z}^d\times\Omega, \\ \mathbf{n}\cdot\nabla_y w = 0 & \text{on } \partial\square\times\mathbb{Z}^d\times\Omega, \end{cases} \tag{6.53}$$

where the subscripts y (resp. z) in the notation $\mathcal{L}_{\square,y}$ (resp. $\mathcal{L}_{\mathrm{spat},z}$) means that the spatial operator $\mathcal{L}_{\square_{n+1}}$ (resp. $\mathcal{L}_{\mathrm{spat},z}$) only acts on the spatial variable y (resp. z). We introduce the notation $\mathbf{f}$ to denote the function

$$\mathbf{f} := \begin{cases} \square\times\mathbb{Z}^d\times\Omega \to \mathbb{R}^{d\times d}, \\ (y,z,\phi)\mapsto \sum_{\operatorname{supp} q\subseteq\square} z(\beta,q)(2\pi)\sin(2\pi(\phi,q))(v,q)\, n_q(y)\otimes n_q(z). \end{cases}$$

Using this notation, the system (6.53) becomes

$$\begin{cases} -\Delta_\phi w + \mathcal{L}_{\square,y} w + \mathcal{L}_{\mathrm{spat},z} w = \mathrm{d}_y\mathrm{d}_z\mathbf{f} & \text{in } \square\times\mathbb{Z}^d\times\Omega, \\ \mathbf{n}\cdot\nabla_y w = 0 & \text{on } \partial\square\times\mathbb{Z}^d\times\Omega. \end{cases} \tag{6.54}$$

To solve the system (6.54), we use the reflection principle. To this end, we need to introduce a few definitions, notation and remarks. We fix a point $z\in\mathbb{Z}^d$ and extend the elliptic operator $\mathcal{L}_\square$, the functions v and $\mathbf{f}(\cdot,z)$, initially defined on the cube $\square$, to the entire space according to the following procedure. We let $\widetilde{\square}$ be the discrete cube $(-l_{n+1},l_{n+1})^d$. For each point $x=(x_1,\ldots,x_d)\in\widetilde{\square}$, we extend $\mathbf{f}$ by setting, for any $(i,j)\in\{1,\ldots,d\}^2$,

$$\mathbf{f}_{ij}(x,z,\phi) = (-1)^{\operatorname{sgn}(x_i)}\mathbf{f}_{ij}(|x_1|,\ldots,|x_d|,z,\phi). \tag{6.55}$$

We also use the reflection to extend the operator $\mathcal{L}_\square$ to the cube $\widetilde{\square}$, and denote this extension by $\mathcal{L}_{\widetilde{\square}}$. We then extend the operator $\mathcal{L}_{\widetilde{\square}}$ and the function $\mathbf{f}$ periodically from the cube $\widetilde{\square}$ to $\mathbb{Z}^d$, and let $\widetilde{w}$ be the solution of the elliptic system

$$-\Delta_\phi\widetilde{w} + \mathcal{L}_{\widetilde{\square},y}\widetilde{w} + \mathcal{L}_{\mathrm{spat},z}\widetilde{w} = \mathrm{d}_y\mathrm{d}_z\mathbf{f} \quad \text{in } \mathbb{Z}^d\times\mathbb{Z}^d\times\Omega. \tag{6.56}$$

Given a point $y_1\in\square$, we denote by $[y_1]\subseteq\mathbb{Z}^d$, the set of vertices $\widetilde{y}_1\in\widetilde{\square}$ whose coordinate are in absolute value equal to the coordinates of y_1 and the reflections of this set. This definition together with (6.55) ensures that for any $y_1\in\square$ and any $\widetilde{y}_1\in[y_1]$ and $i,j\in\{1,\ldots,d\}$

$$|\mathbf{f}_{ij}(\widetilde{y}_1,z,\phi)| = |\mathbf{f}_{ij}(y_1,z,\phi)|.$$

One can verify that, with this construction, the restriction of the function $\widetilde{w}$ to the subcube $\square$ satisfies the elliptic system (6.54); it is thus equal to the function w. We now study the function $\widetilde{w}$. We denote by $\widetilde{\mathcal{G}}_{\mathrm{sec}}$ the Green's matrix associated with the operator $-\Delta_\phi+\mathcal{L}_{\widetilde{\square},y}+\mathcal{L}_{\mathrm{spat},z}$. As was already mentioned, the operator $\mathcal{L}_{\widetilde{\square}}$ is a perturbation of the Laplacian $\frac{1}{2\beta}\Delta$; as a consequence, one can apply the same proofs as the ones written in Section 5, and obtain the same results. In particular the statement of Proposition 5.13 holds for the Green's matrix $\widetilde{\mathcal{G}}_{\mathrm{sec}}$. Using that the function $\widetilde{w}$ solves the system (6.56), we obtain the explicit formula

$$\nabla_y \widetilde{w}(y,z,\phi) = \sum_{y_1, z_1 \in \mathbb{Z}^d} \nabla_y \mathrm{d}^*_{y_1} \mathrm{d}^*_{z_1} \widetilde{\mathcal{G}}_{\mathrm{sec}, \mathbf{f}(y_1,z_1,\cdot)} \left(y,z,\phi; y_1, z_1\right).$$

Using the statement of Proposition 5.13, we obtain the estimate on the $L^2(\mu_\beta)$-norm of the function $\widetilde{w}$, for any $y \in \square$ and any $z \in \mathbb{Z}^d$,

$$\begin{aligned} \|\nabla_y \widetilde{w}(y,z,\cdot)\|_{L^2(\mu_\beta)} &\leq C \sum_{y_1,z_1 \in \mathbb{Z}^d} \frac{\|\mathbf{f}(y_1,z_1,\cdot)\|_{L^2(\mu_\beta)}}{|y_1 - y|^{2d+\frac{3}{4}} + |z_1 - z|^{2d+\frac{3}{4}}} \\ &\leq C \sum_{y_1 \in \square, z_1 \in \mathbb{Z}^d} \|\mathbf{f}(y_1,z_1,\cdot)\|_{L^2(\mu_\beta)} \sum_{\widetilde{y}_1 \in [y_1]} \frac{1}{|\widetilde{y}_1 - y|^{2d+\frac{3}{4}} + |z_1 - z|^{2d+\frac{3}{4}}}. \end{aligned} \tag{6.57}$$

We first estimate the second sum in the right-hand side, and obtain

$$\sum_{\widetilde{y}_1 \in [y_1]} \frac{1}{|\widetilde{y}_1 - y|^{2d+\frac{3}{4}} + |z_1 - z|^{2d+\frac{3}{4}}} \leq \frac{1}{|y_1 - y|^{2d+\frac{3}{4}} + |z_1 - z|^{2d+\frac{3}{4}}} + \frac{1}{3^{nd} \max(3^n, |z_1 - z|)^{d+\frac{3}{4}}}.$$

To compute (6.57), we prove the estimate, for each pair of points $y_1 \in \square$ and $z_1 \in \mathbb{Z}^d$,

$$\|\mathbf{f}(y_1, y_1 + z_1, \cdot)\|_{L^2(\mu_\beta)} \leq C e^{-c\sqrt{\beta}|z_1|} \sum_{y_0 \in \square} e^{-c\sqrt{\beta}|y_0 - y_1|} \|\nabla v(y_0,\cdot)\|_{L^2(\mu_\beta)}. \tag{6.58}$$

Let us make a comment about the estimate (6.58). Due to the exponential decay $|z(\beta,q)| \leq Ce^{-c\sqrt{\beta}\|q\|_1}$, the function $\mathbf{f}$ decays exponentially fast outside the diagonal $\{(y,y) : y \in \square\} \subseteq \mathbb{Z}^{2d}$. This phenomenon can be observed in the inequality (6.58): the exponential term $e^{-c\sqrt{\beta}|z_1|}$ is small when the norm of z_1 is large, i.e., when the point $(y_1, y_1 + z_1)$ is far from the diagonal $\{(y,y) \in \square \times \square\}$. Furthermore, on the diagonal, the term $\|\mathbf{f}(y_1,y_1,\cdot)\|_{L^2(\mu_\beta)}$ is approximately equal to the value $\|\nabla v(y_1,\cdot)\|_{L^2(\mu_\beta)}$; but again the sum over all the charges needs to be taken into consideration and explains the sum over all the radii in the right side of (6.58) with the exponential decay $e^{-c\sqrt{\beta}r}$.

We now prove the estimate (6.58). We start from the inequality, for each pair of points $y_1 \in \square$ and $z_1 \in \mathbb{Z}^d$,

$$\|\mathbf{f}(y_1, y_1 + z_1, \cdot)\|_{L^2(\mu_\beta)} \leq \sum_{q \in \mathcal{Q}} \sum_{y \in \operatorname{supp} n_q} e^{-c\sqrt{\beta}\|q\|_1} \|\nabla v(y,\cdot)\|_{L^2(\mu_\beta)} \|n_q\|_{L^\infty} |n_q(y_1)| |n_q(y_1 + z_1)|. \tag{6.59}$$

We then note that if a charge q is such that the two points y_1 and $y_1 + z_1$ belong to the support of n_q, then the diameter of n_q is larger than $|z_1|$, and thus the diameter of q is larger than $c|z_1|$, for some constant $c(d) > 0$. From this remark, we deduce that

$$\sum_{q \in \mathcal{Q}} e^{-c\sqrt{\beta}\|q\|_1} \|n_q\|_{L^\infty} |n_q(y_1)| |n_q(y_1 + z_1)| \leq C e^{-c\sqrt{\beta}|z_1|}. \tag{6.60}$$

Similarly, if a charge q is such that the three points y_1 and $y_1 + z_1$ and y belong to the support of n_q, then the diameter of n_q is larger than $\max(|z_1|, |y - y_1|) \geq \frac{|z_1| + |y - y_1|}{2}$. This argument implies that the diameter of q has to be larger than $c(|z_1| + |y - y_1|)$, and we deduce that

$$\sum_{q \in \mathcal{Q}} e^{-c\sqrt{\beta}\|q\|_1} \mathbb{1}_{\{y \in \operatorname{supp} n_q\}} \|n_q\|_{L^\infty} |n_q(y_1)| |n_q(y_1 + z_1)| \leq C e^{-c\sqrt{\beta}(|z_1| + |y - y_1|)}. \tag{6.61}$$

Combining the estimates (6.59), (6.60), and (6.61), we obtain

$$\begin{aligned} \|\mathbf{f}(y_1, y_1 + z_1, \cdot)\|_{L^2(\mu_\beta)} &\leq \sum_{q \in \mathcal{Q}} \sum_{y_0 \in \operatorname{supp} n_q} e^{-c\sqrt{\beta}\|q\|_1} \|\nabla v(y_0,\cdot)\|_{L^2(\mu_\beta)} \|n_q\|_{L^\infty} n_q(y_1) n_q(y_1 + z_1) \\ &\leq \sum_{y_0 \in \square} \sum_{q \in \mathcal{Q}} e^{-c\sqrt{\beta}\|q\|_1} \|\nabla v(y_0,\cdot)\|_{L^2(\mu_\beta)} \mathbb{1}_{\{y_0 \in \operatorname{supp} n_q\}} \|n_q\|_{L^\infty} n_q(y_1) n_q(y_1 + z_1) \\ &\leq C e^{-c\sqrt{\beta}|z_1|} \sum_{y_0 \in \square} e^{-c\sqrt{\beta}|y_0 - y_1|} \|\nabla v(y_0,\cdot)\|_{L^2(\mu_\beta)}, \end{aligned}$$

and we have proved the inequality (6.58).

We now come back to the estimate (6.57), fix a point $y \in \square$, and use the estimate (6.58). We obtain

$$\left\| \nabla_y \widetilde{w}(y,z,\phi) \right\|_{L^2(\mu_\beta)} \leq C \sum_{y_0, y_1 \in \square, z_1 \in \mathbb{Z}^d} \frac{e^{-c\sqrt{\beta}(|z_1 - y_1| + |y_0 - y_1|)} \left\| \nabla v\left(y_0, \cdot\right) \right\|_{L^2(\mu_\beta)}}{|y_1 - y|^{2d+\frac{3}{4}} + |z_1 - z|^{2d+\frac{3}{4}}} + C \sum_{y_0, y_1 \in \square, z_1 \in \mathbb{Z}^d} \frac{e^{-c\sqrt{\beta}(|z_1 - y_1| + |y_0 - y_1|)} \left\| \nabla v\left(y_0, \cdot\right) \right\|_{L^2(\mu_\beta)}}{3^{nd} \max(3^n, |z_1 - z|)^{d+\frac{3}{4}}}. \tag{6.62}$$

We first estimate the second term in the right-hand side and write

$$\sum_{y_0, y_1 \in \square, z_1 \in \mathbb{Z}^d} \frac{e^{-c\sqrt{\beta}(|z_1 - y_1| + |y_0 - y_1|)} \left\| \nabla v\left(y_0, \cdot\right) \right\|_{L^2(\mu_\beta)}}{3^{dn} \max(3^n, |z_1 - z|)^{d+\frac{3}{4}}} \leq \frac{C \left\| \nabla v \right\|_{\underline{L}^2(\square, \mu_\beta)}}{\max(3^n, |z|)^{d+\frac{3}{4}}} \leq \frac{C}{\max(3^n, |z|)^{d+\frac{3}{4}}} \leq \frac{C}{|z-y|^{d+\frac{3}{4}}}, \tag{6.63}$$

where we have used that $y \in \frac{1}{3}\square$ to obtain the last inequality. We then estimate the first term in the right-hand side of (6.62) and focus on the sum over the variables y_1 and z_1. The exponential decay of the terms $e^{-c\sqrt{\beta}|z_1 - y_1|}$ and $e^{-c\sqrt{\beta}|y_0 - y_1|}$ forces the sum to contract on the points $y_1 = y_0$ and $z_1 = y_0$. We have the inequality,

$$\sum_{y_1, z_1 \in \mathbb{Z}^d} \frac{e^{-c\sqrt{\beta}(|z_1 - y_1| + |y_0 - y_1|)}}{|y_1 - y|^{2d+\frac{3}{4}} + |z_1 - z|^{2d+\frac{3}{4}}} \leq \frac{C}{|y_0 - y|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}}}.$$

Using the previous estimate, we can simplify the inequality (6.62), and we obtain

$$\left\| \nabla_y \widetilde{w}(y,z,\phi) \right\|_{L^2(\mu_\beta)} \leq C \sum_{y_0 \in \square} \frac{\left\| \nabla v(y_0, \cdot) \right\|_{L^2(\mu_\beta)}}{|y_0 - y|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}}}.$$

We then truncate the sum, depending on whether the point y_0 belongs to the cube $\frac{1}{2}\square$. We write

$$\left\| \nabla_y \widetilde{w}(y,z,\phi) \right\|_{L^2(\mu_\beta)} \leq C \underbrace{\sum_{y_0 \in \frac{1}{2}\square} \frac{\left\| \nabla v\left(y_0, \cdot\right) \right\|_{L^2(\mu_\beta)}}{|y_0 - y|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}}}}_{(6.64)-(i)} + C \underbrace{\sum_{y_0 \in \square \setminus \frac{1}{2}\square} \frac{\left\| \nabla v\left(y_0, \cdot\right) \right\|_{L^2(\mu_\beta)}}{|y_0 - y|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}}}}_{(6.64)-(ii)}. \tag{6.64}$$

We treat the two terms in the right side of (6.64) separately. For the term (6.64)-(i), we use that the map v is a solution of the Helffer-Sjöstrand equation (6.52) in the cube $\square$, and apply Proposition 5.6 with the regularity exponent $\varepsilon = \frac{1}{4}$. We obtain, for each point $y_0 \in \frac{1}{2}\square$,

$$\left\| \nabla v\left(y_0, \cdot\right) \right\|_{L^2(\mu_\beta)} \leq C \left(l_{n+1}\right)^{\frac{1}{2}} \left\| \nabla v \right\|_{\underline{L}^2(\square, \mu_\beta)} \leq C 3^{\frac{n}{2}}, \tag{6.65}$$

where we used Remark 6.7 and the inequality (6.7) in the second inequality. Using the estimate (6.65), we can compute the term (6.64)-(i)

$$\begin{aligned} \sum_{y_0 \in \frac{1}{2}\square} \frac{\left\| \nabla v\left(y_0, \cdot\right) \right\|_{L^2(\mu_\beta)}}{|y_0 - y|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}}} & \leq C 3^{\frac{n}{2}} \sum_{y_0 \in \frac{1}{2}\square} \frac{1}{|y_0 - y|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}}} \\ & \leq C 3^{\frac{n}{2}} \sum_{y_0 \in \mathbb{Z}^d} \frac{1}{|y_0 - y|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}}} \\ & \leq C \frac{3^{\frac{n}{2}}}{|y-z|^{d+\frac{3}{4}}}, \end{aligned} \tag{6.66}$$

where we used the result of Appendix C in the last inequality. We now treat the term (6.64)-(ii). In that case, we use the estimate $|y - y_0| \geq c|y_0|$, valid for any point $y_0 \in \mathbb{Z}^d \setminus \frac{1}{2}\square$ and any point $y \in \frac{1}{3}\square$. We obtain the inequality

$$\sum_{y_0 \in \square \setminus \frac{1}{2}\square} \frac{\left\| \nabla v\left(y_0, \cdot\right) \right\|_{L^2(\mu_\beta)}}{|y_0 - y|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}}} \leq C \sum_{y_0 \in \square \setminus \frac{1}{2}\square} \frac{\left\| \nabla v\left(y_0, \cdot\right) \right\|_{L^2(\mu_\beta)}}{|y_0|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}}}.$$

We then note that, for any point $y_0 \in \square \setminus \frac{1}{2}\square$, and each point $z \in \mathbb{Z}^d$, one has the inequalities

$$c \max(3^n, |z|)^{2d+\frac{3}{4}} \leq |y_0|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}} \leq C \max(3^n, |z|)^{2d+\frac{3}{4}}. \tag{6.67}$$

We thus deduce that

$$(6.68) \qquad \sum_{y_0 \in \square \setminus \frac{1}{2}\square} \frac{\|\nabla v(y_0, \cdot)\|_{L^2(\mu_\beta)}}{|y_0|^{2d+\frac{3}{4}} + |y_0 - z|^{2d+\frac{3}{4}}} \leq \frac{C}{\max(|z|, 3^n)^{d+\frac{3}{4}}} \leq \frac{C}{|z-y|^{d+\frac{3}{4}}}.$$

By combining the estimates (6.63), (6.64), (6.66) and (6.68), we deduce that

$$(6.69) \qquad \|\nabla_y \widetilde{w}(y, z, \cdot)\|_{L^2(\mu_\beta)} \leq \frac{C 3^{\frac{n}{4}}}{|z-y|^{d+\frac{3}{4}}}.$$

We complete the argument by recalling that, for each $y \in \square$, and each $z \in \mathbb{Z}^d$, the function $\widetilde{w}$ is defined so that we have $\nabla_y \widetilde{w}(y, z, \cdot) = \partial_z \nabla v(y, \cdot, \square)$. The inequality (6.69) can thus be rewritten

$$\|\partial_z \nabla v(y, \cdot, \square, p^*)\| \leq \frac{C 3^{\frac{n}{4}}}{|y-z|^{d+\frac{3}{4}}}.$$

The proof of the inequality (6.51), and thus of Step 2 is complete. □

6.3.3. *Control over the L^2-norms of the functions $u - l_p$ and $v - \mathbf{a}_*(\square_n)^{-1} l_{p^*}$.* The objective of this section is to prove that the optimizers $u(\cdot, \cdot, \square_{n+1}^-, p)$ and $v(\cdot, \cdot, \square_{n+2}, p^*)$ are close in the $\underline{L}^2(\square_n, \mu_\beta)$-norm to affine functions. The result relies on the multiscale Poincaré inequality stated in Appendix B, and is quantified in terms of the subadditivity defect τ_n.

Lemma 6.22 (L^2 estimate for the optimizers u and v)**.** *There exist an inverse temperature $\beta_0 := \beta_0(d) < \infty$ and a constant $C := C(d) < \infty$ such that, for each $\beta > \beta_0$, each integer $n \in \mathbb{N}$, and each pair of vectors $p, p^* \in \mathbb{R}^{d \times \binom{d}{2}}$,*

$$(6.70) \qquad \|u(\cdot, \cdot, \square_{n+1}^-, p) - l_p\|^2_{\underline{L}^2(\square_{n+1}, \mu_\beta)} \leq C|p|^2 3^{2n} \left(3^{-\frac{n}{2}} + \sum_{m=0}^{n} 3^{-\frac{n-m}{2}} \tau_m\right),$$

and

$$(6.71) \qquad \left\| v(\cdot, \cdot, \square_{n+2}, p^*) - l_{\bar{\mathbf{a}}_n^{-1} p^*} - (v(\cdot, \cdot, \square_{n+2}, p^*))_{\square_{n+1}, \mu_\beta} \right\|^2_{\underline{L}^2(\square_{n+1}, \mu_\beta)} \leq C|p^*|^2 3^{2n} \left(3^{-\frac{n}{2}} + \sum_{m=0}^{n+1} 3^{-\frac{n-m}{2}} \tau_m\right).$$

Proof. We assume without loss of generality that $|p| = 1$ and $|p^*| = 1$. To ease the notation, we denote by $u := u(\cdot, \cdot, \square_{n+1}^-, p)$ and by $v := v(\cdot, \cdot, \square_{n+2}, p^*)$. The strategy of the proof relies on two ingredients:

- First, we need to estimate the spatial averages of the gradients of the functions $u - l_p$ and $v - l_{\mathbf{a}_*(\square_n)^{-1} p^*}$ and prove that they are small. To be more precise, we estimate these spatial averages in terms of the subadditivity defects τ_n. The proof relies on different arguments depending on which function we consider:
  - For the function u, we use the subadditivity property stated in Proposition 6.13 and the following fact: for any discrete cube $\square \subseteq \mathbb{Z}^d$ and any function $f : \square \to \mathbb{R}$ which is equal to 0 on the boundary of the cube $\square$, one has the identity
  $$\sum_{x \in \square} \nabla f(x) = 0;$$
  - For the function v, we use the subadditivity property stated in Proposition 6.13, and Lemma 6.20 to control the variance of the spatial average of its gradient.
- The multiscale Poincaré inequality, which is stated in Proposition B.1 in Appendix B. This inequality allows to estimate the L^2-norm of a function in terms of the spatial averages of its gradient.

We first focus on the function $u := u(\cdot, \cdot, \square_{n+1}^-, p)$, and prove the inequality (6.70). We first recall that the function u is extended by the affine function l_p outside the cube $\square_{n+1}^-$. We thus have

$$\|u(\cdot, \cdot, \square_{n+1}^-, p) - l_p\|^2_{L^2(\square_{n+1}, \mu_\beta)} = \|u(\cdot, \cdot, \square_{n+1}^-, p) - l_p\|^2_{L^2(\square_{n+1}^-, \mu_\beta)}.$$

By the multiscale Poincaré inequality stated in Proposition B.1 of Appendix B, we have

$$(6.72)\quad \left\| u(\cdot,\cdot,\square_{n+1}^-,p) - l_p \right\|^2_{\underline{L}^2(\square_{n+1},\mu_\beta)} \\ \leq \underbrace{C\left\| \nabla u(\cdot,\cdot,\square_{n+1}^-,p) - p \right\|^2_{\underline{L}^2(\square_{n+1},\mu_\beta)}}_{(6.72)-(i)} + \underbrace{C3^n \sum_{m=0}^{n} \frac{3^m}{|\mathcal{Z}_{m,n}|} \sum_{z\in\mathcal{Z}_{m,n}} \left\langle \left(\frac{1}{|z+\square_m|} \sum_{x\in z+\square_m} \nabla u(\cdot,\cdot,\square_{n+1}^-,p) - p \right)^2 \right\rangle_{\mu_\beta}}_{(6.72)-(ii)}.$$

We bound the term (6.72)-(i) using the estimate (6.7). We obtain the inequality

$$(6.73)\quad \left\| \nabla u(\cdot,\cdot,\square_{n+1}^-,p) - p \right\|^2_{\underline{L}^2(\square_{n+1},\mu_\beta)} \leq 2\left\| \nabla u(\cdot,\cdot,\square_{n+1},p) \right\|^2_{\underline{L}^2(\square_{n+1}^-,\mu_\beta)} + 2|p^2| \leq C|p^2|.$$

To estimate the term (6.72)-(ii), we use the two following ingredients:

- The subadditivity of the energy ν which is stated in Proposition 6.13 and Remark 6.14. It reads, for each integer $m\in\{1,\ldots,n\}$,

$$|\mathcal{Z}_{m,n}|^{-1} \sum_{z\in\mathcal{Z}_{m,n}} \left[\!\left[u(\cdot,\cdot,\square_{n+1}^-,p) - u(\cdot,\cdot,z+\square_m^-,p) \right]\!\right]^2_{\underline{H}^1(z+\square_m^-,\mu_\beta)} \leq C\left(\nu\left(\square_m^-,p\right) - \nu\left(\square_{n+1}^-,p\right) + 3^{-\frac{m}{2}}|p|^2 \right) \\ \leq C\left(\sum_{k=m}^{n} \tau_k + 3^{-\frac{m}{2}}|p|^2 \right).$$

- For each point $z\in\mathcal{Z}_{m,n}$, the function $u(\cdot,z+\square_m^-,p)$ belongs to the space $l_p + H_0^1\left(z+\square_m^-,\mu_\beta\right)$. This implies that, for each realization of the field $\phi\in\Omega$,

$$(6.74)\quad \frac{1}{|z+\square_m|} \sum_{x\in z+\square_m} \nabla u(x,\phi,z+\square_m^-,p) = p.$$

We deduce the inequality, for each integer $m\in\{1,\ldots,n\}$,

$$(6.75)\quad \sum_{z\in\mathcal{Z}_{m,n}} \frac{1}{|\mathcal{Z}_{m,n}|} \left\langle \left(\frac{1}{|z+\square_m|} \sum_{x\in z+\square_m} \nabla u(x,\cdot,\square_{n+1}^-,p) - p \right)^2 \right\rangle_{\mu_\beta} \leq C\left(\sum_{k=m}^{n} \tau_k + 3^{-\frac{m}{2}}|p|^2 \right).$$

Combining the estimates (6.72), (6.73), and (6.75) completes the proof of the estimate (6.70).

We now prove the inequality (6.71). By the multiscale Poincaré inequality, we have

$$(6.76)\quad \left\| v(\cdot,\cdot,\square_{n+2},p^*) - l_{\bar{\mathbf{a}}_n^{-1}p^*} - \left(v(\cdot,\cdot,\square_{n+2},p^*) - l_{\bar{\mathbf{a}}_n^{-1}p^*} \right)_{\square_{n+1}} \right\|^2_{\underline{L}^2(\square_{n+1},\mu_\beta)} \\ \leq \underbrace{C\left\| \nabla v(\cdot,\square_{n+2},p^*) - \bar{\mathbf{a}}_n^{-1}p^* \right\|^2_{\underline{L}^2(\square_{n+1},\mu_\beta)}}_{(6.76)-(i)} + \underbrace{C3^n \sum_{m=0}^{n} \frac{3^m}{|\mathcal{Z}_{m,n}|} \left\langle \left(\frac{1}{|z+\square_m|} \sum_{x\in z+\square_m} \nabla v(\cdot,\square_{n+2},p^*) - \bar{\mathbf{a}}_n^{-1}p^* \right)^2 \right\rangle_{\mu_\beta}}_{(6.76)-(ii)}.$$

We first treat the term on the left side. Since the average value of a linear map on a cube centered at 0 is equal to 0, we have that

$$\left(v(\cdot,\cdot,\square_{n+2},p^*) - l_{\bar{\mathbf{a}}_n^{-1}p^*} \right)_{\square_{n+1}} = \frac{1}{|\square_{n+1}|} \sum_{x\in\square_{n+1}} v(x,\cdot,\square_{n+2},p^*).$$

We then use the estimate (6.6) and the inclusion $\square_{n+1} \subseteq \frac{1}{3}\square_{n+2}$. We obtain

$$(6.77)\quad \left\| \left(v(\cdot,\cdot,\square_{n+2},p^*) - l_{\bar{\mathbf{a}}_n^{-1}p^*} \right)_{\square_{n+1}} - \left(v(\cdot,\cdot,\square_{n+2},p^*) \right)_{\square_{n+1},\mu_\beta} \right\|^2_{L^2(\mu_\beta)} = \mathrm{var}_{\mu_\beta}\left[\left(v(\cdot,\cdot,\square_{n+2},p^*) \right)_{\square_{n+1}} \right] \\ \leq \frac{C}{|\square_n|} \sum_{x\in\square_{n+1}} \mathrm{var}_{\mu_\beta}\left[v(x,\cdot,\square_{n+2},p^*) \right] \\ \leq \frac{C}{|\square_n|} \sum_{x\in\frac{1}{3}\square_{n+2}} \mathrm{var}\left[v(x,\cdot,\square_{n+2},p^*) \right] \\ \leq C|p^*|.$$

We now treat the terms in the right side of (6.76). The term (6.76)-(i) can be estimated with the same argument as in the inequality (6.73). We obtain

$$\left\| \nabla v(\cdot, \cdot, \square_{n+2}, p^*) - \bar{\mathbf{a}}_n^{-1} p^* \right\|_{\underline{L}^2(\square_{n+1}, \mu_\beta)}^2 \leq C. \tag{6.78}$$

To estimate the term (6.76)-(ii), we prove that, for each integer $m \in \{1, \ldots, n\}$,

$$\frac{1}{|\mathcal{Z}_{m,n}|} \left\langle \left(\frac{1}{|z + \square_m|} \sum_{x \in z + \square_m} \nabla v(\cdot, \cdot, \square_{n+2}, p^*) - \bar{\mathbf{a}}_n^{-1} p^* \right)^2 \right\rangle_{\mu_\beta} \leq C 3^{-\frac{m}{2}} + C \sum_{k=m}^{n} \tau_k. \tag{6.79}$$

To this end, we decompose the left side of (6.79) and write

$$\begin{aligned} &\frac{1}{|\mathcal{Z}_{m,n}|} \sum_{z \in \mathcal{Z}_{m,n}} \left\langle \left(\frac{1}{|z + \square_m|} \sum_{x \in z + \square_m} \nabla v(\cdot, \cdot, \square_{n+2}, p^*) - \bar{\mathbf{a}}_n^{-1} p^* \right)^2 \right\rangle_{\mu_\beta} \\ &\qquad \leq 3 \left| \mathcal{Z}_{m,n} \right|^{-1} \sum_{z \in \mathcal{Z}_{m,n}} \left[\!\left[v(\cdot, \cdot, \square_{n+2}, p^*) - v(\cdot, \cdot, z + \square_m, p^*) \right]\!\right]_{\underline{H}^1(z + \square_m, \mu_\beta)}^2 \\ &\qquad \quad + 3 \left| \bar{\mathbf{a}}_n^{-1} p^* - \bar{\mathbf{a}}_m^{-1} p^* \right|^2 \\ &\qquad \quad + 3 \left| \mathcal{Z}_{m,n} \right|^{-1} \sum_{z \in \mathcal{Z}_{m,n}} \left\langle \left(\frac{1}{|z + \square_m|} \sum_{x \in z + \square_m} \nabla v(\cdot, \cdot, z + \square_m, p^*) - \bar{\mathbf{a}}_m^{-1} p^* \right)^2 \right\rangle_{\mu_\beta}. \end{aligned} \tag{6.80}$$

We estimate the first term on the right side by Proposition 6.17, and the second term by Lemma 6.19. We obtain

$$\left| \mathcal{Z}_{m,n} \right|^{-1} \sum_{z \in \mathcal{Z}_{m,n}} \left[\!\left[v(\cdot, \cdot, \square_{n+2}, p^*) - v(\cdot, \cdot, z + \square_m, p^*) \right]\!\right]_{\underline{H}^1(z + \square_m, \mu_\beta)}^2 + \left| \bar{\mathbf{a}}_n^{-1} p^* - \bar{\mathbf{a}}_m^{-1} p^* \right|^2 \leq C 3^{-\frac{m}{2}} + C \sum_{k=m}^{n} \tau_k. \tag{6.81}$$

There remains to estimate the third term in the right side of (6.80). We first recall the identity, for each integer $m \in \mathbb{N}$,

$$\frac{1}{|\square_m|} \sum_{x \in \square_m} \left\langle \nabla v(x, \cdot, \square_m, p^*) \right\rangle_{\mu_\beta} = \bar{\mathbf{a}}_m^{-1} p^*.$$

We use the translation invariance of the measure μ_β and Lemma 6.20. To ease the notation, we note that in dimension larger than 3, we have the estimate $d - \frac{5}{2} \geq \frac{1}{2}$. We obtain

$$\begin{aligned} &\left| \mathcal{Z}_{m,n} \right|^{-1} \sum_{z \in \mathcal{Z}_{m,n}} \left\langle \left(\frac{1}{|z + \square_m|} \sum_{x \in z + \square_m} \nabla v(x, \cdot, z + \square_m, p^*) - \bar{\mathbf{a}}_m^{-1} p^* \right)^2 \right\rangle_{\mu_\beta} = \left\langle \left(\frac{1}{|\square_m|} \sum_{x \in \square_m} \nabla v(x, \cdot, \square_m, p^*) - \bar{\mathbf{a}}_m^{-1} p^* \right)^2 \right\rangle_{\mu_\beta} \\ &\qquad = \mathrm{var}_{\mu_\beta} \left[\frac{1}{|\square_m|} \sum_{x \in \square_n} \nabla v(x, \cdot, \square_m, p^*) \right] \\ &\qquad \leq \mathrm{var}_{\mu_\beta} \left[\frac{1}{|\square_m|} \sum_{x \in \square_m} \nabla v(x, \cdot, \square_{m+1}, p^*) \right] + \tau_m \\ &\qquad \leq C \left(3^{-\frac{m}{2}} + \tau_m \right). \end{aligned} \tag{6.82}$$

Combining the estimates (6.77), (6.78), (6.80), (6.81), and (6.82) completes the proof of (6.71). □

6.3.4. *Control over the energy J.* In this section, we obtain from the previous results and the Caccioppoli inequality a quantitative control over the energy quantity $J(\square_n, p, \bar{\mathbf{a}}_n p)$. The argument needs to take into account the infinite range of the Helffer-Sjöstrand operator and the specific forms of the energies $\mathbf{E}$ and $\mathbf{E}^*$ which causes some technicalities in the analysis. The result is stated in the lemma below.

Lemma 6.23. *There exist an inverse temperature $\beta_0 := \beta_0(d) < \infty$ and a constant $C := C(d) < \infty$ such that for each $\beta \geq \beta_0$, each integer $n \in \mathbb{N}$, and each $p \in \mathbb{R}^{d \times \binom{d}{2}}$,*

$$J(\square_n, p, \bar{\mathbf{a}}_n p) \leq C |p|^2 \left(3^{-\frac{n}{2}} + \sum_{m=0}^{n+1} 3^{-\frac{n-m}{2}} \tau_m \right). \tag{6.83}$$

Proof. The strategy of the proof relies on three ingredients: the Caccioppoli inequality stated in Proposition 5.1, the one-sided convex duality formula (6.20) stated in Proposition 6.12, and the L^2-norm estimate on the optimizers u and v stated in Lemma 6.22.

We fix a slope $p \in \mathbb{R}^d$ and assume without loss of generality that $|p| = 1$. By Proposition 6.12, we have the identity

$$J(\square_n, p, \bar{\mathbf{a}}_n p) = \mathbf{E}^*_{\square_n}\left[u(\cdot,\cdot,\square_n^-, p) - v(\cdot,\cdot,\square_n, \bar{\mathbf{a}}_n p)\right] + O\left(C 3^{-\frac{m}{2}}\right).$$

To prove the estimate (6.83), it is thus sufficient to prove the estimate

$$\mathbf{E}^*_{\square_n}\left[u(\cdot,\cdot,\square_n^-, p) - v(\cdot,\cdot,\square_n, \bar{\mathbf{a}}_n p)\right] \leq C\left(3^{-\frac{n}{2}} + \sum_{m=0}^{n+1} 3^{-\frac{n-m}{2}} \tau_m\right). \tag{6.84}$$

Using the coercivity of the energy $\mathbf{E}^*_{\square_n}$ stated in (6.2), we see that to prove the inequality (6.84), it is sufficient to prove the estimate

$$\left[\!\left[u(\cdot,\cdot,\square_n^-, p) - v(\cdot,\cdot,\square_n, \bar{\mathbf{a}}_n p)\right]\!\right]^2_{\underline{H}^1(\square_n,\mu_\beta)} \leq C\left(3^{-\frac{n}{2}} + \sum_{m=0}^{n+1} 3^{-\frac{n-m}{2}} \tau_m\right), \tag{6.85}$$

and by Propositions 6.13 and 6.17, we see that to prove (6.85) it is sufficient to prove

$$\left[\!\left[u(\cdot,\cdot,\square_{n+1}^-, p) - v(\cdot,\cdot,\square_{n+2}, \bar{\mathbf{a}}_n p)\right]\!\right]^2_{\underline{H}^1(\square_n,\mu_\beta)} \leq C\left(3^{-\frac{n}{2}} + \sum_{m=0}^{n+1} 3^{-\frac{n-m}{2}} \tau_m\right). \tag{6.86}$$

We now focus on the proof of (6.86). In the rest of the proof, we make use of the notation $u := u(\cdot,\cdot,\square_{n+1}^-, p)$ and $v := v(\cdot,\cdot,\square_{n+2}, \bar{\mathbf{a}}_n p) - \left(v(\cdot,\cdot,\square_{n+2}, \bar{\mathbf{a}}_n p)\right)_{\square_{n+1},\mu_\beta}$. By Lemma 6.22, we have the $\underline{L}^2(\square_{n+1},\mu_\beta)$-estimate

$$\|u - v\|^2_{\underline{L}^2(\square_{n+1},\mu_\beta)} \leq 2\|u - l_p\|^2_{\underline{L}^2(\square_{n+1},\mu_\beta)} + 2\|v - l_p\|^2_{\underline{L}^2(\square_{n+1},\mu_\beta)} \leq C3^{2n}\left(3^{-\frac{n}{2}} + \sum_{m=0}^{n+1} 3^{-\frac{n-m}{2}} \tau_m\right). \tag{6.87}$$

We recall the following notation: for each integer $k \in \mathbb{N}$, we denote by $\square_{n+2}^k$ the interior cube $\square_{n+2}^k := \{x \in \square_{n+2} : \operatorname{dist}(x, \partial\square_{n+2}) \geq k\}$. By the first variation formula stated in Proposition 6.12, the maps u and v are solutions of the equations

$$\mathcal{L}u = 0 \text{ in } \square_{n+1}^- \times\Omega \quad \text{and} \quad \mathcal{L}_{\square_{n+2}} v = 0 \text{ in } \square_{n+2} \times\Omega,$$

where we recall the definition of the Helffer-Sjöstrand operator $\mathcal{L}_{\square_{n+2}}$

$$\mathcal{L}_{\square_{n+2}} := -\Delta_\phi - \frac{1}{2\beta}\Delta + \frac{1}{2\beta}\sum_{k\geq1}\frac{(-1)^{k+1}}{\beta^{\frac{k}{2}}}\nabla^{k+1}\cdot\left(\mathbb{1}_{\square^k_{n+2}}\nabla^{k+1}\right) - \frac{1}{\beta^{\frac{5}{4}}}\nabla\cdot\left(\mathbb{1}_{\square_{n+2}\setminus\square^-_{n+2}}\nabla\right) + \sum_{\operatorname{supp} q\subseteq\square_{n+2}}\nabla_q\cdot\mathbf{a}_q\nabla_q.$$

One can adapt the proof of the Caccioppoli inequality (Proposition 5.1) to the operator $\mathcal{L}_{\square_{n+2}}$ and obtain the following statement. There exists a constant $C := C(d) < \infty$ such that for any vector fields $F : \square_{n+2}\times\Omega \to \mathbb{R}^{d\times\binom{d}{2}}$ and $G : \square_{n+2}\times\Omega\to\mathbb{R}^d$, any ball $B(x,r)$ such that $B(x,2r)$ is included in the cube $\square_{n+2}$, and every solution $w : B(x,2r)\times\Omega\to\mathbb{R}^{\binom{d}{2}}$ of the equation

$$\mathcal{L}_{\square_{n+2}} w = \nabla\cdot F + \mathrm{d}G \text{ in } B(x,2r)\times\Omega,$$

one has the estimate

$$\begin{aligned}[t] [\![w]\!]_{\underline{H}^1(B_r(x),\mu_\beta)} \leq{}& \frac{C}{R}\|w\|_{\underline{L}^2(B_{2r}(x),\mu_\beta)} \\ &+ \|F\|_{\underline{L}^2(B_{2r}(x),\mu_\beta)} + \|G\|_{\underline{L}^2(B_{2r}(x),\mu_\beta)} + \sum_{y\in\square_{n+2}\setminus B_{2r}(x)} e^{-c(\ln\beta)|y-x|}\|w(y,\cdot)\|_{L^2(\mu_\beta)}. \end{aligned} \tag{6.88}$$

We then note that, by the definition of the operator $\mathcal{L}_{\square_{n+2}}$, the function u satisfies the equation

$$\mathcal{L}_{\square_{n+2}} u = \nabla\cdot F + \mathrm{d}G \text{ in } \square_{n+1}^- \times\Omega,$$

where the vector fields F and G are defined by the formulae, for each $x \in \square_{n+1}^-$,

$$F(x) := -\frac{1}{2\beta}\sum_{k\geq\operatorname{dist}(x,\partial\square_{n+2})}\frac{1}{\beta^{\frac{k}{2}}}(-\Delta)^k\nabla u(x) \quad \text{and} \quad G(x) := \sum_{\operatorname{supp} q\notin\square_{n+2}}\mathbf{a}_q\nabla_q u\times n_q(x).$$

We estimate the $L^2(\square_{n+1},\mu_\beta)$-norm of the functions F and G. We first note that every point x in the cube $\square_{n+1}$ satisfies the inequality $\operatorname{dist}(x,\partial\square_{n+2}) \geq c3^n$. Using the boundedness of the discrete Laplacian operator, the upper bound on the L^2-norm of the gradient of the function u stated in (6.7), and choosing the inverse temperature β large enough, we have

$$\|F\|_{\underline{L}^2(\square_{n+1},\mu_\beta)} \leq \sum_{k\geq c3^n} \frac{1}{\beta^{\frac{k}{2}}} \left\|\Delta^k \nabla u\right\|_{\underline{L}^2(\square_{n+1},\mu_\beta)} \leq \sum_{k\geq c3^n} \frac{C^k}{\beta^{\frac{k}{2}}} \|\nabla u\|_{\underline{L}^2(\square^-_{n+2},\mu_\beta)} \leq Ce^{-c(\ln\beta)3^{\frac{n}{2}}}. \tag{6.89}$$

Using a similar argument, we note that for each point x in the interior cube $\square_{n+1}$, if a charge $q \in \mathcal{Q}$ is such that its support is not included in the cube $\square_{n+2}$ and such that the point x belongs to the support of n_q, then its diameter must be larger than $c3^n$. We then use the exponential decay on the coefficient $\mathbf{a}_q$ and the estimate (6.7) to obtain

$$\|G\|_{\underline{L}^2(\square_{n+1},\mu_\beta)} = \left\| \sum_{\operatorname{supp} q \not\subseteq \square_{n+2}} \mathbf{a}_q(\nabla_q u)\, n_q \right\|_{\underline{L}^2(\square_n,\mu_\beta)} \leq Ce^{-c\sqrt{\beta}3^n} \|\nabla u\|_{\underline{L}^2(\square^-_{n+2},\mu_\beta)} \leq Ce^{-c\sqrt{\beta}3^n}. \tag{6.90}$$

We now apply the Caccioppoli inequality (5.5) to the function $w := u - v$, which is solution of the equation $\mathcal{L}_{\square_{n+2}}(u-v) = \nabla\cdot F + \mathrm{d}G$ in the set $\square^-_{n+1}\times\Omega$. We obtain

$$\begin{aligned}
&\beta \sum_{y\in\mathbb{Z}^d} \|\partial_y (u-v)\|_{\underline{L}^2(\square_n,\mu_\beta)} + \|\nabla(u-v)\|_{\underline{L}^2(\square_n,\mu_\beta)} \\
&\qquad \leq \underbrace{C3^{-2n}\|u-v\|^2_{\underline{L}^2(\square^-_{n+1},\mu_\beta)}}_{(6.91)-(i)} + \underbrace{\|F\|^2_{\underline{L}^2(\square^-_{n+1},\mu_\beta)} + \|G\|^2_{\underline{L}^2(\square^-_{n+1},\mu_\beta)}}_{(6.91)-(ii)} \\
&\qquad\quad + \underbrace{\left(\sum_{x\in\square_{n+2}\setminus\square^-_{n+1}} e^{-c(\ln\beta)|x|}\|u(x,\cdot)-v(x,\cdot)\|_{L^2(\mu_\beta)}\right)^2}_{(6.91)-(iii)}.
\end{aligned} \tag{6.91}$$

We estimate the term (6.91)-(i) thanks to the inequality (6.87). We obtain

$$C3^{-2n}\|u-v\|^2_{\underline{L}^2(\square^-_{n+1},\mu_\beta)} \leq C\left(3^{-\frac{n}{2}} + \sum_{m=0}^{n+1} 3^{-\frac{n-m}{2}}\tau_m\right). \tag{6.92}$$

We estimate the term (6.91)-(ii) by the inequalities (6.89) and (6.90). We obtain

$$\|F\|^2_{\underline{L}^2(\square^-_{n+1},\mu_\beta)} + \|G\|_{\underline{L}^2(\square^-_{n+1},\mu_\beta)} \leq Ce^{-c(\ln\beta)3^{\frac{n}{2}}}. \tag{6.93}$$

For the term (6.91)-(iii), we use the estimate (6.87), the observation $\tau_n \leq C$, and note that if a point x lies outside the cube $\square_n$, then its norm must be larger than $c3^n$. We obtain

$$\begin{aligned}
\sum_{x\in\square_{n+2}\setminus\square^-_{n+1}} e^{-c\sqrt{\beta}|x|}\|(u-v)(x,\cdot)\|_{L^2(\mu_\beta)} &\leq Ce^{-c\sqrt{\beta}3^n}\sum_{x\in\square_{n+2}}\|\nabla(u-v)(x,\cdot)\|_{L^2(\mu_\beta)} \\
&\leq Ce^{-c\sqrt{\beta}3^n}3^{\frac{dn}{2}}\|u-v\|_{\underline{L}^2(\square_{n+2},\mu_\beta)} \\
&\leq Ce^{-c\sqrt{\beta}3^n}.
\end{aligned} \tag{6.94}$$

Combining the estimates (6.91), (6.92), (6.93), and (6.94) completes the proof of Lemma 6.23. □

6.3.5. *Quantitative rate of convergence for the energy J.* In this section, we use Lemma 6.23 together with an iterative argument to obtain an algebraic rate of convergence for the quantity $J(\square_n, p, \bar{\mathbf{a}}_n p)$. The strategy implemented in the proof is essentially the one described in the paragraph following Proposition 6.10 up to a technical difficulty: the term in the right side of the estimate (6.83) of Lemma 6.23 is not the subadditivity defect τ_n but a weighted average the subadditivity defects. This additional technicality requires to make use of a weighted quantity denoted by $\widetilde{F}_n$ in the proof below.

Proposition 6.24. *There exist a constant $C := C(d) < \infty$ and an exponent $\alpha := \alpha(d) > 0$ such that, for each integer $n\in\mathbb{N}$, and each $p\in\mathbb{R}^{d\times\binom{d}{2}}$,*

$$J(\square_n, p, \bar{\mathbf{a}}_n p) \leq C|p|^2 3^{-\alpha n}.$$

We record, as a corollary, that the quantitative rate of convergence established in Proposition 6.24 implies a quantitative estimate on the subadditivity defect τ_n.

Corollary 6.25. *There exist a constant $C := C(d) < \infty$ and an exponent $\alpha := \alpha(d) > 0$ such that, for each integer $n \in \mathbb{N}$,*

$$-C3^{-\frac{n}{2}} \leq \tau_n \leq C3^{-\alpha n}. \tag{6.95}$$

Proof of Proposition 6.24 and Corollary 6.25. We let B_1 be the unit ball in $\mathbb{R}^{d\binom{d}{2}}$. We denote by C_0 the constant which appears in the right side of the identity (6.20), and define, for each integer $n \in \mathbb{N}$,

$$F_n := \sup_{p \in B_1} \nu\left(\square_n^-, p\right) + \nu^*\left(\square_n, \bar{\mathbf{a}}_n p\right) - \bar{\mathbf{a}}_n |p|^2 + C_0 |p|^2 3^{-\frac{n}{2}}.$$

We note that by the inequality (6.5), we have the upper bound, for each integer $n \in \mathbb{N}$, $F_n \leq C$. By Proposition 6.12 and Lemma 6.23, we have, for each integer $n \in \mathbb{N}$,

$$0 \leq F_n \leq C\left(3^{-\frac{n}{2}} + \sum_{m=0}^{n+1} 3^{-\frac{n-m}{2}} \tau_m\right). \tag{6.96}$$

Additionally, we obtain from the subadditivity properties stated in Propositions 6.13 and 6.17

$$F_{n+1} \leq F_n + C3^{-\frac{n}{2}}. \tag{6.97}$$

Combining the estimates (6.96) and (6.97) implies that

$$0 \leq F_{n+1} \leq C\left(3^{-\frac{n+1}{2}} + \sum_{m=0}^{n+1} 3^{-\frac{n-m}{2}} \tau_m\right).$$

By definition of the subadditivity defect τ_n, and the fact that the maps $p \to \nu\left(\square_n^-, p\right) - \nu\left(\square_{n+1}^-, p\right) + C|p|^2 3^{-\frac{n}{2}}$ and $p^* \to \nu^*\left(\square_n, p^*\right) - \nu^*\left(\square_{n+1}, p^*\right) + C\left|p^*\right|^2 3^{-\frac{n}{2}}$ are quadratic and non-negative, we have

$$\begin{aligned} \tau_n &\leq C \sum_{k=1}^{d} \left(\nu\left(\square_n^-, e_k\right) - \nu\left(\square_{n+1}^-, e_k\right) + \nu^*\left(\square_n, e_k\right) - \nu^*\left(\square_{n+1}, e_k\right)\right) + C3^{-\frac{n}{2}} \\ &\leq C\left(F_n - F_{n+1} + 3^{-\frac{n}{2}}\right). \end{aligned} \tag{6.98}$$

We then define $\widetilde{F}_n := 3^{-\frac{n}{4}} \sum_{k=0}^{n} 3^{\frac{k}{4}} F_k$. From the estimates (6.96), (6.98), and the inequality $F_0 \leq C$, we deduce that

$$\begin{aligned} \widetilde{F}_n - \widetilde{F}_{n+1} = 3^{-\frac{n}{4}} \sum_{k=0}^{n} 3^{\frac{k}{4}} \left(F_k - F_{k+1}\right) - 3^{-\frac{(n+1)}{4}} F_0 &\geq 3^{-\frac{n}{4}} \sum_{k=0}^{n} 3^{\frac{k}{4}} \left(\frac{1}{C}\tau_k - 3^{-\frac{k}{2}}\right) - C3^{-\frac{n}{4}} \\ &\geq \frac{1}{C} \sum_{k=0}^{n} 3^{-\frac{(n-k)}{4}} \tau_k - \sum_{k=0}^{n} 3^{-\frac{(n-k)}{4}} 3^{-\frac{k}{2}} - C3^{-\frac{n}{4}} \\ &\geq \frac{1}{C} \sum_{k=0}^{n} 3^{-\frac{(n-k)}{4}} \tau_k - C3^{-\frac{n}{4}}. \end{aligned} \tag{6.99}$$

We then compute, by using the inequalities (6.97) and (6.96),

$$\begin{aligned} \widetilde{F}_{n+1} = 3^{-\frac{n+1}{4}} \sum_{k=0}^{n+1} 3^{\frac{k}{4}} F_k = 3^{-\frac{n}{4}} \sum_{k=0}^{n} 3^{\frac{k}{4}} F_{k+1} + 3^{-\frac{n+1}{4}} F_0 &\leq 3^{-\frac{n}{4}} \sum_{k=0}^{n} 3^{\frac{k}{4}} \left(F_k + C3^{-\frac{n}{2}}\right) + C3^{-\frac{n}{4}} \\ &\leq \widetilde{F}_n + C3^{-\frac{n}{4}}. \end{aligned}$$

We use the estimate (6.96) and write

$$\begin{aligned} \widetilde{F}_{n+1} \leq 3^{-\frac{n}{4}} \sum_{k=0}^{n} 3^{\frac{k}{4}} F_k + C3^{-\frac{n}{4}} &\leq 3^{-\frac{n}{4}} \sum_{k=0}^{n} 3^{\frac{k}{4}} \left(C3^{-\frac{k}{2}} + \sum_{m=0}^{k} 3^{-\frac{k-m}{2}} \tau_m\right) + C3^{-\frac{n}{4}} \\ &\leq C3^{-\frac{n}{4}} \sum_{k=0}^{n} 3^{-\frac{k}{4}} + 3^{-\frac{n}{4}} \sum_{k=0}^{n} 3^{-\frac{k}{4}} \sum_{m=0}^{k} 3^{-\frac{m}{2}} \tau_m + C3^{-\frac{n}{4}} \\ &\leq C \sum_{k=0}^{n} 3^{-\frac{n-k}{4}} \tau_k + C3^{-\frac{n}{4}}. \end{aligned} \tag{6.100}$$

By combining the estimates (6.99) and (6.100), we have obtained

$$\widetilde{F}_{n+1} \leq C\left(\widetilde{F}_n - \widetilde{F}_{n+1}\right) + C3^{-\frac{n}{4}}.$$

The previous inequality can be rewritten

$$\widetilde{F}_{n+1} \leq \frac{C}{C+1} \widetilde{F}_n + C 3^{-\frac{n}{4}}. \tag{6.101}$$

We set $\alpha_0 := \frac{1}{\ln 3} \ln \frac{C}{C+1}$ so that we have $3^{\alpha_0} = \frac{C}{C+1}$, and define the exponent $\alpha := \min\left(\alpha_0, \frac{1}{8}\right)$. We iterate the inequality (6.101), and note that the inequality $F_0 \leq C$ implies the inequality $\widetilde{F}_0 \leq C$. We obtain

$$\widetilde{F}_n \leq 3^{-\alpha_0 n} \widetilde{F}_0 + C \sum_{k=0}^{n} 3^{-\alpha_0 k} 3^{-\frac{n-k}{4}} \leq C 3^{-\alpha n}.$$

Finally, by the definition of the weighted sum $\widetilde{F}_n$, we have the inequality $F_n \leq \widetilde{F}_n$. The proof of Proposition 6.24 is complete.

There only remains to prove Corollary 6.25. The lower bound in (6.95) is a direct consequence subadditivity properties stated in Propositions 6.13 and 6.17. For the upper bound, we use the inequality (6.98) together with the estimates $F_n \leq C 3^{-\alpha n}$ and $F_{n+1} \geq 0$. □

6.3.6. *Quantitative rate of convergence for the subadditive quantities ν and ν^*.* In this section, we deduce Proposition 6.10 from Proposition 6.24.

Proof of Proposition 6.10. Before starting the proof, we collect some ingredients which were proved in this section:

- By Proposition 6.12 and Definition 6.38, we have the identities, for each integer $n \in \mathbb{N}$, and each $p, p^* \in \mathbb{R}^d$,

$$\nu\left(\square_n^-, p\right) = \frac{1}{2} p \cdot \mathbf{a}\left(\square_n^-\right) p \quad \text{and} \quad \nu^*\left(\square_n, p^*\right) = \frac{1}{2} p^* \cdot \bar{\mathbf{a}}_n^{-1} p^*; \tag{6.102}$$

- By Property (4) of Proposition 6.12, there exist two strictly positive constants c, C depending only on the dimension d such that, for every cube $\square \subseteq \mathbb{Z}^d$,

$$c I_{d \times \binom{d}{2}} \leq \mathbf{a}(\square), \mathbf{a}_*(\square) \leq C I_{d \times \binom{d}{2}}; \tag{6.103}$$

- By Corollaries 6.15 and 6.18, we have the convergences

$$\mathbf{a}\left(\square_n^-\right) \underset{n \to \infty}{\longrightarrow} \bar{\mathbf{a}} \quad \text{and} \quad \bar{\mathbf{a}}_n^{-1} \underset{n \to \infty}{\longrightarrow} \bar{\mathbf{a}}_*^{-1}; \tag{6.104}$$

- By the one-sided convex duality estimate (6.20) and Proposition 6.24, we have the inequalities, for each $p \in \mathbb{R}^{d \times \binom{d}{2}}$,

$$-C|p|^2 3^{-\frac{n}{2}} \leq \nu\left(\square_n, p\right) + \nu^*\left(\square_n, \bar{\mathbf{a}}_n p\right) - \bar{\mathbf{a}}_n |p|^2 \leq C|p|^2 3^{-\alpha n},$$

  which can be rewritten, by using (6.102),

$$\left|\mathbf{a}\left(\square_n^-\right) - \bar{\mathbf{a}}_n\right| \leq C 3^{-\alpha n}; \tag{6.105}$$

- By Lemma 6.19 and Corollary 6.25, we have the inequality, for each pair of integers $(m, n) \in \mathbb{N}$ such that $m \leq n$,

$$\left|\bar{\mathbf{a}}_n^{-1} - \bar{\mathbf{a}}_m^{-1}\right|^2 \leq \sum_{k=m}^{n} \tau_k + C 3^{-\frac{m}{2}} \leq \sum_{k=m}^{n} C 3^{-\alpha k} + C 3^{-\frac{m}{2}} \leq C 3^{-\alpha m}. \tag{6.106}$$

We now combine the four previous results to complete the proof of Proposition 6.10. First by sending n to infinity in the inequality (6.105), and using the convergence (6.104), we obtain the identity $\bar{\mathbf{a}} = \bar{\mathbf{a}}_*^{-1}$. Then by sending n to infinity in the inequality (6.106), we obtain the inequality, for each integer $m \in \mathbb{N}$,

$$\left|\bar{\mathbf{a}}_m^{-1} - \bar{\mathbf{a}}^{-1}\right| \leq C 3^{-\alpha m}. \tag{6.107}$$

We then combine the inequality (6.103) with the inequality (6.107) to obtain

$$\left|\bar{\mathbf{a}}_m - \bar{\mathbf{a}}\right| \leq C 3^{-\alpha m}. \tag{6.108}$$

Combining the estimates (6.105) and (6.108), we deduce that, for each integer $n \in \mathbb{N}$,

$$\left|\bar{\mathbf{a}}\left(\square_n^-\right) - \bar{\mathbf{a}}\right| \leq \left|\bar{\mathbf{a}}\left(\square_n^-\right) - \bar{\mathbf{a}}_n\right| + \left|\bar{\mathbf{a}}_n - \bar{\mathbf{a}}\right| \leq C 3^{-\alpha n}. \tag{6.109}$$

Proposition 6.12 is then a consequence of the estimates (6.108), (6.109), and the representation formulae (6.102). □

6.4. **Definition of the first-order corrector and quantitative sublinearity.** An important ingredient to prove the quantitative homogenization of the mixed derivative of the Green's matrix associated with the Helffer-Sjöstrand operator (which is the subject of Section 7) is the first-order corrector. The objective of this section is to introduce this function, and to deduce from the algebraic rate of convergence on the energy ν established in Proposition 6.10 two properties on this map:

- The quantitative sublinearity of the corrector, this result is stated in the equation (6.110);
- A quantitative estimate on the H^{-1}-norm of the flux of the corrector, this result is stated in the estimate (6.111).

The corrector which is introduced in this section is a finite-volume version of the corrector (see Definition 6.26), the reason justifying this choice is that it is simpler to construct from the subadditive energy ν and allows the arguments developed in Section 7 to work. We do not try to construct the infinite-volume corrector as it would require to prove a quantitative homogenization theorem and establish a large-scale regularity theory (following the techniques of [8, Section 3]), and the development of this technology is not necessary to prove Theorem 1. Nevertheless, the specific structure of the problem (and the strong regularity properties established in Section 5) allows to define the gradient of the infinite-volume corrector with a simple argument; the construction is carried out in Proposition 6.29.

6.4.1. *Finite-volume corrector.* This section is devoted to the definition and the study of the finite-volume corrector.

Definition 6.26 (Finite-volume corrector)**.** For each integer $n \in \mathbb{N}$, and each slope $p \in \mathbb{R}^{d\times\binom{d}{2}}$, we define the finite-volume corrector at scale 3^n to be the function $\chi_{n,p} : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ defined by the formula

$$\chi_{n,p} := u\left(\cdot,\cdot,\square_n^-,p\right) - l_p.$$

We recall that the corrector extended by 0 outside the trimmed cube $\square_n^-$. Given two integers $(i,j) \in \{1,\ldots,d\}\times\{1,\ldots,\binom{d}{2}\}$, we denote by $e_{ij} \in \mathbb{R}^{d\times\binom{d}{2}}$ the vector $e_{ij} = (0,\ldots,e_i,\ldots,0)$, and denote by $\chi_{n,ij} := \chi_{n,e_{ij}}$.

Remark 6.27. The finite volume corrector $\chi_{n,p}$ is the solution of the equation

$$\begin{cases} -\Delta_\phi \chi_{n,p} - \frac{1}{2\beta}\Delta\chi_{n,p} + \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}\left(-\Delta\right)^{n+1}\chi_{n,p} + \sum_{q\in\mathcal{Q}}\nabla_q^*\cdot\mathbf{a}_q\nabla_q\left(l_p+\chi_{n,p}\right) = 0 \text{ in } \square_n^- \times\Omega, \\ \chi_{n,p} = 0 \text{ on } \partial\square_n^- \times\Omega. \end{cases}$$

By the identity $\nabla_q\left(l_p+\chi_{n,p}\right) = \left(n_q, \mathrm{d}^* l_p + \mathrm{d}^*\chi_{n,p}\right)$, we see that the corrector depends only on the value of $\mathrm{d}^* l_p$. In particular, if $\mathrm{d}^* l_p = 0$ then $\chi_{n,p} = 0$. As the vectors $\mathrm{d}^* l_p$ belong to the space $\mathbb{R}^d$, the collection of correctors $\left(\chi_p\right)_{p\in\mathbb{R}^{d\times\binom{d}{2}}}$ forms a d-dimensional vector space from which we extract a basis: for each integer $i\in\{1,\ldots,d\}$, we select a vector $p_i \in \mathbb{R}^{d\times\binom{d}{2}}$ such that $\mathrm{d}^* l_{p_i} = e_i$ and denote by $\nabla\chi_i = \nabla\chi_{p_i}$.

The following proposition establishes quantitative sublinearity of the corrector and provides a quantitative estimate for the H^{-1}-norm of its flux.

Proposition 6.28 (Quantitative sublinearity)**.** *There exist a constant $C := C(d)$, an exponent $\alpha(d) > 0$, and an inverse temperature $\beta_0(d) < \infty$ such that, for every inverse temperature $\beta > \beta_0$, and every vector $p \in \mathbb{R}^{d\times\binom{d}{2}}$, the finite-volume corrector satisfies the following estimates*

$$\left\|\chi_{n,p}\right\|_{\underline{L}^2\left(\square_n^-,\mu_\beta\right)} \leq C|p|3^{(1-\alpha)n} \tag{6.110}$$

and

$$\left\|\frac{1}{2}\left(p+\nabla\chi_{n,p}\right) + \beta\sum_{q\in\mathcal{Q}}\mathbf{a}_q\nabla_q\left(l_p+\chi_{n,p}\right)L^t_{2,\mathrm{d}^*}\left(n_q\right) - \bar{\mathbf{a}}p\right\|_{\underline{H}^{-1}\left(\square_n^-,\mu_\beta\right)} \leq C|p|3^{(1-\alpha)n}. \tag{6.111}$$

Proof. The estimate (6.110) is obtained by combining Lemma 6.22 and Corollary 6.25. The proof of the estimate (6.111) regarding the flux is more involved and we split it into two steps. The argument requires to take into account the infinite range of the sum over the charges (by using the boundary layer BL_n and the exponential decay of the coefficient $\mathbf{a}_q$), which makes the proof technical. Since similar technicalities have already been treated in the previous sections, and the analysis does not contain any new arguments, we omit some of the details and only write a (detailed) sketch of the proof.

Step 1. In this step, we prove that, to prove (6.111) it is sufficient to prove the estimate, for each $p^* \in \mathbb{R}^{d \times \binom{d}{2}}$,

$$\left\| \frac{1}{2} \nabla v \left(\cdot, \cdot, \square_n, p^* \right) + \beta \sum_{\operatorname{supp} q \subseteq \square_n} \mathbf{a}_q \nabla_q v \left(\cdot, \cdot, \square_n, p^* \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) - p^* \right\|_{\underline{H}^{-1}(\square_n^-, \mu_\beta)} \leq C |p^*| 3^{(1-\alpha)n}. \tag{6.112}$$

We fix a vector $p^* \in \mathbb{R}^{d \times \binom{2}{d}}$ and recall that, by definition of the first order corrector, $l_p + \chi_{n,p} = u \left(\cdot, \cdot, \square_n^-, p \right)$. To ease the notation, we denote by $u := u \left(\cdot, \cdot, \square_n^-, p \right)$ and by $v := v \left(\cdot, \cdot, \square_n, \bar{\mathbf{a}} p \right)$. First, we note that Proposition 6.12 implies the inequality $|\mathbf{a} \left(\square_m^- \right) - \bar{\mathbf{a}}| \leq C 3^{-\alpha m}$. Combining this result with the estimate (6.7), we obtain the inequality, for each vector $p \in \mathbb{R}^{d \times \binom{d}{2}}$,

$$\left\| \nabla v \left(\cdot, \cdot, \square_n, \bar{\mathbf{a}} p \right) - \nabla v \left(\cdot, \cdot, \square_n, \bar{\mathbf{a}}_n p \right) \right\|_{\underline{L}^2(\square_n, \mu_\beta)} = \left\| \nabla v \left(\cdot, \cdot, \square_n, \bar{\mathbf{a}} p - \bar{\mathbf{a}}_n p \right) \right\|_{\underline{L}^2(\square_n, \mu_\beta)} \leq C 3^{-\alpha n} |p|. \tag{6.113}$$

We use the inequality (6.113) with the estimate (6.85) stated in the proof of Proposition 6.23 and Corollary 6.25. We deduce that

$$\begin{aligned} \left\| \nabla u - \nabla v \right\|_{\underline{L}^2(\square_n, \mu_\beta)} &\leq \left\| \nabla u - \nabla v \left(\cdot, \cdot, \square_n, \bar{\mathbf{a}}_n p \right) \right\|_{\underline{L}^2(\square_n, \mu_\beta)} + \left\| \nabla v \left(\cdot, \cdot, \square_n, \bar{\mathbf{a}}_n p \right) - \nabla v \right\|_{\underline{L}^2(\square_n, \mu_\beta)} \\ &\leq C 3^{-\alpha n} |p|. \end{aligned} \tag{6.114}$$

Using the estimate (6.114), we can write

$$\begin{aligned} &\left\| \frac{1}{2} \nabla u + \beta \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(\nabla_q u \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) - \bar{\mathbf{a}} p \right\|_{\underline{H}^{-1}(\square_n^-, \mu_\beta)} \\ &\quad \leq \left\| \frac{1}{2} \nabla v + \beta \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(\nabla_q v \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) - \bar{\mathbf{a}} p \right\|_{\underline{H}^{-1}(\square_n^-, \mu_\beta)} \\ &\qquad + \left\| \frac{1}{2} \nabla \left(u - v \right) + \beta \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(\nabla_q \left(u - v \right) \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) \right\|_{\underline{H}^{-1}(\square_n, \mu_\beta)} \\ &\quad \leq \left\| \frac{1}{2} \nabla v + \beta \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(\nabla_q v \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) - \bar{\mathbf{a}} p \right\|_{\underline{H}^{-1}(\square_n, \mu_\beta)} \\ &\qquad + C 3^n \left\| \frac{1}{2} \nabla \left(u - v \right) + \beta \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(\nabla_q \left(u - v \right) \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) - \bar{\mathbf{a}} p \right\|_{\underline{L}^2(\square_n, \mu_\beta)}. \end{aligned}$$

Using the estimate $|\mathbf{a}_q| \leq e^{-c\sqrt{\beta}\|q\|_1}$, we see that

$$\left\| \frac{1}{2} \nabla \left(u - v \right) + \beta \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(\nabla_q \left(u - v \right) \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) \right\|_{\underline{L}^2(\square_n, \mu_\beta)} \leq C \left\| \nabla \left(u - v \right) \right\|_{L^2(\square_n, \mu_\beta)} \leq C 3^{-\alpha n} |p|.$$

A combination of the two previous displays shows

$$\left\| \frac{1}{2} \nabla u + \beta \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(\nabla_q u \right) n_q - \bar{\mathbf{a}} p \right\|_{\underline{H}^{-1}(\square_n, \mu_\beta)} \leq \left\| \frac{1}{2} \nabla v + \beta \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(\nabla_q v \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) - \bar{\mathbf{a}} p \right\|_{\underline{H}^{-1}(\square_n, \mu_\beta)} + C 3^{(1-\alpha)n} |p|. \tag{6.115}$$

The estimate (6.115) implies that to prove the inequality (6.111), it is sufficient to prove (6.112).

Step 2. Proving the estimate (6.112). The argument is similar to the proof presented in Lemma 6.22. To ease the notation, we denote by $v := v \left(\cdot, \cdot, \square_n, p^* \right)$ and by $v_{z,m} := v \left(\cdot, \cdot, z + \square_m, p^* \right)$, and assume without loss of generality that $|p^*| = 1$. We use the H^{-1}-version of the multiscale Poincaré inequality stated in Proposition B.1

of Appendix B. We obtain

$$\left\| \frac{1}{2}\nabla v + \beta \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q v L^t_{2,\mathrm{d}^*}(n_q) - p^* \right\|^2_{\underline{H}^{-1}(\square_n^-,\mu_\beta)} \leq C \left\| \frac{1}{2}\nabla v + \beta \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q v L^t_{2,\mathrm{d}^*}(n_q) - p^* \right\|^2_{\underline{L}^2(\square_n^-,\mu_\beta)} + C3^n \sum_{m=0}^{n} \sum_{z\in\mathcal{Z}_{m,n}} \frac{3^m}{|\mathcal{Z}_{m,n}|} \left\langle \left(\frac{1}{|z+\square_m|} \sum_{x\in z+\square_m} \frac{1}{2}\nabla v(x,\cdot) + \beta \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q v L^t_{2,\mathrm{d}^*}(n_q(x)) - p^* \right)^2 \right\rangle_{\mu_\beta}. \tag{6.116}$$

The first term in the right side of (6.116) can be estimated by the estimate (6.7). We obtain

$$\left\| \frac{1}{2}\nabla v + \beta \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q v L^t_{2,\mathrm{d}^*}(n_q) - p^* \right\|_{\underline{L}^2(\square_n^-,\mu_\beta)} \leq C. \tag{6.117}$$

To estimate the second term in the right side of (6.116), we proceed as in Lemma 6.22, and use the subadditivity estimate stated in Proposition 6.17 and Corollary 6.18. We obtain

$$\sum_{z\in\mathcal{Z}_{m,n}} \frac{1}{|\mathcal{Z}_{m,n}|} \left\langle \left(\frac{1}{|z+\square_m|} \sum_{x\in z+\square_m} \frac{1}{2}\nabla v(x,\cdot) + \beta \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q v L^t_{2,\mathrm{d}^*}(n_q(x)) - p^* \right)^2 \right\rangle_{\mu_\beta} \leq \sum_{z\in\mathcal{Z}_{m,n}} \frac{1}{|\mathcal{Z}_{m,n}|} \left\langle \left(\frac{1}{|z+\square_m|} \sum_{x\in z+\square_m} \frac{1}{2}\nabla v_{z,m}(x,\cdot) + \beta \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q v_{z,m} L^t_{2,\mathrm{d}^*}(n_q(x)) - p^* \right)^2 \right\rangle_{\mu_\beta} + C3^{-\alpha m}. \tag{6.118}$$

We then use the two following results:

- The identity, for each point $z \in \mathcal{Z}_{m,n}$,
$$\left\langle \frac{1}{|z+\square_m|} \sum_{x\in z+\square_m} \left(\frac{1}{2}\nabla v_{z,m}(x,\cdot) + \beta \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q v_{z,m} L^t_{2,\mathrm{d}^*}(n_q(x)) \right) \right\rangle_{\mu_\beta} = p^*;$$
- The variance estimate
$$\mathrm{var}\left[\frac{1}{|z+\square_m|} \sum_{x\in z+\square_m} \left(\frac{1}{2}\nabla v_{z,m}(x,\cdot) + \beta \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q v_{z,m} L^t_{2,\mathrm{d}^*}(n_q(x)) \right) \right] \leq C3^{-\frac{m}{2}},$$
which is a consequence of Lemma 6.20, the inequality $d - \frac{5}{2} \geq \frac{1}{2}$ valid in dimension larger than 3, and the translation invariance of the measure μ_β.

We obtain the estimate

$$\left\langle \left(\frac{1}{|z+\square_m|} \sum_{x\in z+\square_m} \frac{1}{2}\nabla v_{z,m}(x,\cdot) + \beta \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q v_{z,m} L^t_{2,\mathrm{d}^*}(n_q(x)) - p^* \right)^2 \right\rangle_{\mu_\beta} \leq C3^{-\frac{m}{2}}. \tag{6.119}$$

Combining the estimates (6.116), (6.117), (6.118), and (6.119), we have obtained

$$\left\| \frac{1}{2}\nabla v + \beta \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q v L^t_{2,\mathrm{d}^*}(n_q) - p^* \right\|_{\underline{H}^{-1}(\square_n^-,\mu_\beta)} \leq C3^{(1-\alpha)n}.$$

The proof of Proposition 6.28 is complete. □

6.4.2. *Gradient of the infinite-volume corrector.* The next proposition establishes the existence and stationarity of the spatial gradient of the infinite-volume corrector.

Proposition 6.29 (Existence of the gradient of the infinite-volume corrector and stationarity)**.** *There exists a stationary random field $\nabla\chi : \mathbb{Z}^d \times \Omega \to \mathbb{R}$ satisfying the following property, for each $p \in \mathbb{R}^d$, and each integer $n \in \mathbb{N}$,*

$$\|\nabla\chi_{n,p} - \nabla\chi_p\|_{\underline{L}^2(\square_m,\mu_\beta)} \leq C3^{-n\alpha}.$$

Remark 6.30. The property stated in Remark 6.27 about the finite volume corrector also applies to the infinite volume corrector.

Let us first present the main idea of the argument. By assuming that the inverse temperature is large enough, one has $C^{0,1-\varepsilon}$-regularity estimates for the solutions of the Helffer-Sjöstrand equation, following the arguments given in Section 5.2. By Proposition 6.10, one also has an algebraic rate of convergence for the subadditive energy ν with exponent α. The exponent ε depends on the inverse temperature β and tends to 0 as β tends to infinity, while the exponent α depends only on the dimension, and remains bounded away from zero when the inverse temperature tends to infinity. It is thus possible to choose β sufficiently large so that the exponent ε is smaller than the exponent $\alpha/2$, and to leverage on this property, the $C^{0,1-\varepsilon}$-regularity estimate presented in Proposition 5.4, and the Caccioppoli inequality to prove the existence of the gradient of the infinite-volume corrector.

Proof. We fix a vector $p \in \mathbb{R}^{d\times\binom{d}{2}}$ and assume without loss of generality that $|p| = 1$. We decompose the proof into two steps. In the first step, we prove that, for each point $x \in \mathbb{Z}^d$, the sequence $(\nabla\chi_{n,p}(x,\cdot))_{n\in\mathbb{N}}$ is Cauchy in the space $L^2(\mu_\beta)$. This implies that it converges, and we define the gradient of the infinite-volume corrector to be its limit. In the second step we prove that the function $\nabla\chi_p$ is stationary.

Step 1. We prove the inequality, for each point $x \in \mathbb{Z}^d$ integer $n \in \mathbb{N}$ such that $x \in \square_n^-$,

$$\|\nabla\chi_{n,p}(x,\cdot) - \nabla\chi_{n+1,p}(x,\cdot)\|_{\underline{L}^2(\mu_\beta)} \leq C3^{-\frac{\alpha}{2}n}. \tag{6.120}$$

We now fix a point $x \in \mathbb{Z}^d$ and prove the estimate (6.120). By the definition of the correctors stated in Definition 6.26, the functions χ_n and χ_{n+1} are solutions of the Helffer-Sjöstrand equations

$$\mathcal{L}(l_p + \chi_{n,p}) = 0 \text{ in } \square_n^- \times\Omega \text{ and } \mathcal{L}(l_p + \chi_{n+1,p}) = 0 \text{ in } \square_{n+1}^- \times\Omega.$$

In particular, the difference $\chi_{n+1,p} - \chi_{n,p}$ is solution of the equation $\mathcal{L}(\chi_{n+1,p} - \chi_{n,p}) = 0$ in the set $\square_n^- \times \Omega$. We can thus apply Proposition 5.6 to obtain, for each integer $n \in \mathbb{N}$ such that $x \in \square_{n-1}$,

$$\begin{aligned} \|\nabla\chi_{n,p}(x,\cdot) - \nabla\chi_{n+1,p}(x,\cdot)\|_{L(\mu_\beta)} &\leq \sup_{y\in\square_{n-1}} \|\nabla\chi_{n,p}(y,\cdot) - \nabla\chi_{n+1,p}(y,\cdot)\|_{L^2(\mu_\beta)} \\ &\leq C3^{(\varepsilon-1)n} \left\|\chi_{n,p} - \chi_{n+1,p} - (\chi_{n,p} - \chi_{n+1,p})_{\square_n^-}\right\|_{\underline{L}^2(\square_n^-,\mu_\beta)} \\ &\leq C3^{(\varepsilon-1)n} \|\chi_{n,p} - \chi_{n+1,p}\|_{\underline{L}^2(\square_n^-,\mu_\beta)}. \end{aligned} \tag{6.121}$$

By combining the estimate (6.121) and Proposition 6.28, we obtain the estimate, for each pair of integers $n \in \mathbb{N}$ such that $x \in \square_{n-1}$,

$$\|\nabla\chi_{n,p}(x,\cdot) - \nabla\chi_{n+1,p}(x,\cdot)\|_{L^2(\mu_\beta)} \leq C3^{(\varepsilon-\alpha)n}.$$

Using the assumption $\varepsilon \leq \frac{\alpha}{2}$, we obtain

$$\|\nabla\chi_{n,p}(x,\cdot) - \nabla\chi_{n+1,p}(x,\cdot)\|_{L^2(\mu_\beta)} \leq C3^{-\frac{\alpha}{2}n}. \tag{6.122}$$

The inequality (6.122) implies that, the sequence $(\nabla\chi_{n,p}(x,\cdot))_{n\in\mathbb{N}}$ is Cauchy in the space $L^2(\mu_\beta)$. This implies that it converges in the space $L^2(\mu_\beta)$. We define the gradient of the corrector $\nabla\chi_p(x,\cdot)$ to be the limiting object.

From the estimate (6.122), we also deduce that, for each pair of integers $n \in \mathbb{N}$,

$$\|\nabla\chi_{n,p}(x,\cdot) - \nabla\chi_p(x,\cdot)\|_{L^2(\mu_\beta)} \leq C3^{-\frac{\alpha}{2}n}.$$

The proof of Step 1 is complete.

Step 2. In this step, we prove the stationarity of the infinite-volume gradient corrector. For $z \in \mathbb{Z}^d$, we will make use of the notation τ_z for the translation of the field introduced in Section 2. We prove the identity, for each $(x,\phi) \in \mathbb{Z}^d \times \Omega$,

$$\nabla\chi_p(x,\phi) = \nabla\chi_p(z+x,\tau_z\phi). \tag{6.123}$$

To prove the equality (6.123), we first note that, by the definition of the function u, we have the equality, for each point $z \in \mathbb{Z}^d$, each cube $\square \subseteq \mathbb{Z}^d$, and each pair $(x,\phi) \in (y+\square)\times\Omega$,

$$u(x,\phi,y+\square,p) = u(x-y,\tau_{-y}\phi,\square,p). \tag{6.124}$$

Using the identity (6.124), the result established in Step 1, and the translation invariance of the measure μ_β, we obtain that the sequence $(\nabla u(x,\cdot,y+\square_n,p) - p)_{n\in\mathbb{N}}$ converges in the space $L^2(\mu_\beta)$ to the random variable $\phi \to \nabla \chi_p (x-y, \tau_{-y}\phi)$. Thus to prove the identity (6.123), it is sufficient to prove that the sequence $(\nabla u(x,\cdot,y+\square_n,p)-p)_{n\in\mathbb{N}}$ also converges in $L^2(\mu_\beta)$ to the gradient of the corrector $\phi \to \nabla\chi_p(x,\phi)$. This is what we now prove.

We first note that the proof of Proposition 6.12 can be adapted so as to have the following result. For each $y \in \mathbb{Z}^d$, and each integer n such that $3^{\frac{n}{2}} \geq 2|y|$, one has the estimate

$$\sum_{z\in\mathcal{Z}_n} \left\|\nabla u(\cdot,\cdot,y+z+\square_n,p) - \nabla u(\cdot,\cdot,\square_{n+1},p)\right\|^2_{\underline{L}^2(y+z+\square_n,\mu_\beta)} \leq C\left(\nu(\square_n,p)-\nu(\square_{n+1},p)+3^{-\frac{n}{2}}\right). \tag{6.125}$$

The proof is identical; indeed under the assumption $3^{\frac{n}{2}} \geq 2|y|$, one can partition the triadic cube $(y+\square_{n+1})$ into the collection of triadic cubes $(y+z+\square_n)_{z\in\mathcal{Z}_n}$ and a boundary layer of width of size $3^{\frac{n}{2}}$. One can then rewrite the proof of Proposition 6.12 to obtain the estimate (6.125). We then use Proposition 6.12 (or more precisely Corollary 6.25), and obtain the inequality

$$\left\|\nabla u(\cdot,\cdot,y+\square_n,p) - \nabla u(\cdot,\cdot,\square_{n+1},p)\right\|^2_{\underline{L}^2(y+\square_n,\mu_\beta)} \leq C3^{-\alpha n}.$$

Using the $C^{1-\varepsilon}$-regularity estimate stated in Proposition 5.6, the assumption $\varepsilon \leq \frac{\alpha}{2}$, and an argument similar to the one presented in Step 1, we obtain, for each integer $n\in\mathbb{N}$ such that $3^{\frac{n}{2}} \geq 2|y|$, and each point $x \in \square_{n-1}$,

$$\left\|\nabla u(x,\cdot,y+\square_n,p) - \nabla u(x,\cdot,\square_{n+1},p)\right\|^2_{L^2(\mu_\beta)} \leq C3^{-\frac{\alpha}{2}n}.$$

Using the definition of the finite-volume corrector given in Definition 6.26 and the inequality (6.120), we deduce that

$$\left\|\nabla u(x,\cdot,y+\square_n,p) - p - \nabla\chi_p(x,\cdot)\right\|^2_{L^2(\mu_\beta)} \leq C3^{-\frac{\alpha}{2}n}.$$

The previous inequality implies that the sequence $(\nabla u(x,\cdot,y+\square_n,p)-p)_{n\in\mathbb{N}}$ converges in the space $L^2(\mu_\beta)$ to the random variable $\phi \to \nabla\chi_p(x,\phi)$. The proof of Proposition 6.29 is complete. □

7. Quantitative homogenization of the Green's matrix

7.1. **Statement of the main result.** The objective of this section is to prove the homogenization of the mixed gradient of the Green's matrix stated in Theorem 2. We first introduce the notation $\bar{\mathbf{a}}_\beta := \bar{\mathbf{a}}/\beta$ and the Green's matrix associated with the homogenized operator $\nabla \cdot \bar{\mathbf{a}}_\beta \nabla$: we denote by $\overline{G} : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2} \times \binom{d}{2}}$ the fundamental solution of the elliptic system

$$-\nabla \cdot \bar{\mathbf{a}}_\beta \nabla \overline{G} = \delta_0 \text{ in } \mathbb{Z}^d. \tag{7.1}$$

The matrix $\bar{\mathbf{a}}_\beta$ is a small perturbation of the matrix $\frac{1}{2\beta} I_d$ and the size of the perturbation is of order $\beta^{-\frac{3}{2}} \ll \beta^{-1}$. The solvability of the equation is thus ensured by the arguments of Section 5; more specifically, a Nash-Aronson estimate holds for the heat-kernel associated with the operator $-\nabla \cdot \bar{\mathbf{a}}_\beta \nabla$ which can then be integrated over time. We rewrite the statement of Theorem 2 below

Theorem 2 (Homogenization of the mixed derivative of the Green's matrix)**.** *Fix a charge $q_1 \in \mathcal{Q}$ such that 0 belongs to the support of n_{q_1} and let $\mathcal{U}_{q_1}$ be the solution of the Helffer-Sjöstrand equation*

$$\mathcal{L}\mathcal{U}_{q_1} = \cos\left(2\pi\left(\phi, q_1\right)\right) q_1 \ \textit{in } \mathbb{Z}^d \times \Omega. \tag{7.2}$$

For each integer $k \in \left\{1, \ldots, \binom{d}{2}\right\}$, we define the function $\overline{G}_{q_1,k} : \mathbb{Z}^d \to \mathbb{R}$ by the formula

$$\overline{G}_{q_1,k} = \sum_{1 \leq i \leq d} \sum_{1 \leq j \leq \binom{d}{2}} \left\langle \cos\left(2\pi\left(\phi, q_1\right)\right) \left(n_{q_1}, \mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{ij}\right)\right\rangle_{\mu_\beta} \nabla_i \overline{G}_{jk}.$$

Then, there exist an inverse temperature $\beta_0 := \beta_0(d) < \infty$, an exponent $\gamma := \gamma(d) > 0$, and a constant C_{q_1} which satisfies the estimate $|C_{q_1}| \leq C \left\|q_1\right\|_1^k$ for some constant $C := C(d, \beta) < \infty$ and exponent $k := k(d) < \infty$, such that for each $\beta \geq \beta_0$ and each radius $R \geq 1$, one has the inequality

$$\left\| \nabla \mathcal{U}_{q_1} - \sum_{1 \leq i \leq d} \sum_{1 \leq j \leq \binom{d}{2}} \left(e_{ij} + \nabla \chi_{ij}\right) \nabla_i \overline{G}_{q_1,j} \right\|_{\underline{L}^2(A_R, \mu_\beta)} \leq \frac{C_{q_1}}{R^{d+\gamma}}. \tag{7.3}$$

Remark 7.1. Since the codifferential d^* is a linear functional of the gradient, the map $\mathrm{d}^* \chi_{ij}$ is well-defined even if we have only defined the gradient of the infinite-volume corrector: we have the identity $\mathrm{d}^* \chi_{ij} = L_{2,\mathrm{d}^*}\left(\nabla^* \chi_{ij}\right)$.

Remark 7.2. We recall that in this section, the constants are allowed to depend on the dimension d and on the inverse temperature β.

Remark 7.3. We recall the definition of the annulus $A_R := B_{2R} \setminus B_R$; its volume is of order R^d.

Remark 7.4. The double sum $\sum_{1 \leq i \leq d} \sum_{1 \leq j \leq \binom{d}{2}}$ appears frequently in the proofs of this section; to ease the notation, we denote it by $\sum_{i,j}$.

Remark 7.5. Since the form q_1 can be written $\mathrm{d} n_{q_1}$, we expect the two gradients $\nabla \mathcal{U}_{q_1}$ and $\nabla \overline{G}_{q_1}$ to behave like the mixed derivative of the Green's function, i.e., they should be of order R^{-d} in the annulus A_R. The proposition asserts that the difference between the two terms $\nabla \mathcal{U}_{q_1}$ and $\sum_{i,j} \left(e_{ij} + \nabla \chi_{ij}\right) \nabla_i \overline{G}_{q_1,j}$ is quantitatively smaller than the typical size of the two terms considered separately.

7.2. **Outline of the argument.** The strategy of the proof of Theorem 2 relies on a classical strategy in homogenization: the two-scale-expansion. The proofs presented in the Section make essentially use of two ingredients established in Sections 5 and 6:

- The quantitative sublinearity of the finite-volume corrector and the estimate on the H^{-1}-norm of the flux stated in Proposition 6.28;
- The $C^{0,1-\varepsilon}$-regularity theory established in Section 5.

We now give an outline of the proof of Theorem 2. The argument is split into two sections:

- In Section 7.3, we perform the two-scale expansion and obtain a result of homogenization for the gradient of the Green's matrix as stated in Proposition 7.6;
- In Section 7.4, we use the result of Proposition 7.6 and perform the two-scale expansion a second time to obtain the quantitative homogenization of the mixed derivative of the Green's matrix stated in Theorem 2.

7.2.1. *Homogenization of the gradient of the Green's matrix.* In this subsection, we present an outline of the proof of Section 7.3; the objective is to establish the quantitative homogenization of the gradient of the Green's matrix stated in Proposition 7.6 below.

Proposition 7.6 (Homogenization of the Green's matrix). *Let $\mathcal{G} : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2} \times \binom{d}{2}}$ be the Green's matrix associated with the Helffer-Sjöstrand equation*

$$\mathcal{L}\mathcal{G} = \delta_0 \quad in \quad \mathbb{Z}^d \times \Omega. \tag{7.4}$$

Then, there exist an inverse temperature $\beta_0(d) < \infty$, an exponent $\gamma := \gamma(d) > 0$, and a constant $C := C(d) < \infty$ such that, for any $\beta > \beta_0$, any radius $R \geq 1$, and any integer $k \in \left\{1, \ldots, \binom{d}{2}\right\}$,

$$\left\| \nabla \mathcal{G}_{\cdot k} - \sum_{i,j} \left(e_{ij} + \nabla \chi_{ij} \right) \nabla_i \overline{G}_{jk} \right\|_{\underline{L}^2(A_R, \mu_\beta)} \leq \frac{C}{R^{d-1+\gamma}}. \tag{7.5}$$

To set up the argument, we first select an inverse temperature β large enough, depending only on the dimension d, such that the quantitative sublinearity of the finite-volume corrector and of its flux stated in Proposition 6.28 holds with exponent $\alpha > 0$. Following the argument explained at the beginning of Section 6.4, we can choose the parameter β large enough so that all the results presented in Section 5 pertaining to the $C^{0,1-\varepsilon}$-regularity theory for the Helffer-Sjöstrand operator $\mathcal{L}$ are valid with a regularity exponent ε which is small compared to the exponent α (we assume for instance that the ratio between ε and α is smaller than $100d^2$). We also fix an exponent δ which is both larger than ε and smaller than α and corresponds to the size of a mollifier exponent which needs to be taken into account in the argument (we assume for instance that the ratios between the exponents α and δ and between the exponents ε and δ are both smaller than $10d$). We have thus three exponents in the argument; they can be ordered by the following relations

$$0 < \underbrace{\varepsilon}_{\text{regularity}} \ll \underbrace{\delta}_{\text{mollifier exponent}} \ll \underbrace{\alpha}_{\text{homogenization}} \ll 1. \tag{7.6}$$

We additionally assume that the exponents ε, δ and α are chosen in a way that they depend only on the dimension d. The exponent γ in the statement of Proposition 7.7 depends only ε, δ and α (and thus only on the dimension d).

We now give an outline of the proof of the inequality (7.5). The first step of the argument is to approximate the Green's matrices $\mathcal{G}$ and $\overline{G}$; the main issue is that the spatial Dirac function δ_0 in the definitions of the Green's matrices $\mathcal{G}$ in (7.4) and $\overline{G}$ in (7.1) is too singular and causes some problems in the analysis. To remedy this, we replace the Dirac function δ_0 by a smoother function, and make use of the mollifier exponent δ: we let ρ_δ be a discrete function from $\mathbb{Z}^d$ to $\mathbb{R}^{\binom{d}{2} \times \binom{d}{2}}$, we denote its components by $(\rho_{\delta,ij})_{1 \leq i,j \leq \binom{d}{2}}$, and assume that they satisfy the four properties

$$\operatorname{supp} \rho_\delta \subseteq B_{R^{1-\delta}}, \ 0 \leq \rho_{\delta,ij} \leq C R^{-(1-\delta)d}, \ \sum_{x \in \mathbb{Z}^d} \rho_{\delta,ij}(x) = \mathbb{1}_{\{i=j\}}, \text{ and } \forall k \in \mathbb{N}, \ \left| \nabla^k \rho_{\delta,ij} \right| \leq \frac{C}{R^{(d+k)(1-\delta)}}, \tag{7.7}$$

which implies in particular that $\rho_{\delta,ij} = 0$ if $i \neq j$. We define the functions $\mathcal{G}_\delta : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2} \times \binom{d}{2}}$ and $\overline{G}_\delta : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2} \times \binom{d}{2}}$ to be the solution of the systems, for each integer $k \in \left\{1, \ldots, \binom{d}{2}\right\}$,

$$\mathcal{L}\mathcal{G}_{\delta,\cdot k} = \rho_{\delta,\cdot k} \text{ in } \Omega \times \mathbb{Z}^d, \qquad -\nabla \cdot \left(\bar{\mathbf{a}}_\beta \nabla \overline{G}_{\delta,\cdot k} \right) = \rho_{\delta,\cdot k} \text{ in } \mathbb{Z}^d. \tag{7.8}$$

We then prove, by using the regularity theory established in Section 5, that the functions $\mathcal{G}_\delta$, $\overline{G}_\delta$ are good approximations of the functions $\mathcal{G}$, $\overline{G}$. This is the subject of Lemma 7.7 where we show that there exists an exponent $\gamma := \gamma(d, \beta, \delta, \varepsilon) > 0$ such that

$$\left\| \nabla \mathcal{G}_\delta - \nabla \mathcal{G} \right\|_{L^\infty(A_R, \mu_\beta)} \leq C R^{1-d-\gamma} \quad \text{and} \quad \left\| \nabla \overline{G}_\delta - \nabla \overline{G} \right\|_{L^\infty(A_R, \mu_\beta)} \leq C R^{1-d-\gamma}. \tag{7.9}$$

By the estimates (7.9), we see that to prove Proposition 7.7 it is sufficient to prove the inequality, for each integer $k \in \{1, \ldots, \binom{d}{2}\}$,

$$\left\| \nabla \mathcal{G}_{\delta,\cdot k} - \sum_{i,j} \left(e_{ij} + \nabla \chi_{ij} \right) \nabla_i \overline{G}_{\delta,jk} \right\|_{\underline{L}^2(A_R, \mu_\beta)} \leq C R^{1-d-\gamma}. \tag{7.10}$$

We now sketch the proof of the inequality (7.10). We let m be the integer uniquely defined by the inequalities $3^m \leq R^{1+\delta} < 3^{m+1}$, and consider the collection of finite-volume correctors $(\chi_{m,ij})_{1 \leq i \leq d, 1 \leq j \leq \binom{d}{2}}$. We then define the two-scale expansion $\mathcal{H}_\delta : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2} \times \binom{d}{2}}$ according to the formula, for each $k \in \left\{1, \ldots, \binom{d}{2}\right\}$,

$$\mathcal{H}_{\delta,\cdot k} := \overline{G}_{\delta,\cdot k} + \sum_{i,j} \left(\nabla_i \overline{G}_{\delta,jk} \right) \chi_{m,ij}. \tag{7.11}$$

We now fix an integer $k \in \left\{1, \ldots, \binom{d}{2}\right\}$. The strategy is to compute the value of $\mathcal{L}\mathcal{H}_{\delta,\cdot k}$ by using the explicit formula on the map $\mathcal{H}_{\delta,\cdot k}$ stated in (7.11), and to prove that it is quantitatively close to the map $\rho_{\delta,\cdot k}$ in the correct functional space; precisely, we prove the H^{-1}-estimate,

$$\left\| \mathcal{L}\mathcal{H}_{\delta,\cdot k} - \rho_{\delta,\cdot k} \right\|_{\underline{H}^{-1}\left(B_{R^{1+\delta}}, \mu_\beta\right)} \leq C R^{1-d-\gamma}. \tag{7.12}$$

Obtaining this result relies on the quantitative behavior of the corrector and of the flux established in Proposition 6.28. Once one has a good control over the H^{-1}-norm of $\mathcal{L}\mathcal{H}_{\delta,\cdot k} - \rho_{\delta,\cdot k}$, the inequality (7.10) can be deduced from the following two arguments:

- We use that the function $\mathcal{G}_{\delta,\cdot k}$ satisfies the equation $\mathcal{L}\mathcal{G}_{\delta,\cdot k} = \rho_{\delta,\cdot k}$ to obtain that the H^{-1}-norm of the term $\mathcal{L}\left(\mathcal{H}_{\delta,\cdot k} - \mathcal{G}_{\delta,\cdot k}\right)$ is small. We then introduce a cutoff function $\eta : \mathbb{Z}^d \to \mathbb{R}$ which satisfies:

$$\operatorname{supp} \eta \subseteq A_R, \ 0 \leq \eta \leq 1, \ \eta = 1 \text{ on } \left\{ x \in \mathbb{Z}^d \, : \, 1.1R \leq |x| \leq 1.9R \right\}, \text{ and } \forall k \in \mathbb{N}, \ \left| \nabla^k \eta \right| \leq \frac{C}{R^k}, \tag{7.13}$$

  and use the function $\eta\left(\mathcal{H}_{\delta,\cdot k} - \mathcal{G}_{\delta,\cdot k}\right)$ as a test function in the definition of the H^{-1}-norm of the inequality (7.12). We obtain that the L^2-norm of the difference $\left(\nabla\mathcal{H}_{\delta,\cdot k} - \nabla\mathcal{G}_{\delta,\cdot k}\right)$ is small (the cutoff function is used to ensure that the function $\eta\left(\mathcal{H}_{\delta,\cdot k} - \mathcal{G}_{\delta,\cdot k}\right)$ is equal to 0 on the boundary of the ball $B_{R^{1+\delta}}$ and can thus be used as a test function). The precise estimate we obtain is the following

$$\left\| \nabla\mathcal{H}_{\delta,\cdot k} - \nabla\mathcal{G}_{\delta,\cdot k} \right\|_{L^2(\mathbb{Z}^d, \mu_\beta)} \leq C R^{\frac{d}{2}+1-d-\gamma}; \tag{7.14}$$

- By using the identity (7.11), we can compute an explicit formula for the gradient of the two-scale expansion $\mathcal{H}_{\delta,\cdot k}$. We then use the quantitative sublinearity of the corrector stated in Proposition 6.28 and the property of the gradient of the infinite volume corrector stated in Proposition 6.29 to deduce that the L^2-norm of the difference $\nabla\mathcal{H}_{\delta,\cdot k} - \sum_{i,j}(e_{ij} + \nabla\chi_{ij})\nabla_i\overline{G}_{jk}$ is small; the precise result we obtain is the following

$$\left\| \nabla\mathcal{H}_{\delta,\cdot k} - \sum_{i,j}(e_{ij} + \nabla\chi_{ij})\nabla_i\overline{G}_{jk} \right\|_{L^2(\mathbb{Z}^d, \mu_\beta)} \leq C R^{\frac{d}{2}+1-d-\gamma}. \tag{7.15}$$

The inequality (7.10) is then a consequence of the inequalities (7.14) and (7.15).

7.2.2. *Homogenization of the mixed derivative of the Green's matrix.* In this subsection, we present the arguments of Section 7.4. The objective there is to use Proposition 7.6 to prove Theorem 2. The proof is decomposed into four steps:

- In Step 1, we use Proposition 7.6 and the symmetry of the Helffer-Sjöstrand operator $\mathcal{L}$ to prove the inequality in expectation

$$\left(R^{-d} \sum_{z \in A_R} \left| \left\langle \mathcal{U}_{q_1}(z, \cdot) \right\rangle_{\mu_\beta} - \overline{G}_{q_1}(z) \right|^2 \right)^{\frac{1}{2}} \leq \frac{C}{R^{d-1+\gamma}}; \tag{7.16}$$

- In Step 2, we prove the variance estimate, for each point $z \in \mathbb{Z}^d$,

$$\operatorname{var}\left[\mathcal{U}_{q_1}(z, \cdot)\right] \leq \frac{C_{q_1}}{|z|^{2d-2\varepsilon}}. \tag{7.17}$$

  Since we expect the function $z \mapsto \mathcal{U}_{q_1}(z)$ to decay like $|z|^{1-d}$; its variance should be of order $|z|^{2-2d}$. The estimate (7.17) states that the variance of the random variable $\phi \to \mathcal{U}_{q_1}(z, \phi)$ is (quantitatively) smaller than its size; this means that the random variable $\mathcal{U}_{q_1}(z)$ concentrates around its expectation. We then use the result established in Step 1 to refine the result: since by (7.16), one knows that the expectation of the map $\mathcal{U}_{q_1}(z)$ is close to the function $\overline{G}_{q_1}$, one deduces that the function $\mathcal{U}_{q_1}$ is close to the function $\overline{G}_{q_1}$ in the $\underline{L}^2(A_R, \mu_\beta)$-norm. The precise estimate we obtain is the following

$$\left\| \mathcal{U}_{q_1} - \overline{G}_{q_1} \right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{C_{q_1}}{R^{d-1-\gamma}}. \tag{7.18}$$

The proof of the inequality (7.17) does not rely on tools from stochastic homogenization; we appeal to the Brascamp-Lieb inequality and use the properties of the second-order Helffer-Sjöstrand equation introduced in Section 5.4.

- In Step 3, we prove the estimate (7.3), the proof is similar to the argument presented in the proof of Proposition 7.6 and relies on a two-scale expansion; it is decomposed into two substeps.

  In Substep 3.1, We define the two-scale expansion $\mathcal{H}_{q_1}$ by the formula

$$\mathcal{H}_{q_1} := \overline{G}_{q_1} + \sum_{i,j} \nabla_i \overline{G}_{q_1,j} \chi_{m,ij}. \tag{7.19}$$

  We then use that the function $\overline{G}_{q_1}$ is a solution to the equation $\nabla \cdot \bar{\mathbf{a}}_\beta \nabla \overline{G}_{q_1} = 0$ in the annulus A_R to prove that the $\underline{H}^{-1}(A_R, \mu_\beta)$-norm of the term $\mathcal{L}\mathcal{H}_{q_1}$ over the annulus A_R is small; we show

$$\left\| \mathcal{L}\mathcal{H}_{q_1} \right\|_{\underline{H}^{-1}(A_R,\mu_\beta)} \leq \frac{C_{q_1}}{R^{d+\gamma}}. \tag{7.20}$$

  The proof is essentially a notational modification of the proof of the estimate (7.12), and is even simpler since we do not have to take into account the exponent δ and the function ρ_δ.

  In Substep 3.2, we use that the function $\mathcal{U}_{q_1}$ satisfies the identity $\mathcal{L}\mathcal{U}_{q_1} = 0$ in the set $A_R \times \Omega$ to deduce that the $\underline{H}^{-1}(A_R, \mu_\beta)$-norm of the term $\mathcal{L}(\mathcal{H}_{q_1} - \mathcal{U}_{q_1}) = \mathcal{L}\mathcal{H}_{q_1}$ is small. We then consider the map η defined in (7.13) and use the function $\eta(\mathcal{H}_{q_1} - \mathcal{U}_{q_1})$ as a test function in the definition of the $\underline{H}^{-1}(A_R, \mu_\beta)$-norm of the term $\mathcal{L}(\mathcal{H}_{q_1} - \mathcal{U}_{q_1})$. We obtain that the $\underline{L}^2(A_R^1, \mu_\beta)$-norm of the difference $\nabla \mathcal{H}_{q_1} - \nabla \mathcal{U}_{q_1}$ is small, where we used the notation $A_R^1 := \left\{ x \in \mathbb{Z}^d : 1.1R \leq |x| \leq 1.9R \right\}$. This is the subject of Substep 3.2 where we prove

$$\left\| \nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A_R^1,\mu_\beta)} \leq \frac{C_{q_1}}{R^{d+\gamma}}. \tag{7.21}$$

- Step 4 is the conclusion of the argument, we use the explicit formula for the two-scale expansion $\mathcal{H}_{q_1}$ given in (7.19), the quantitative sublinearity of the corrector stated in Proposition 6.28, and the quantitative estimate for the difference of the finite and infinite-volume gradient of the corrector stated in Proposition 6.29 to prove the estimate

$$\left\| \nabla \mathcal{H}_{q_1} - \sum_{i,j} \left(e_{ij} + \nabla \chi_{ij} \right) \nabla_i \overline{G}_{q_1,j} \right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{C_{q_1}}{R^{d+\gamma}}. \tag{7.22}$$

  The argument is a notational modification of the one used to prove (7.19). We finally combine the estimates (7.21) and (7.22) to obtain the estimate (7.3), and complete the proof of Theorem 2.

7.3. **Two-scale expansion and homogenization of the gradient of the Green's matrix.** This section is devoted to the proof of Proposition 7.6. We collect some preliminary results in Section 7.3.1 and prove Theorem 2 in Sections 7.3.2, 7.3.3 and 7.3.4 following the outline given in Section 7.2.

7.3.1. *Preliminary estimates.* In this section, we collect some preliminary properties which are used in the proof of Proposition 7.6.

We first introduce some notation for the exponent γ. As was already mentioned, this exponent depends on the parameters α, δ and ε; in the argument, we need to keep track of its order of magnitude and we proceed as follows:

- We use the notation γ_1 when the exponent is of order 1; a typical example is the exponent $\gamma_1 := 1 - c_0\alpha - c_1\delta - c_2\varepsilon$ for some constants c_0, c_1, c_2 depending only on the dimension d;
- We use the notation γ_α when the exponent is of order α; a typical example is the exponent $\gamma_\alpha := \alpha - c_0\delta - c_1\varepsilon$ for some constants c_0, c_1 depending only on the dimension d;
- We use the notation γ_δ when the exponent is of order δ; a typical example is the exponent $\gamma_\delta := \delta - c_0\varepsilon$ for some constant c_0 depending only on the dimension d.

We always have the ordering

$$0 < \gamma_\varepsilon \ll \gamma_\delta \ll \gamma_\alpha \ll \gamma_1.$$

We also allow the value of the exponents $\gamma_\varepsilon, \gamma_\delta, \gamma_\alpha, \gamma_1$ to vary from line to line in the argument as long as the order of magnitude is preserved. In particular, we may write

$$\gamma_1 = \gamma_1 - \alpha, \ \gamma_\alpha = \gamma_\alpha - \delta \text{ and } \gamma_\delta = \gamma_\delta - \varepsilon.$$

We are now able to collect and prove some regularity estimates pertaining to the Green's matrices $\mathcal{G}$, $\mathcal{G}_\delta$, $\overline{G}$ and $\overline{G}_\delta$.

Proposition 7.7. *The following properties hold:*

- *There exists an exponent $\gamma_\delta > 0$ such that one has the L^∞-estimates*

$$\|\nabla \mathcal{G}(x,\cdot) - \nabla \mathcal{G}_\delta(x,\cdot)\|_{L^\infty(A_R,\mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\delta}} \quad \text{and} \quad \left\|\nabla \overline{G} - \nabla \overline{G}_\delta\right\|_{L^\infty(A_R)} \leq \frac{C}{R^{d-1+\gamma_\delta}}; \tag{7.23}$$

- *The Green's matrix $\mathcal{G}_\delta$ satisfies the following L^∞-estimates*

$$\|\mathcal{G}_\delta\|_{L^\infty(\mathbb{Z}^d,\mu_\beta)} \leq \frac{C}{R^{(1-\delta)(d-2)}} \text{ and } \|\nabla \mathcal{G}_\delta\|_{L^\infty(\mathbb{Z}^d,\mu_\beta)} \leq \frac{C}{R^{(1-\delta)(d-1-\varepsilon)}}, \tag{7.24}$$

as well as the estimates

$$\|\mathcal{G}_\delta\|_{L^\infty(A_{R^{1+\delta}},\mu_\beta)} \leq \frac{C}{R^{(1+\delta)(d-2)}} \text{ and } \|\nabla \mathcal{G}_\delta\|_{L^\infty(A_{R^{1+\delta}},\mu_\beta)} \leq \frac{C}{R^{(1+\delta)(d-1-\varepsilon)}}; \tag{7.25}$$

- *The homogenized Green's matrix $\overline{G}_\delta$ satisfies the regularity estimate, for each integer $k \in \mathbb{N}$,*

$$\left\|\nabla^k \overline{G}_\delta\right\|_{L^\infty(\mathbb{Z}^d,\mu_\beta)} \leq \frac{C}{R^{(1-\delta)(d-2+k)}}, \tag{7.26}$$

as well as the estimate

$$\left\|\nabla^k \overline{G}_\delta\right\|_{L^\infty(A_{R^{1+\delta}},\mu_\beta)} \leq \frac{C}{R^{(1+\delta)(d-2+k)}}. \tag{7.27}$$

Proof of Proposition 7.7. The proof relies on the regularity estimates established in Section 5. We first note that, by definitions of the functions $\mathcal{G}$ and $\mathcal{G}_\delta$, we have the identities

$$\mathcal{G}(x,\phi) = \mathcal{G}_1(x,\phi;0) \quad \text{and} \quad \mathcal{G}_\delta(x,\phi) = \sum_{y \in B_{R^{1-\delta}}} \mathcal{G}_1(x,\phi;y)\,\rho_\delta(y), \tag{7.28}$$

where the product in the right side of (7.28) is the standard matrix product between $\mathcal{G}_1(x,\phi;y)$ and $\rho_\delta(y)$. Using that the map ρ_δ has total mass 1 and the regularity estimate on the Green's matrix stated in Proposition 5.13, we obtain, for each point $x \in A_R$,

$$\begin{aligned}
\|\nabla_x \mathcal{G}(x,\cdot;0) - \nabla_x \mathcal{G}_\delta(x,\cdot;y)\|_{L^\infty(\mu_\beta)} &\leq \sum_{y \in B_{R^{1-\delta}}} \rho_\delta(y)\,\|\nabla_x \mathcal{G}_1(x,\cdot;0) - \nabla_x \mathcal{G}_1(x,\cdot;y)\|_{L^\infty(\mu_\beta)} \\
&\leq R^{1-\delta} \sup_{y \in B_{R^{1-\delta}}} \|\nabla_x \nabla_y \mathcal{G}_1(x,\cdot;y)\|_{L^\infty(\mu_\beta)} \\
&\leq R^{1-\delta} \sup_{y \in B_{R^{1-\delta}}} |x-y|^{-d-\varepsilon} \\
&\leq R^{1-\delta} R^{-d-\varepsilon}.
\end{aligned}$$

This computation implies the estimate (7.23) with the exponent $\gamma_\delta = \delta - \varepsilon$ which is strictly positive by the assumption (7.6).

The estimate on the homogenized Green's matrix is similar and even simpler since we only have to work with the Green's matrix associated with the discrete elliptic operator $\nabla \cdot \bar{\mathbf{a}}_\beta \nabla$ on $\mathbb{Z}^d$; we omit the details.

The proof of the inequality (7.24) relies on the estimates on the Green's matrix and its gradient established in Proposition 3.17. We use the identity (7.28) and write, for each point $x \in \mathbb{Z}^d$,

$$\begin{aligned}
\|\mathcal{G}_\delta(x,\cdot)\|_{L^\infty(\mu_\beta)} = \sum_{y \in B_{R^{1-\delta}}} |\rho_\delta(y)|\,\|\mathcal{G}_1(x,\phi;y)\|_{L^\infty(\mu_\beta)} \leq \frac{1}{R^{(1-\delta)d}} \sum_{y \in B_{R^{1-\delta}}} \frac{C}{|x-y|^{d-2}} &\leq \frac{1}{R^{(1-\delta)d}} \sum_{y \in B_{R^{1-\delta}}} \frac{C}{|y|^{d-2}} \\
&\leq \frac{1}{R^{(1-\delta)(d-2)}}.
\end{aligned}$$

A similar computation shows the bound for the gradient of the Green's matrix and the bounds (7.25) in the annulus $A_{R^{1+\delta}}$.

To prove the regularity estimate (7.26), we use the definition of the map $\overline{G}_\delta$ given in (7.8) and note that

$$-\nabla \cdot \bar{\mathbf{a}}_\beta \nabla \left(\nabla^k \overline{G}_\delta \right) = \nabla^k \rho_\delta \text{ in } \mathbb{Z}^d.$$

We then use the properties of the function ρ_δ stated in (7.7) and standard estimates on the homogenized Green's matrix $\overline{G}$. We obtain, for each point $x \in \mathbb{Z}^d$,

$$\left| \nabla^k \overline{G}_\delta(x) \right| \leq \sum_{y \in B_{R^{1-\delta}}} \left| \nabla^k \rho_\delta(y) \right| \left| \overline{G}(x-y) \right| \leq \frac{C}{R^{(d+k)(1-\delta)}} \sum_{y \in B_{R^{1-\delta}}} \frac{1}{|x-y|^{d-2}} \leq \frac{C}{R^{(1-\delta)(d-2+k)}}.$$

There only remains to prove the estimate (7.27). To this end, we select a point $x \in A_{R^{1+\delta}}$ and write

$$\begin{aligned} \left| \nabla^k \overline{G}_\delta(x) \right| = \left| \sum_{y \in B_{R^{1-\delta}}} \nabla^k \overline{G}(x-y) \rho_\delta(y) \right| \leq \sum_{y \in B_{R^{1-\delta}}} \frac{|\rho_\delta(y)|}{|x-y|^{d-2+k}} &\leq \frac{C}{R^{(d-2+k)(1+\delta)}} \sum_{y \in B_{R^{1-\delta}}} |\rho_\delta(y)| \\ &\leq \frac{C}{R^{(1+\delta)(d-2+k)}}. \end{aligned}$$

□

We have now collected all the necessary preliminary ingredients for the proof of Proposition 7.6 and devote the rest of Section 7.3 to its demonstration.

7.3.2. *Estimating the weak norm of $\mathcal{L}\mathcal{H}_\delta - \rho_\delta$.* In this section, we fix an integer $k \in \left\{1, \ldots, \binom{d}{2}\right\}$, let $\mathcal{H}_{\delta,\cdot k}$ be the two-scale expansion introduced in (7.11) and prove that there exists an exponent $\gamma_\alpha > 0$ such that

$$\left\| \mathcal{L}\mathcal{H}_{\delta,\cdot k} - \rho_{\delta,\cdot k} \right\|_{\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\alpha}}. \tag{7.29}$$

The strategy is to use the explicit formula for the map $\mathcal{H}_{\delta,\cdot k}$ to compute the value of the term $\mathcal{L}\mathcal{H}_{\delta,\cdot k}$. We then prove that its $\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)$-norm is small by using the quantitative properties of the corrector stated in Proposition 6.28. We first write

$$\mathcal{L}\mathcal{H}_{\delta,\cdot k} = \underbrace{-\Delta_\phi \mathcal{H}_{\delta,\cdot k}}_{\text{Substep 1.1}} + \underbrace{\frac{1}{2\beta} \sum_{n \geq 1} \frac{1}{\beta^{\frac{n}{2}}} (-\Delta)^{n+1} \mathcal{H}_{\delta,\cdot k}}_{\text{Substep 1.2}} - \underbrace{\frac{1}{2\beta} \Delta \mathcal{H}_{\delta,\cdot k} + \sum_{q \in \mathcal{Q}} \nabla_q^* \cdot \mathbf{a}_q \nabla_q \mathcal{H}_{\delta,\cdot k}}_{\text{Substep 1.3}}. \tag{7.30}$$

We treat the three terms in the right side in three distinct substeps.

Substep 1.1 In this substep, we treat the term $-\Delta_\phi \mathcal{H}_{\delta,\cdot k}$. Since the homogenized Green's matrix $\overline{G}_{\delta,\cdot k}$ does not depend on the field ϕ, we have the formula

$$-\Delta_\phi \mathcal{H}_{\delta,\cdot k} = \sum_{i,j} \nabla_i \overline{G}_{\delta,jk} \left(\Delta_\phi \chi_{m,ij} \right). \tag{7.31}$$

Sustep 1.2. In this substep, we study the iteration of the Laplacian of the two-scale expansion. We prove the identity

$$\sum_{n \geq 1} \frac{1}{\beta^{\frac{n}{2}}} (-\Delta)^{n+1} \mathcal{H}_{\delta,\cdot k} = \sum_{i,j} \sum_{n \geq 1} \frac{1}{\beta^{\frac{n}{2}}} \nabla_i \overline{G}_{\delta,jk} (-\Delta)^{n+1} \chi_{m,ij} + R_{\Delta^n}, \tag{7.32}$$

where R_{Δ^n} is an error term which satisfies the $\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)$-estimate

$$\left\| R_{\Delta^n} \right\|_{\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\alpha}}. \tag{7.33}$$

We use the following identity for the iteration of the Laplacian on a product of functions: given two smooth functions $f, g \in C^\infty\left(\mathbb{R}^d\right)$, we have the identity

$$\Delta^n (fg) = \sum_{r=0}^{n} \sum_{l=0}^{r} \binom{n-r}{l} \left(\nabla^r \Delta^l f \right) \cdot \left(\nabla^r \Delta^{n-r-l} g \right). \tag{7.34}$$

We note that this formula is valid for continuous functions (with the continuous Laplacian), it can be adapted to the discrete setting by taking into considerations translations of the functions f and g. Since this adaptation does not affect the overall strategy of the proof, we ignore this technical difficulty in the rest of the argument and apply the formula (7.34) to the two-scale expansion $\mathcal{H}_{\delta,\cdot k}$ as such. We obtain

$$\Delta^n \mathcal{H}_{\delta,\cdot k} = \Delta^n \overline{G}_{\delta,\cdot k} + \sum_{i,j} \sum_{r=0}^{n} \sum_{l=0}^{r} \binom{n-r}{l} \left(\nabla^r \Delta^l \nabla_i \overline{G}_{\delta,jk} \right) \cdot \left(\nabla^r \Delta^{n-r-l} \chi_{m,ij} \right). \tag{7.35}$$

We first focus on the term $\Delta^n \overline{G}_{\delta,\cdot k}$ in the identity (7.35) and prove that it is small in the $\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)$-norm. Using the regularity estimate (7.26), we have, for each integer $n \geq 2$,

$$\begin{aligned} \left\| \Delta^n \overline{G}_{\delta,\cdot k} \right\|_{\underline{H}^{-1}(B_{R^{1+\delta}},\mu_\beta)} \leq C R^{1+\delta} \left\| \Delta^n \overline{G}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} &\leq C R^{1+\delta} \left\| \Delta^n \overline{G}_{\delta,\cdot k} \right\|_{L^\infty(B_{R^{1+\delta}})} \\ &\leq \frac{C^{2n} R^{1+\delta}}{R^{(1-\delta)(d-2+2n)}} \\ &\leq \frac{C^{2n} R^{1+\delta}}{R^{(1-\delta)(d-2+4)}} \\ &\leq \frac{C^{2n}}{R^{d-1+\gamma_1}}, \end{aligned} \tag{7.36}$$

where we have set $\gamma_1 := 2 + \delta(d+1) > 0$.

Using the regularity estimate (7.26) a second time, we can estimate the terms of the right side of the identity (7.35) with more than 3 derivatives on the homogenized Green's matrix $\overline{G}_\delta$. We obtain the following inequality: for each $(i,j) \in \{1,\ldots,d\} \times \{1,\ldots,\binom{d}{2}\}$ and each $(r,l) \in \{1,\ldots,d\}^2$ such that $l \leq n-k$ and $k + 2l \geq 2$,

$$\begin{aligned} &\left\| \left(\nabla^r \Delta^l \nabla_i \overline{G}_{\delta,jk} \right) \cdot \left(\nabla^r \Delta^{n-r-l} \chi_{m,ij} \right) \right\|_{\underline{H}^{-1}(B_{R^{1+\delta}},\mu_\beta)} \\ &\quad \leq C R^{1+\delta} \left\| \left(\nabla^r \Delta^l \nabla_i \overline{G}_{\delta,jk} \right) \cdot \left(\nabla^r \Delta^{n-r-l} \chi_{m,ij} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \\ &\quad \leq C R^{1+\delta} \left\| \nabla^r \Delta^l \nabla_i \overline{G}_{\delta,jk} \right\|_{L^\infty(B_{R^{1+\delta}})} \times \left\| \nabla^r \Delta^{n-r-l} \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \\ &\quad \leq \frac{C^{r+2l} R^{1+\delta}}{R^{(1-\delta)(d-1+2l+r)}} \left\| \nabla^r \Delta^{n-r-l} \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)}. \end{aligned} \tag{7.37}$$

We use that the discrete operator $\nabla^r \Delta^{n-r-l}$ is bounded in the space $L^2(B_{R^{1+\delta}})$ and Proposition 6.28 to estimate the L^2-norm of the corrector. We obtain

$$\left\| \nabla^r \Delta^{n-r-l} \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \leq C^{2n-2l} \left\| \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \leq C^{2n-2l} R^{(1+\delta)(1-\alpha)}. \tag{7.38}$$

Putting the estimates (7.37) and (7.38) together and using the inequality $3 \leq 2l + r \leq 2n$, we deduce that

$$\left\| \left(\nabla^r \Delta^l \nabla_i \overline{G}_{\delta,jk} \right) \cdot \left(\nabla^r \Delta^{n-r-l} \chi_{m,ij} \right) \right\|_{\underline{H}^{-1}(B_{R^{1+\delta}},\mu_\beta)} \leq \frac{C^{2n} R^{1+\delta}}{R^{(1-\delta)(d+2)}} R^{(1+\delta)(1-\alpha)} \leq \frac{C^{2n}}{R^{d-1+\gamma_1}}, \tag{7.39}$$

where we have set $\gamma_1 := 1 + \alpha - \alpha\delta + \delta(d-1) + \delta > 0$.

We then estimate the H^{-1}-norm of the terms corresponding to the parameters $r = 1$ and $l = 0$ in the sum in the right side of the identity (7.35). To estimate it, we select a function $h \in H_0^1(B_{R^{1+\delta}}, \mu_\beta)$ such that $\|h\|_{\underline{H}^1(B_{R^{1+\delta}},\mu_\beta)} \leq 1$. We use the function h as a test function, perform an integration by parts in the first line, use the Cauchy-Schwarz inequality in the second line and the continuity of the discrete Laplacian (as an operator acting on $L^2(\mathbb{Z}^d)$) in the third line

$$\begin{aligned} &\frac{1}{R^{(1+\delta)d}} \sum_{x \in B_{R^{1+\delta}}} \left\langle \left(\nabla \nabla_i \overline{G}_{\delta,jk}(x,\cdot) \right) \cdot \left(\nabla \Delta^{n-1} \chi_{m,ij}(x,\cdot) \right) h(x,\cdot) \right\rangle_{\mu_\beta} \\ &\quad = \frac{1}{R^{(1+\delta)d}} \sum_{x \in B_{R^{1+\delta}}} \left\langle \chi_{m,ij}(x,\cdot) \nabla \cdot \Delta^{n-1} \left(\left(\nabla \nabla_i \overline{G}_{\delta,jk}(x,\cdot) \right) h(x,\cdot) \right) \right\rangle_{\mu_\beta} \\ &\quad \leq \left\| \chi_{m,ij} \right\|_{L^2(B_{R^{1+\delta}},\mu_\beta)} \left\| \nabla \cdot \Delta^{n-1} \left(\nabla \nabla_i \overline{G}_{\delta,jk} h \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \\ &\quad \leq C^n \left\| \chi_{m,ij} \right\|_{L^2(B_{R^{1+\delta}},\mu_\beta)} \left\| \nabla \cdot \left(\nabla \nabla_i \overline{G}_{\delta,jk} h \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)}. \end{aligned} \tag{7.40}$$

Using the regularity estimate for the homogenized Green's matrix stated in (7.26) and the inequality $\|h\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \leq C R^{1+\delta}$ (which is a consequence of the assumption $\|h\|_{\underline{H}^1(B_{R^{1+\delta}},\mu_\beta)} \leq 1$ and the Poincaré inequality), we obtain

$$\left\| \nabla \cdot \left(\left(\nabla \nabla_i \overline{G}_{\delta,jk} \right) h \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \left\| \nabla^3 \overline{G}_{\delta,jk} h \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} + \left\| \nabla^2 \overline{G}_{\delta,jk} \nabla h \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C R^{1+\delta}}{R^{(1-\delta)(d+1)}} + \frac{C}{R^{(1-\delta)d}} \leq \frac{C}{R^{d-\delta(d+2)}}. \tag{7.41}$$

We then combine the estimate (7.40) with the inequality (7.41) and the quantitative sublinearity of the corrector to obtain

$$\frac{1}{R^{(1+\delta)d}} \sum_{x \in B_{R^{1+\delta}}} \left\langle \left(\nabla \nabla_i \overline{G}_{\delta,jk}(x) \right) \cdot \left(\nabla \Delta^{n-1} \chi_{m,ij}(x,\cdot) \right) h(x,\cdot) \right\rangle_{\mu_\beta} \leq \frac{C^n R^{(1+\delta)(1-\alpha)}}{R^{d-\delta(d+2)}} \leq \frac{C^n}{R^{d-1+\gamma_\alpha}}, \tag{7.42}$$

where we have set $\gamma_\alpha := \alpha(1+\delta) - \delta(d+3) > 0$.

By combining the identity (7.35) with the estimates (7.36), (7.39), (7.42) and choosing the inverse temperature β large enough so that the series $\left(\frac{C^n}{\beta^{\frac{n}{2}}} \right)_{n \in \mathbb{N}}$ is summable, we obtain the main result (7.32) and (7.33) of this substep.

Substep 1.3. In this substep, we study the term pertaining to the charges in the identity (7.30). We prove the expansion

$$\frac{-1}{2\beta} \Delta \mathcal{H}_{\delta,\cdot k} + \sum_{q \in \mathcal{Q}} \nabla_q^* \cdot \mathbf{a}_q \nabla_q \mathcal{H}_{\delta,\cdot k} = -\nabla \cdot \bar{\mathbf{a}}_\beta \nabla \overline{G}_{\delta,\cdot k} - \sum_{i,j} \frac{1}{2\beta} \nabla_i \overline{G}_{\delta,jk} \Delta \chi_{m,ij} + \sum_{i,j} \sum_{q \in \mathcal{Q}} \nabla_i \overline{G}_{\delta,jk} \nabla_q^* \cdot \mathbf{a}_q \nabla_q \chi_{m,ij} + R_Q. \tag{7.43}$$

where R_Q is an error term which satisfies the $\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)$-norm estimate

$$\| R_Q \|_{\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\alpha}}. \tag{7.44}$$

We first compute the gradient and the Laplacian of the two-scale expansion $\mathcal{H}_{\delta,\cdot k}$ using the notation of (A.6) to expand the gradient of a product. We obtain the formulae

$$\nabla \mathcal{H}_{\delta,\cdot k} = \nabla \overline{G}_{\delta,\cdot k} + \sum_{i,j} \left[\nabla \nabla_i \overline{G}_{\delta,jk} \otimes \chi_{m,ij} + \nabla_i \overline{G}_{\delta,jk} \nabla \chi_{m,ij} \right], \tag{7.45}$$

and

$$\Delta \mathcal{H}_{\delta,\cdot k} = \Delta \overline{G}_{\delta,\cdot k} + \sum_{i,j} \nabla \cdot \left(\nabla \nabla_i \overline{G}_{\delta,jk} \otimes \chi_{m,ij} \right) + \left(\nabla \nabla_i \overline{G}_{\delta,jk} \right) \cdot \left(\nabla \chi_{m,ij} \right) + \left(\nabla_i \overline{G}_{\delta,jk} \right) \Delta \chi_{m,ij}. \tag{7.46}$$

We first treat the term $\Delta \mathcal{H}_{\delta,\cdot k}$ and use the two following ingredients:

(i) We introduce the notation $R_{Q,1} := \sum_{i,j} \nabla \cdot \left(\nabla \nabla_i \overline{G}_{\delta,jk} \otimes \chi_{m,ij} \right)$. By using the regularity estimate (7.26) on the homogenized Green's matrix and the quantitative sublinearity of the corrector, we prove that this term is an error term and estimate its $\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)$-norm according to the following computation

$$\| R_{Q,1} \|_{\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)} \leq C \left\| \sum_{i,j} \nabla \nabla_i \overline{G}_{\delta,jk} \otimes \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \sum_{i,j} C \left\| \nabla \nabla_i \overline{G}_{\delta,jk} \right\|_{L^\infty(B_{R^{1+\delta}}, \mu_\beta)} \| \chi_{m,ij} \|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C R^{(1+\delta)(1-\alpha)}}{R^{(1-\delta)d}} \leq \frac{C}{R^{d-1+\gamma_\alpha}},$$

where we have set $\gamma_\alpha := \alpha(1+\delta) - \delta(d+1) > 0$.

(ii) Second, we use the identity $\Delta \overline{G}_{\delta,\cdot k} = \nabla \cdot \nabla \overline{G}_{\delta,\cdot k} = \sum_{i,j} \nabla \cdot \left(\nabla_i \overline{G}_{\delta,jk} e_{ij} \right)$.

We obtain

$$\Delta \mathcal{H}_{\delta,\cdot k} = \nabla \cdot \left(\sum_{i,j} \nabla_i \overline{G}_{\delta,jk} \left(e_{ij} + \nabla \chi_{m,ij} \right) \right) + \sum_{i,j} \left(\nabla_i \overline{G}_{\delta,jk} \right) \Delta \chi_{m,ij} + R_{Q,1}. \tag{7.47}$$

We then treat the term pertaining to the charges; the objective is to prove the identity

$$\sum_{q\in\mathcal{Q}} \nabla_q^* \cdot \mathbf{a}_q \nabla_q \mathcal{H}_{\delta,\cdot k} = \sum_{i,j} \nabla\nabla_i \overline{G}_{\delta,jk} \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q \left(l_{e_{ij}} + \chi_{m,ij} \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) + \sum_{q\in\mathcal{Q}} \nabla_i \overline{G}_{\delta,jk} \nabla_q^* \cdot \mathbf{a}_q \nabla_q \left(l_{e_{ij}} + \chi_{m,ij} \right) + R_{Q,2}, \tag{7.48}$$

where $R_{Q,2}$ is an error term which satisfies the estimate

$$\| R_{Q,2} \|_{\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\alpha}}. \tag{7.49}$$

To prove this result, we select a test function $h : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}}$ which belongs to the space $H^1_0 \left(B_{R^{1+\delta}}, \mu_\beta \right)$ and satisfies the estimate $\|h\|_{\underline{H}^1(B_{R^{1+\delta}}, \mu_\beta)} \leq 1$. For each charge $q \in \mathcal{Q}$, we select a point x_q which belongs to the support of the charge q arbitrarily. We then write

$$\sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q \mathcal{H}_{\delta,\cdot k} \nabla_q h = \sum_{q\in\mathcal{Q}} \mathbf{a}_q \left(n_q, \mathrm{d}^* \mathcal{H}_{\delta,\cdot k} \right) \left(n_q, \mathrm{d}^* h \right). \tag{7.50}$$

We use the exact formula for $\mathcal{H}_\delta$ and apply the codifferential. We obtain

$$\begin{aligned} \mathrm{d}^* \mathcal{H}_{\delta,\cdot k} = L_{2,\mathrm{d}^*} \left(\nabla \mathcal{H}_{\delta,\cdot k} \right) &= L_{2,\mathrm{d}^*} \left(\nabla \overline{G}_{\delta,\cdot k} + \sum_{i,j} \left[\nabla\nabla_i \overline{G}_{\delta,jk} \otimes \chi_{m,ij} + \nabla_i \overline{G}_{\delta,jk} \nabla \chi_{m,ij} \right] \right) \\ &= \mathrm{d}^* \overline{G}_{\delta,\cdot k} + \sum_{i,j} \nabla_i \overline{G}_{\delta,jk} \mathrm{d}^* \chi_{m,ij} + \sum_{i,j} L_{2,\mathrm{d}^*} \left(\nabla\nabla_i \overline{G}_{\delta,jk} \otimes \chi_{m,ij} \right). \end{aligned} \tag{7.51}$$

We record the following formula

$$\begin{aligned} \mathrm{d}^* \overline{G}_{\delta,\cdot k} = L_{2,\mathrm{d}^*} \left(\nabla \overline{G}_{\delta,\cdot k} \right) = L_{2,\mathrm{d}^*} \left(\sum_{i,j} \nabla_i \overline{G}_{\delta,jk} e_{ij} \right) &= L_{2,\mathrm{d}^*} \left(\sum_{i,j} \nabla_i \overline{G}_{\delta,jk} \nabla l_{e_{ij}} \right) \\ &= \sum_{i,j} \nabla_i \overline{G}_{\delta,jk} L_{2,\mathrm{d}^*} \left(\nabla l_{e_{ij}} \right) \\ &= \sum_{i,j} \nabla_i \overline{G}_{\delta,jk} \mathrm{d}^* l_{e_{ij}}. \end{aligned} \tag{7.52}$$

Putting the identities (7.51) and (7.52) back into (7.50), we obtain

$$\begin{aligned} \sum_{q\in\mathcal{Q}} \mathbf{a}_q \nabla_q \mathcal{H}_{\delta,\cdot k} \nabla_q h = & \underbrace{\sum_{i,j} \sum_{q\in\mathcal{Q}} \mathbf{a}_q \left(n_q, \nabla_i \overline{G}_{\delta,jk} \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, \mathrm{d}^* h \right)}_{(7.53)-(i)} \\ & + \underbrace{\sum_{q\in\mathcal{Q}} \mathbf{a}_q \left(n_q, L_{2,\mathrm{d}^*} \left(\nabla\nabla_i \overline{G}_{\delta,jk} \otimes \chi_{m,ij} \right) \right) \left(n_q, \mathrm{d}^* h \right)}_{(7.53)-(ii)}. \end{aligned} \tag{7.53}$$

The second term (7.53)-(ii) is an error term which is small and can be estimated thanks to the regularity estimate (7.26) and Young's inequality. We obtain

$$\begin{aligned} & \left| \left\langle \sum_{q\in\mathcal{Q}} \mathbf{a}_q \left(n_q, L_{2,\mathrm{d}^*} \left(\nabla\nabla_i \overline{G}_{\delta,jk} \otimes \chi_{m,ij} \right) \right) \left(n_q, \mathrm{d}^* h \right) \right\rangle_{\mu_\beta} \right| \\ & \leq \sum_{q\in\mathcal{Q}} e^{-c\sqrt{\beta}\|q\|_1} \|n_q\|_2^2 \left\| \nabla^2 \overline{G}_{\delta,jk} \right\|_{L^\infty(\mathbb{Z}^d, \mu_\beta)} \|\chi_{m,ij}\|_{L^2(\operatorname{supp} n_q, \mu_\beta)} \|\nabla h\|_{L^2(\operatorname{supp} n_q, \mu_\beta)} \\ & \leq \frac{C}{R^{(1-\delta)(d+1)}} \sum_{q\in\mathcal{Q}} e^{-c\sqrt{\beta}\|q\|_1} \|n_q\|_2^2 \|\chi_{m,ij}\|_{L^2(\operatorname{supp} n_q, \mu_\beta)} \|\nabla h\|_{L^2(\operatorname{supp} n_q, \mu_\beta)} \\ & \leq \frac{C}{R^{(1-\delta)(d+1)}} \sum_{q\in\mathcal{Q}} e^{-c\sqrt{\beta}\|q\|_1} \|n_q\|_2^2 \left(R^{-1+\alpha} \|\chi_{m,ij}\|^2_{L^2(\operatorname{supp} n_q, \mu_\beta)} + R^{1-\alpha} \|\nabla h\|^2_{L^2(\operatorname{supp} n_q, \mu_\beta)} \right). \end{aligned}$$

We then use the inequality, for each point $x \in \mathbb{Z}^d$,

$$\sum_{q \in \mathcal{Q}} e^{-c\sqrt{\beta}\|q\|_1} \|n_q\|_2^2 \mathbb{1}_{\{x \in \operatorname{supp} n_q\}} \leq C. \tag{7.54}$$

We deduce that

$$\left| \left\langle \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q, L_{2,\mathrm{d}^*} \left(\nabla \nabla_i \overline{G}_{\delta,jk} \otimes \chi_{m,ij} \right) \right) \left(n_q, \mathrm{d}^* h \right) \right\rangle_{\mu_\beta} \right| \\ \leq \frac{CR^{-1+\alpha}}{R^{(1-\delta)(d+1)}} \|\chi_{m,ij}\|_{L^2(B_{R^{1+\delta}},\mu_\beta)}^2 + \frac{CR^{1-\alpha}}{R^{(1-\delta)(d+1)}} \|\nabla h\|_{L^2(B_{R^{1+\delta}},\mu_\beta)}^2 .$$

We then use Proposition 6.28 and the assumption $\|\nabla h\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \leq 1$. We obtain

$$\left| \frac{1}{R^{(1+\delta)d}} \left\langle \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q, L_{2,\mathrm{d}^*} \left(\nabla \nabla_i \overline{G}_{\delta,jk} \otimes \chi_{m,ij} \right) \right) \left(n_q, \mathrm{d}^* h \right) \right\rangle_{\mu_\beta} \right| \leq \frac{CR^{1-\alpha}}{R^{(1-\delta)d}} \leq \frac{C}{R^{d-1+\gamma_\alpha}}, \tag{7.55}$$

where we have set $\gamma_\alpha = \alpha - \delta(d+1) > 0$.

To treat the term (7.53)-(i), we make use of the point x_q and write

$$\begin{aligned} \sum_{q \in \mathcal{Q}} \mathbf{a}_q &\left(n_q, \nabla_i \overline{G}_{\delta,jk} \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, \mathrm{d}^* h \right) \\ &= \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q, \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, \nabla_i \overline{G}_{\delta,jk} \mathrm{d}^* h \right) \\ &+ \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q, \left(\nabla_i \overline{G}_{\delta,jk} - \nabla_i \overline{G}_{\delta,jk}(x_q) \right) \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, \mathrm{d}^* h \right) \\ &+ \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, \left(\nabla_i \overline{G}_{\delta,jk} - \nabla_i \overline{G}_{\delta,jk}(x_q) \right) \mathrm{d}^* h \right). \end{aligned}$$

The terms on the second and third lines are error terms which are small, they can be estimated by the regularity estimate (7.26) on the gradient of the homogenized Green's matrix and Young's inequality as follows

$$\begin{aligned} &\left| \left\langle \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q, \left(\nabla_i \overline{G}_{\delta,jk} - \nabla_i \overline{G}_{\delta,jk}(x_q) \right) \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, \mathrm{d}^* h \right) \right\rangle_{\mu_\beta} \right| \\ &\leq \sum_{q \in \mathcal{Q}} e^{-c\sqrt{\beta}\|q\|_1} \|n_q\|_2 \left\| \nabla \overline{G}_{\delta,jk} - \nabla \overline{G}_{\delta,jk}(x_q) \right\|_{L^\infty(\operatorname{supp} n_q,\mu_\beta)} \|\nabla \chi_{m,ij}\|_{L^2(\operatorname{supp} n_q,\mu_\beta)} \|\nabla h\|_{L^2(\operatorname{supp} n_q,\mu_\beta)} \\ &\leq \frac{C}{R^{(1-\delta)d}} \sum_{q \in \mathcal{Q}} e^{-c\sqrt{\beta}\|q\|_1} \|n_q\|_2 \operatorname{diam} n_q \left(\|\nabla \chi_{m,ij}\|_{L^2(\operatorname{supp} n_q,\mu_\beta)}^2 + \|\nabla h\|_{L^2(\operatorname{supp} n_q,\mu_\beta)}^2 \right). \end{aligned}$$

We then apply the estimate (7.54), the bound $\|\nabla \chi_{m,ij}\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \leq C$ on the gradient of the corrector and the assumption $\|\nabla h\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \leq 1$ to conclude that

$$\left| \frac{1}{R^{(1+\delta)d}} \left\langle \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q, \left(\nabla_i \overline{G}_{\delta,jk} - \nabla_i \overline{G}_{\delta,jk}(x_q) \right) \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, \mathrm{d}^* h \right) \right\rangle_{\mu_\beta} \right| \leq \frac{C}{R^{(1-\delta)d}} \leq \frac{C}{R^{(d-1)+\gamma_1}}, \tag{7.56}$$

where we have set $\gamma_1 = 1 - \delta > 0$. The same argument proves the inequality

$$\left| \frac{1}{R^{(1+\delta)d}} \left\langle \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, \left(\nabla_i \overline{G}_{\delta,jk} - \nabla_i \overline{G}_{\delta,jk}(x_q) \right) \mathrm{d}^* h \right) \right\rangle_{\mu_\beta} \right| \leq \frac{C}{R^{(d-1)+\gamma_1}}, \tag{7.57}$$

with the same exponent $\gamma_1 > 0$. Combining the identity (7.53) with the estimates (7.55), (7.56), (7.57), we have obtained the following result: for each function $h \in H_0^1(B_{R^{1+\delta}}, \mu_\beta)$ such that $\|h\|_{\underline{H}^1(B_{R^{1+\delta}},\mu_\beta)} \leq 1$, one has the expansion

$$\frac{1}{R^{(1+\delta)d}} \sum_{q \in \mathcal{Q}} \mathbf{a}_q \nabla_q \mathcal{H}_{\delta,\cdot k} \nabla_q h = \frac{1}{R^{(1+\delta)d}} \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q, \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, \nabla_i \overline{G}_{\delta,jk} \mathrm{d}^* h \right) + O\left(\frac{C}{R^{d-1+\gamma_\alpha}} \right).$$

We then use the identity $\nabla_i \overline{G}_{\delta,jk} \mathrm{d}^* h = \mathrm{d}^* \left(\nabla_i \overline{G}_{\delta,jk} h \right) - L_{2,\mathrm{d}^*} \left(\nabla \nabla_i \overline{G}_{\delta,jk} \otimes h \right)$ which is established in (7.52). We deduce that

$$\sum_{q \in \mathcal{Q}} \mathbf{a}_q \nabla_q \mathcal{H}_{\delta,\cdot k} \nabla_q h = \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q, \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, \mathrm{d}^* \left(\nabla_i \overline{G}_{\delta,jk} h \right) \right) + \sum_{q \in \mathcal{Q}} \mathbf{a}_q \left(n_q, \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{m,ij} \right) \right) \left(n_q, L_{2,\mathrm{d}^*} \left(\nabla \nabla_i \overline{G}_{\delta,jk} \otimes h \right) \right) + O\left(\frac{C}{R^{d-1+\gamma_\alpha}} \right). \tag{7.58}$$

This implies the identity (7.48) and the estimate (7.49).

We now complete the proof of (7.43). To prove this identity, it is sufficient, in view of (7.47) and (7.48), to prove the estimate

$$\frac{1}{2\beta} \sum_{i,j} \left(\nabla \nabla_i \overline{G}_{\delta,jk} \right) \cdot \left(e_{ij} + \nabla \chi_{m,ij} \right) + \sum_{i,j} \left(\nabla \nabla_i \overline{G}_{\delta,jk} \right) \sum_{q \in \mathcal{Q}} \mathbf{a}_q \nabla_q \left(l_{e_{ij}} + \chi_{m,ij} \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) = -\nabla \cdot \left(\bar{\mathbf{a}}_\beta \nabla \overline{G}_{\delta,\cdot k} \right) + R_{Q,3}, \tag{7.59}$$

where the term $R_{Q,3}$ satisfies the estimate

$$\left\| R_{Q,3} \right\|_{\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\alpha}}. \tag{7.60}$$

The proof relies on the quantitative estimate for the $\underline{H}^{-1}\left(B_{R^{1+\delta}}, \mu_\beta\right)$-norm of the flux corrector stated in Proposition 6.28 and the regularity estimate (7.26) on the homogenized matrix $\overline{G}_\delta$ and the identity

$$\nabla \cdot \bar{\mathbf{a}}_\beta \nabla \overline{G}_{\delta,\cdot k} = \nabla \cdot \sum_{i,j} \nabla_i \overline{G}_{\delta,jk} \bar{\mathbf{a}}_\beta e_{ij} = \sum_{i,j} \nabla \nabla_i \overline{G}_{\delta,jk} \cdot \bar{\mathbf{a}}_\beta e_{ij}.$$

We select a function $h : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}}$ which belongs to the space $H^1_0\left(B_{R^{1+\delta}}, \mu_\beta\right)$ and such that $\|h\|_{\underline{H}^1(B_{R^{1+\delta}}, \mu_\beta)} \leq 1$. We use it as a test function and write

$$\begin{aligned} &\frac{1}{R^{(1+\delta)d}} \Bigg| \Bigg\langle \sum_{x \in B_{R^{1+\delta}}} \sum_{i,j} \frac{1}{2\beta} \left(\nabla \nabla_i \overline{G}_{\delta,jk}(x) \right) \cdot \left(e_{ij} + \nabla \chi_{m,ij}(x,\cdot) \right) h(x,\cdot) \\ &\quad + \sum_{x \in B_{R^{1+\delta}}} \sum_{i,j} \sum_{q \in \mathcal{Q}} \mathbf{a}_q \nabla_q \left(l_{e_{ij}} + \chi_{m,ij} \right) \left(n_q, L_{2,\mathrm{d}^*} \left(\nabla \nabla_i \overline{G}_{\delta,jk} \otimes h \right) \right) - \sum_{x \in B_{R^{1+\delta}}} \nabla \cdot \left(\bar{\mathbf{a}}_\beta \nabla \overline{G}_{\delta,\cdot k} \right)(x) h(x,\cdot) \Bigg\rangle_{\mu_\beta} \Bigg| \\ &\quad \leq C \sum_{i,j} \left\| \frac{1}{2\beta} \left(e_{ij} + \nabla \chi_{m,ij} \right) + \sum_{q \in \mathcal{Q}} \mathbf{a}_q \nabla_q \left(l_{e_{ij}} + \chi_{m,ij} \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) - \bar{\mathbf{a}}_\beta e_{ij} \right\|_{\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)} \left\| \nabla \nabla_i \overline{G}_{\delta,jk} \otimes h \right\|_{\underline{H}^1(B_{R^{1+\delta}}, \mu_\beta)}, \end{aligned} \tag{7.61}$$

where we have used the (tautological) identities, for each point $x \in B_{R^{1+\delta}}$, and each field $\phi \in \Omega$,

$$\left(\nabla \nabla_i \overline{G}_{\delta,jk}(x) \right) \cdot \left(e_{ij} + \nabla \chi_{m,ij}(x,\phi) \right) h(x,\phi) = \left(e_{ij} + \nabla \chi_{m,ij}(x,\phi) \right) \cdot \left(\nabla \nabla_i \overline{G}_{\delta,jk}(x) \otimes h(x,\phi) \right),$$

and

$$\nabla \cdot \left(\bar{\mathbf{a}}_\beta \nabla \overline{G}_{\delta,\cdot k} \right)(x) h(x,\phi) = \left(\sum_{i,j} \nabla \nabla_i \overline{G}_{\delta,jk}(x) \cdot \bar{\mathbf{a}}_\beta e_{ij} \right) h(x,\phi) = \sum_{i,j} \bar{\mathbf{a}}_\beta e_{ij} \cdot \left(\nabla \nabla_i \overline{G}_{\delta,jk}(x) \otimes h(x,\phi) \right).$$

We then use Proposition 6.28 to write

$$\sum_{i,j} \left\| \frac{1}{2\beta} \left(e_{ij} + \nabla \chi_{m,ij} \right) + \sum_{q \in \mathcal{Q}} \mathbf{a}_q \nabla_q \left(l_{e_{ij}} + \chi_{n,ij} \right) L^t_{2,\mathrm{d}^*} \left(n_q \right) - \bar{\mathbf{a}}_\beta e_{ij} \right\|_{\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)} \leq C R^{(1+\delta)(1-\alpha)}, \tag{7.62}$$

and the regularity estimate (7.26) to write

(7.63)
$$\begin{aligned}\left\|\nabla\nabla_i\overline{G}_{\delta,jk}h\right\|_{\underline{H}^1(B_{R^{1+\delta}},\mu_\beta)} &\leq \frac{1}{R^{1+\delta}}\left\|\nabla^2\overline{G}_{\delta,\cdot k}\right\|_{L^\infty(B_{R^{1+\delta}},\mu_\beta)}\|h\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} + \left\|\nabla^3\overline{G}_{\delta,\cdot k}\right\|_{L^\infty(B_{R^{1+\delta}},\mu_\beta)}\|h\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \\ &\quad + \left\|\nabla^2\overline{G}_{\delta,\cdot k}\right\|_{L^\infty(B_{R^{1+\delta}},\mu_\beta)}\|\nabla h\|_{\underline{L}^2(B_{R^{1+\delta}},\mu_\beta)} \\ &\leq C\left(\frac{1}{R^{d(1-\delta)}} + \frac{1}{R^{(d-1)(1-\delta)}} + \frac{R^{1+\delta}}{R^{(d+1)(1-\delta)}}\right) \\ &\leq \frac{1}{R^{d-\delta(d+2)}}.\end{aligned}$$

Combining the estimates (7.61), (7.62), and (7.63), we have obtained that, for each function $h \in H_0^1(B_{R^{1+\delta}},\mu_\beta)$ such that $\|h\|_{\underline{H}^1(B_{R^{1+\delta}},\mu_\beta)} \leq 1$,

$$\begin{aligned}(7.64)\qquad &\frac{1}{R^{(1+\delta)d}}\Bigg|\Bigg\langle \sum_{x\in B_{R^{1+\delta}}}\sum_{i,j}\left(\nabla\nabla_i\overline{G}_{\delta,jk}(x)\right)\cdot\left(e_{ij}+\nabla\chi_{m,ij}(x,\cdot)\right) \\ &\qquad + \sum_{i,j}\sum_{q\in\mathcal{Q}}\mathbf{a}_q\nabla_q\left(l_{e_{ij}}+\chi_{m,ij}\right)\left(n_q, L_{2,\mathrm{d}^*}\left(\nabla\nabla_i\overline{G}_{\delta,jk}\otimes h\right)\right) - \sum_{x\in B_{R^{1+\delta}}}\bar{\mathbf{a}}\Delta\overline{G}_{\delta,\cdot k}(x)\cdot h(x,\cdot)\Bigg\rangle_{\mu_\beta}\Bigg| \\ &\leq \frac{CR^{(1+\delta)(1-\alpha)}}{R^{d-\delta(d+2)}} \\ &\leq \frac{C}{R^{d-1+\gamma_\alpha}},\end{aligned}$$

where we have set $\gamma := \alpha(1+\delta) - \delta(d+3) > 0$. Since the inequality (7.64) is valid for any function $h \in H_0^1(B_{R^{1+\delta}},\mu_\beta)$ satisfying $\|h\|_{\underline{H}^1(B_{R^{1+\delta}},\mu_\beta)} \leq 1$, the estimate (7.64) is equivalent to the identity (7.59) and the $H^{-1}(B_{R^{1+\delta}},\mu_\beta)$-norm estimate (7.60). The proof of (7.59), and thus of (7.43), is complete.

Substep 1.4 In this substep, we conclude Step 1 and prove the estimate (7.29). We use the identity (7.30) and the identities (7.31) proved in Substep 1, (7.32) proved in Substep 2 and (7.43) proved in Substep 3. We obtain

$$\begin{aligned}(7.65)\quad \mathcal{L}\mathcal{H}_{\delta,\cdot k} &= -\nabla\cdot\bar{\mathbf{a}}_\beta\nabla\overline{G}_{\delta,\cdot k} \\ &\quad + \sum_{i,j}\nabla_i\overline{G}_{\delta,jk}\left(\Delta_\phi\chi_{m,ij} + \frac{1}{2\beta}\Delta\chi_{m,ij} + \sum_{q\in\mathcal{Q}}\nabla_q^*\cdot\mathbf{a}_q\nabla_q\left(l_{e_{ij}}+\chi_{m,ij}\right) + \frac{1}{2\beta}\sum_{n\geq1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1}\chi_{m,ij}\right) \\ &\quad + R_Q + R_{\Delta^n}.\end{aligned}$$

We then treat the three lines of the previous display separately. For the first line, we use the identity

$$(7.66)\qquad -\nabla\cdot\bar{\mathbf{a}}_\beta\nabla\overline{G}_{\delta,\cdot k} = \rho_{\delta,\cdot k} \text{ in } \mathbb{Z}^d.$$

For the second line, we use that, by the definition of the finite-volume corrector given in Definition 6.26, this map is a solution of the Helffer-Sjöstrand equation $\mathcal{L}\left(l_{e_{ij}}+\chi_{m,ij}\right) = 0$ in the set $B_{R^{1+\delta}}\times\Omega$. We obtain

$$\begin{aligned}(7.67)\quad &\sum_{i,j}\nabla_i\overline{G}_{\delta,jk}\left(\Delta_\phi\chi_{m,ij} + \frac{1}{2\beta}\Delta\chi_{m,ij} + \sum_{q\in\mathcal{Q}}\nabla_q^*\cdot\mathbf{a}_q\nabla_q\left(l_{e_{ij}}+\chi_{m,ij}\right) + \frac{1}{2\beta}\sum_{n\geq1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1}\chi_{m,ij}\right) \\ &\qquad = \sum_{i,j}\nabla_i\overline{G}_{\delta,jk}\mathcal{L}\left(l_{e_{ij}}+\chi_{m,ij}\right) \\ &\qquad = 0.\end{aligned}$$

For the third line, we use the estimates (7.33) and (7.44) on the error terms R_Q and R_{Δ^n} respectively. We obtain

$$(7.68)\qquad \left\|R_Q + R_{\Delta^n}\right\|_{\underline{H}^{-1}(B_{R^{1+\delta}},\mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\alpha}}.$$

A combination of the identities (7.65), (7.66), (7.67) and the estimate (7.68) proves the inequality

$$\left\|\mathcal{L}\mathcal{H}_{\delta,\cdot k} - \rho_{\delta,\cdot k}\right\|_{\underline{H}^{-1}(B_{R^{1+\delta}},\mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\alpha}}.$$

The proof of the estimate (7.29) is complete.

7.3.3. *Estimating the L^2-norm of the term $\nabla \mathcal{G}_\delta - \nabla \mathcal{H}_\delta$.* The objective of this section is to prove that the gradient of the Green's matrix $\nabla \mathcal{G}_\delta$ and the gradient of the two-scale expansion $\nabla \mathcal{H}_\delta$ are close in the $\underline{L}^2 (A_R, \mu_\beta)$-norm. More specifically, we prove that there exists an exponent $\gamma_\delta > 0$ such that one has the estimate

$$\|\nabla \mathcal{G}_\delta - \nabla \mathcal{H}_\delta\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{C}{R^{(d-1)+\gamma_\delta}}. \tag{7.69}$$

To prove this inequality, we work on the larger set $B_{R^{1+\delta}/2}$ and prove the estimate

$$\|\nabla \mathcal{G}_\delta - \nabla \mathcal{H}_\delta\|_{\underline{L}^2\left(B_{R^{1+\delta}/2},\mu_\beta\right)} \leq \frac{C}{R^{(1+\delta)(d-1-\varepsilon/2)}}. \tag{7.70}$$

The inequality (7.69) implies (7.70); indeed, by using that the annulus A_R is included in the ball $B_{R^{1+\delta}}$, we can compute

$$\begin{aligned}
\|\nabla \mathcal{G}_\delta - \nabla \mathcal{H}_\delta\|_{\underline{L}^2(A_R,\mu_\beta)} &\leq \left(\frac{\left|B_{R^{1+\delta}/2}\right|}{|A_R|} \right)^{\frac{1}{2}} \|\nabla \mathcal{G}_\delta - \nabla \mathcal{H}_\delta\|_{\underline{L}^2\left(B_{R^{1+\delta}/2},\mu_\beta\right)} \\
&\leq C \left(\frac{R^{d(1+\delta)}}{R^d} \right)^{\frac{1}{2}} \frac{C}{R^{(1+\delta)(d-1-\varepsilon/2)}} \\
&\leq \frac{C}{R^{d-1+\gamma_\delta}},
\end{aligned}$$

where we have set $\gamma_\delta := \delta(\frac{d}{2} - 1 - \varepsilon/2) > 0$. We now focus on the proof of the estimate (7.70) and fix an integer $k \in \left\{1, \ldots, \binom{d}{2}\right\}$ to write the proof. The strategy is to use the identity $\mathcal{L}\mathcal{G}_{\delta,\cdot k} = \rho_{\delta,\cdot k}$ to rewrite the estimate (7.29) in the following form

$$\|\mathcal{L}\left(\mathcal{H}_{\delta,\cdot k} - \mathcal{G}_{\delta,\cdot k}\right)\|_{\underline{H}^{-1}(B_{R^{1+\delta}},\mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\alpha}}. \tag{7.71}$$

We then use the function $\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}$ as a test function in the definition of the H^{-1}-norm in the inequality (7.71) to obtain the H^1-estimate stated in (7.70), as described in the outline of the proof at the beginning of this section. The overall strategy is relatively straightforward; however, one has to deal with the following technical difficulty. By definition of the H^{-1}-norm, one needs to use a function in $H^1_0 (B_{R^{1+\delta}}, \mu_\beta)$ as a test function; in particular the function must be equal to 0 outside the ball $B_{R^{1+\delta}}$. This condition is not verified by the function $\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}$ which is thus not a suitable test function. To overcome this issue, we introduce a cutoff function $\eta : \mathbb{Z}^d \to \mathbb{R}$ supported in the ball $B_{R^{1+\delta}}$ which satisfies the properties

$$0 \leq \eta \leq \mathbb{1}_{B_{R^{1+\delta}}}, \ \eta = 1 \text{ in } B_{\frac{R^{1+\delta}}{2}}, \text{ and } \forall k \in \mathbb{N}, \left|\nabla^k \eta\right| \leq \frac{C}{R^{(1+\delta)k}}, \tag{7.72}$$

and use the function $\eta\left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)$ as a test function. The main difficulty is thus to treat the cutoff function. This difficulty is similar to the one treated in the proof of the Caccioppoli inequality stated in Proposition 5.1 and we will omit some of the technical details of the argument.

We first write

$$\begin{aligned}
&\frac{1}{R^{(1+\delta)d}} \sum_{x \in B_{R^{1+\delta}}} \left\langle \eta\left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right) \mathcal{L}\left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right) \right\rangle_{\mu_\beta} \leq \|\mathcal{L}\left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\|_{\underline{H}^{-1}(B_{R^{1+\delta}},\mu_\beta)} \|\eta\left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\|_{\underline{H}^1(B_{R^{1+\delta}},\mu_\beta)} \\
&\qquad \leq \frac{C}{R^{d-1+\gamma_\alpha}} \|\eta\left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\|_{\underline{H}^1(B_{R^{1+\delta}},\mu_\beta)}.
\end{aligned} \tag{7.73}$$

We then treat the terms in the left and right sides of the inequality (7.73) separately. Regarding the left side, we prove the estimate

$$\|\eta\left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\|_{\underline{H}^1(B_{R^{1+\delta}},\mu_\beta)} \leq \frac{C}{R^{d-1-\delta d}}. \tag{7.74}$$

The proof relies on the properties of the cutoff function η stated in (7.72), the regularity estimate on the Green's matrix stated in Proposition 7.7, the L^∞-bound on the homogenized Green's matrix $\overline{G}_\delta$ stated in (7.26) and the bounds on the corrector and its gradient recalled below

$$\left\| \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq C R^{(1+\delta)(1-\alpha)}, \ \left\| \nabla \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq C \text{ and, } \sum_{x \in \mathbb{Z}^d} \left\| \partial_x \chi_{m,ij} \right\|^2_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq C.$$

We first write

$$\left\| \eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{H}^1(B_{R^{1+\delta}}, \mu_\beta)} \leq \underbrace{\frac{1}{R^{1+\delta}} \left\| \eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)}}_{(7.75)-(i)} + \underbrace{\left\| \nabla \eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)}}_{(7.75)-(ii)}$$
$$+ \underbrace{\left\| \eta \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)}}_{(7.75)-(iii)} + \underbrace{\beta \sum_{x \in \mathbb{Z}^d} \left\| \eta \left(\partial_x \mathcal{G}_{\delta,\cdot k} - \partial_x \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)}}_{(7.75)-(iv)}, \tag{7.75}$$

and treats the four terms in the right side separately. For the term (7.75)-(i), we use that the function η is non-negative and smaller than 1 to write

$$\frac{1}{R^{1+\delta}} \left\| \eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{1}{R^{1+\delta}} \left(\left\| \mathcal{G}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} + \left\| \mathcal{H}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \right).$$

We then estimate the L^2-norm of the Green's matrix $\mathcal{G}_\delta$ thanks to the estimate

$$\left\| \mathcal{G}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \left\| \mathcal{G}_{\delta,\cdot k} \right\|_{L^\infty(\mathbb{Z}^d, \mu_\beta)} \leq \frac{C}{R^{(1-\delta)(d-2)}}.$$

The L^2-norm of the two-scale expansion $\mathcal{H}_\delta$ can be estimated according to the following computation

$$\begin{aligned} \left\| \mathcal{H}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} &\leq \left\| \overline{G}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} + \sum_{i,j} \left\| \nabla_i \overline{G}_{\delta,jk} \right\|_{L^\infty(B_{R^{1+\delta}}, \mu_\beta)} \left\| \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \\ &\leq \frac{C}{R^{(1-\delta)(d-2)}} + \frac{C R^{(1+\delta)(1-\alpha)}}{R^{(1-\delta)(d-1)}} \\ &\leq \frac{C}{R^{(1-\delta)(d-2)}}, \end{aligned}$$

where we have used the inequality $\alpha \gg \delta$ in the third inequality. A combination of the three previous displays shows the estimate

$$\frac{1}{R^{1+\delta}} \left\| \eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C}{R^{1+\delta} \times R^{(1-\delta)(d-2)}} \leq \frac{C}{R^{d-1-\delta(d-3)}}. \tag{7.76}$$

The proof of the term (7.75)-(ii) is identical, we use the estimate $|\nabla \eta| \leq \frac{C}{R^{1+\delta}}$ and apply the estimate obtained for the term (7.75)-(ii). We obtain

$$\left\| \nabla \eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C}{R^{d-1-\delta(d-3)}}. \tag{7.77}$$

For the term (7.75)-(iii), we first write

$$\left\| \eta \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \left\| \nabla \mathcal{G}_{\delta,\cdot k} \right\|_{L^\infty(\mathbb{Z}^d, \mu_\beta)} + \left\| \nabla \mathcal{H}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)}.$$

The L^∞-norm of the Green's matrix $\nabla \mathcal{G}_{\delta,\cdot k}$ is estimated by Proposition 7.7. We have

$$\left\| \nabla \mathcal{G}_{\delta,\cdot k} \right\|_{L^\infty(\mathbb{Z}^d, \mu_\beta)} \leq \frac{C}{R^{(1-\delta)(d-1-\varepsilon)}}.$$

For the L^2-norm of the two-scale expansion $\mathcal{H}$, we use the formula (7.45) and write

$$\begin{aligned} \left\| \nabla \mathcal{H}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} &\leq \left\| \nabla \overline{G}_{\delta,\cdot k} \right\|_{L^\infty(\mathbb{Z}^d)} + \sum_{i,j} \left\| \nabla \nabla_i \overline{G}_{\delta,jk} \right\|_{L^\infty(\mathbb{Z}^d)} \left\| \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \\ &\quad + \sum_{i,j} \left\| \nabla_i \overline{G}_{\delta,jk} \right\|_{L^\infty(\mathbb{Z}^d)} \left\| \nabla \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \\ &\leq \frac{C}{R^{(1-\delta)(d-1-\varepsilon)}} + \frac{C R^{(1+\delta)(1-\alpha)}}{R^{(1-\delta)(d-\varepsilon)}} + \frac{C}{R^{(1-\delta)(d-1-\varepsilon)}} \\ &\leq \frac{C}{R^{d-1-\varepsilon-\delta(d-1-\varepsilon)}}. \end{aligned}$$

A combination of the three previous displays together with the inequality $\delta \gg \varepsilon$ yields the estimate

$$\left\| \eta \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C}{R^{d-1-\delta d}}. \tag{7.78}$$

There remains to estimate the term (7.75)-(iv). We first write

$$\beta \sum_{x \in \mathbb{Z}^d} \left\| \eta \left(\partial_x \mathcal{G}_{\delta,\cdot k} - \partial_x \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \underbrace{\beta \sum_{x \in \mathbb{Z}^d} \left\| \eta \partial_x \mathcal{G}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)}}_{(7.79)-(i)} + \underbrace{\beta \sum_{x \in \mathbb{Z}^d} \left\| \eta \partial_x \mathcal{H}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)}}_{(7.79)-(ii)} \tag{7.79}$$

and estimate the two terms in the right side separately. For the term (7.79)-(i), we use that the map $\mathcal{G}_{\delta,\cdot k}$ is a solution of the equation $\mathcal{L}\mathcal{G}_{\delta,\cdot k} = \rho_{\delta,\cdot k}$ and use the map $\eta^2 \mathcal{G}_{\delta,\cdot k}$ as a test function. We obtain

$$\begin{aligned} \beta \sum_{x \in \mathbb{Z}^d} \left\| \eta \partial_x \mathcal{G}_{\delta,\cdot k} \right\|^2_{L^2(B_{R^{1+\delta}}, \mu_\beta)} = & -\frac{1}{2} \sum_{x \in \mathbb{Z}^d} \left\langle \nabla \mathcal{G}_{\delta,\cdot k}(x,\cdot) \cdot \nabla \left(\eta^2 \mathcal{G}_{\delta,\cdot k} \right)(x,\cdot) \right\rangle_{\mu_\beta} - \beta \sum_{q \in \mathcal{Q}} \left\langle \nabla_q \mathcal{G}_{\delta,\cdot k} \cdot \mathbf{a}_q \nabla_q \left(\eta^2 \mathcal{G}_{\delta,\cdot k} \right) \right\rangle_{\mu_\beta} \\ & - \frac{1}{2} \sum_{n \geq 1} \sum_{x \in \mathbb{Z}^d} \frac{1}{\beta^{\frac{n}{2}}} \left\langle \nabla^{n+1} \mathcal{G}_{\delta,\cdot k}(x,\cdot) \cdot \nabla^{n+1} \left(\eta^2 \mathcal{G}_{\delta,\cdot k} \right)(x,\cdot) \right\rangle_{\mu_\beta} + \beta \sum_{x \in \mathbb{Z}^d} \rho_{\delta,\cdot k}(x) \eta^2(x) \cdot \left\langle \mathcal{G}_{\delta,\cdot k}(x,\cdot) \right\rangle_{\mu_\beta}. \end{aligned}$$

We then estimate the four terms in the right sides using the pointwise estimates on the function $\mathcal{G}_\delta$ and its gradient stated in Proposition 7.7, the properties on the functions ρ_δ and η stated in (7.7) and (7.72) respectively. We omit the technical details and obtain the estimate

$$\sum_{x \in \mathbb{Z}^d} \left\| \eta \partial_x \mathcal{G}_{\delta,\cdot k} \right\|^2_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C}{R^{2(1-\delta)(d-1-\varepsilon)}}. \tag{7.80}$$

The term (7.79)-(ii) involving the two-scale expansion is the easiest one to estimate; using the explicit formula for the map $\mathcal{H}_{\delta,\cdot k}$ and the fact that the function $\overline{G}_{\delta,\cdot k}$ does not depend on the field ϕ, we have the identity

$$\partial_x \mathcal{H}_{\delta,\cdot k} := \sum_{i,j} \nabla_i \overline{G}_{\delta,jk} \partial_x \chi_{m,ij}.$$

We deduce that

$$\begin{aligned} \sum_{x \in \mathbb{Z}^d} \left\| \eta \partial_x \mathcal{H}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} & \leq \sum_{x \in \mathbb{Z}^d} \left\| \partial_x \mathcal{H}_{\delta,\cdot k} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \\ & \leq C \sum_{i,j} \left\| \nabla \overline{G}_{\delta,\cdot k} \right\|_{L^\infty(\mathbb{Z}^d)} \sum_{x \in \mathbb{Z}^d} \left\| \partial_x \chi_{m,ij} \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \\ & \leq \frac{C}{R^{(1-\delta)(d-1-\varepsilon)}}. \end{aligned} \tag{7.81}$$

Combining the inequalities (7.79), (7.80) and (7.81) yields

$$\sum_{x \in \mathbb{Z}^d} \left\| \eta \left(\partial_x \mathcal{G}_{\delta,\cdot k} - \partial_x \mathcal{H}_{\delta,\cdot k} \right) \right\|_{\underline{L}^2(B_{R^{1+\delta}}, \mu_\beta)} \leq \frac{C}{R^{d-1-\varepsilon-\delta(d-1-\varepsilon)}} \leq \frac{C}{R^{d-1-\delta d}}. \tag{7.82}$$

The inequality (7.74) is then obtained by combining the estimates (7.76), (7.77), (7.78) and (7.82). We then put the inequality back into the inequality (7.73) and deduce that

$$\frac{1}{R^{(1+\delta)d}} \sum_{x \in B_{R^{1+\delta}}} \left\langle \eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k} \right) \mathcal{L} \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k} \right) \right\rangle_{\mu_\beta} \leq \frac{C}{R^{d-1+\gamma_\alpha} \times R^{d-1-\delta d}} \leq \frac{C}{R^{2d-2+\gamma_\alpha}}, \tag{7.83}$$

where we have used in the second inequality that the exponent γ_α is of order α and is thus much larger than the value δd.

In the rest of this step, we treat the left side of (7.83) and prove the inequality

$$\left\| \nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k} \right\|^2_{\underline{L}^2\left(B_{\frac{R^{1+\delta}}{2}}, \mu_\beta\right)} \leq \frac{1}{R^{(1+\delta)d}} \sum_{x \in B_{R^{1+\delta}}} \left\langle \eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k} \right) \mathcal{L} \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k} \right) \right\rangle_{\mu_\beta} + \frac{C}{R^{(1+\delta)(2d-2-\varepsilon)}}. \tag{7.84}$$

First, by definition of the Helffer-Sjöstrand operator $\mathcal{L}$, we have the identity

$$\begin{aligned}(7.85)\quad & \sum_{x\in\mathbb{Z}^d} \left\langle \eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right) \mathcal{L} \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\right\rangle_{\mu_\beta} \\ & = \sum_{x,y\in\mathbb{Z}^d} \eta(x) \left\langle \left(\partial_y \mathcal{G}_{\delta,\cdot k}(x,\cdot) - \partial_y \mathcal{H}_{\delta,\cdot k}(x,\cdot)\right)^2 \right\rangle_{\mu_\beta} \\ & + \frac{1}{2\beta} \sum_{x\in\mathbb{Z}^d} \left\langle \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot) \cdot \nabla \left(\eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\right)(x,\cdot) \right\rangle_{\mu_\beta} \\ & + \sum_{q\in\mathcal{Q}} \left\langle \nabla_q \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right) \cdot \mathbf{a}_q \nabla_q \left(\eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\right)\right\rangle_{\mu_\beta} \\ & + \frac{1}{2\beta} \sum_{n\geq 1} \sum_{x\in\mathbb{Z}^d} \frac{1}{\beta^{\frac{n}{2}}} \left\langle \nabla^{n+1} \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot) \cdot \nabla^{n+1} \left(\eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\right)(x,\cdot) \right\rangle_{\mu_\beta}.\end{aligned}$$

We then estimate the four terms on the right side separately. For the first one, we use that it is non-negative

$$\sum_{x,y\in\mathbb{Z}^d} \eta(x)^2 \left\langle \left(\partial_y \mathcal{G}_{\delta,\cdot k}(x,\cdot) - \partial_y \mathcal{H}_{\delta,\cdot k}(x,\cdot)\right)^2 \right\rangle_{\mu_\beta} \geq 0.$$

For the second one, we expand the gradient of the product $\eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}\right)$ and write

$$\begin{aligned}(7.86)\quad & \sum_{x\in\mathbb{Z}^d} \left\langle \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot) \cdot \nabla \left(\eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\right)(x,\cdot) \right\rangle_{\mu_\beta} \\ & = \sum_{x\in\mathbb{Z}^d} \eta(x) \left\langle \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot) \cdot \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot) \right\rangle_{\mu_\beta} \\ & \quad + \sum_{x\in\mathbb{Z}^d} \left\langle \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot) \cdot \nabla \eta(x) \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot) \right\rangle_{\mu_\beta}.\end{aligned}$$

we divide the identity (7.86) by the volume factor $R^{(1+\delta)d}$ and use the properties of the function η stated in (7.72). In particular, we use that the gradient of η is supported in the annulus $A_{R^{1+\delta}} := B_{R^{1+\delta}} \setminus B_{\frac{R^{1+\delta}}{2}}$ and obtain

$$\begin{aligned}(7.87)\quad & \frac{1}{R^{(1+\delta)d}} \sum_{x\in\mathbb{Z}^d} \left\langle \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot) \cdot \nabla \left(\eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\right)(x,\cdot) \right\rangle_{\mu_\beta} \geq c \left\| \eta \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right)\right\|^2_{\underline{L}^2\left(B_{R^{1+\delta}},\mu_\beta\right)} \\ & \qquad - \frac{C}{R^{1+\delta}} \left\| \nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right\|_{\underline{L}^2\left(A_{R^{1+\delta}},\mu_\beta\right)} \left\| \mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right\|_{\underline{L}^2\left(A_{R^{1+\delta}},\mu_\beta\right)}.\end{aligned}$$

By a computation similar to the one performed for the term (7.75)-(iii), but using the estimates (7.25) and (7.27) for the Green's matrices in the distant annulus $A_{R^{1+\delta}}$, instead of the L^∞-estimates (7.24) and (7.26). We obtain

$$(7.88)\quad \left\| \nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right\|_{\underline{L}^2\left(A_{R^{1+\delta}},\mu_\beta\right)} \leq \frac{C}{R^{(1+\delta)(d-1-\varepsilon)}} \text{ and } \left\| \mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right\|_{\underline{L}^2\left(A_{R^{1+\delta}},\mu_\beta\right)} \leq \frac{C}{R^{(1+\delta)(d-2)}}.$$

A combination of the inequalities (7.87) and (7.88) proves the estimate

$$\begin{aligned}(7.89)\quad & \frac{1}{R^{(1+\delta)d}} \sum_{x\in\mathbb{Z}^d} \left\langle \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot) \cdot \nabla \left(\eta \left(\mathcal{G}_{\delta} - \mathcal{H}_{\delta,\cdot k}\right)\right)(x,\cdot) \right\rangle_{\mu_\beta} + \frac{C}{R^{(1+\delta)(2d-2-\varepsilon)}} \\ & \qquad \geq c \left\| \eta \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right)\right\|^2_{\underline{L}^2\left(B_{R^{1+\delta}},\mu_\beta\right)}.\end{aligned}$$

The other terms in the right side of the identity (7.85) involving the sum over the iteration of the Laplacian and over the charges $q \in \mathcal{Q}$ are treated similarly and we omit the details. The results obtained are stated below

$$\begin{aligned}(7.90)\quad & \frac{1}{R^{(1+\delta)d}} \sum_{q\in\mathcal{Q}} \left\langle \nabla_q \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right) \cdot \mathbf{a}_q \nabla_q \left(\eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\right)\right\rangle_{\mu_\beta} + \frac{C}{R^{(1+\delta)(2d-2-\varepsilon)}} \\ & \qquad \geq -Ce^{-c\sqrt{\beta}} \left\| \eta \left(\nabla \mathcal{G}_{\delta,\cdot k} - \nabla \mathcal{H}_{\delta,\cdot k}\right)\right\|^2_{L^2\left(\mathbb{Z}^d,\mu_\beta\right)}\end{aligned}$$

and

$$\frac{1}{R^{(1+\delta)d}} \sum_{n\geq 1} \sum_{x\in\mathbb{Z}^d} \frac{1}{\beta^{\frac{n}{2}}} \left\langle \nabla^{n+1} \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot) \cdot \nabla^{n+1} \left(\eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\right)(x,\cdot)\right\rangle_{\mu_\beta} + \frac{C}{R^{(1+\delta)(2d-2-\varepsilon)}} \\ \geq \sum_{n\geq 1} \sum_{x\in\mathbb{Z}^d} \frac{1}{\beta^{\frac{n}{2}}} \left\langle \eta(x) \left|\nabla^{n+1} \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)(x,\cdot)\right|^2 \right\rangle_{\mu_\beta} \\ \geq 0. \tag{7.91}$$

We then combine the identity (7.85) with the estimates (7.89), (7.90) and (7.91) and assume that the inverse temperature β is large enough. We obtain

$$\frac{1}{R^{(1+\delta)d}} \sum_{x\in\mathbb{Z}^d} \left\langle \eta \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}\right) \mathcal{L} \left(\mathcal{G}_{\delta,\cdot k} - \mathcal{H}_{\delta,\cdot k}\right)\right\rangle_{\mu_\beta} + \frac{C}{R^{(1+\delta)(2d-2-\varepsilon)}} \geq c \left\|\eta \left(\nabla\mathcal{G}_{\delta,\cdot k} - \nabla\mathcal{H}_{\delta,\cdot k}\right)\right\|^2_{\underline{L}^2(\mathbb{Z}^d,\mu_\beta)} \\ \geq c \left\|\nabla\mathcal{G}_{\delta,\cdot k} - \nabla\mathcal{H}_{\delta,\cdot k}\right\|^2_{\underline{L}^2\left(B_{R^{1+\delta}/2},\mu_\beta\right)}.$$

The proof of the inequality (7.84) is then complete. To complete the proof of Step 2, we combine the estimates (7.83) and (7.84). We obtain

$$\left\|\nabla\mathcal{G}_{\delta,\cdot k} - \nabla\mathcal{H}_{\delta,\cdot k}\right\|^2_{\underline{L}^2\left(B_{\frac{R^{1+\delta}}{2}},\mu_\beta\right)} \leq \frac{C}{R^{2d-2+\gamma_\alpha}} + \frac{C}{R^{(1+\delta)(2d-2-\varepsilon)}} \leq \frac{C}{R^{(1+\delta)(2d-2-\varepsilon)}}, \tag{7.92}$$

where the last inequality is a consequence of the fact that γ_α is of order α and of the ordering $\alpha \gg \delta \gg \varepsilon$. Since the inequality (7.92) is valid for any integer $k \in \left\{1,\ldots,\binom{d}{2}\right\}$, the proof of the estimate (7.70) is complete.

7.3.4. *Homogenization of the gradient of the Green's matrix.* In this section, we post-process the conclusion (7.70) of Section 7.3.3 and prove that the gradient of the Green's matrix $\nabla\mathcal{G}_{\cdot k}$ is close to the map $\sum_{i,j}\left(e_{ij} + \nabla\chi_{ij}\right)\nabla_i\overline{G}_{jk}$. The objective is to prove that there exists an exponent $\gamma_\delta > 0$ such that

$$\left\|\nabla\mathcal{G}_{\cdot k} - \sum_{i,j}\left(e_{ij} + \nabla\chi_{ij}\right)\nabla_i\overline{G}_{jk}\right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\delta}}. \tag{7.93}$$

We first use the regularity estimates stated in Proposition 7.7 and the L^2-bound on the gradient of the infinite-volume corrector, for each $x \in \mathbb{Z}^d$, each pair of integers $(i,j) \in \{1,\ldots,d\} \times \left\{1,\ldots,\binom{d}{2}\right\}$, $\|\nabla\chi_{ij}(x,\cdot)\|_{L^2(\mu_\beta)} \leq C$. We write

$$\left\|\nabla\left(\mathcal{G}_{\cdot k} - \mathcal{G}_{\delta,\cdot k}\right) - \sum_{i,j}\left(e_{ij} + \nabla\chi_{ij}\right)\nabla_i\left(\overline{G}_{\delta,jk} - \overline{G}_{jk}\right)\right\|_{\underline{L}^2(A_R,\mu_\beta)} \\ \leq \left\|\nabla\left(\mathcal{G}_{\cdot k} - \mathcal{G}_{\delta,\cdot k}\right)\right\|_{\underline{L}^2(A_R,\mu_\beta)} + \sum_{i,j}\left\|\left(e_{ij} + \nabla\chi_{ij}\right)\right\|_{\underline{L}^2(A_R,\mu_\beta)}\left\|\nabla_i\left(\overline{G}_{\delta,jk} - \overline{G}_{jk}\right)\right\|_{L^\infty(A_R,\mu_\beta)} \\ \leq \frac{C}{R^{d-1+\gamma_\delta}}. \tag{7.94}$$

Using the inequality (7.94), we see that, to prove (7.93), it is sufficient to prove the estimate

$$\left\|\nabla\mathcal{G}_{\delta,\cdot k} - \sum_{i,j}\left(e_{ij} + \nabla\chi_{ij}\right)\nabla_i\overline{G}_{\delta,jk}\right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\delta}}. \tag{7.95}$$

We then use the main estimate (7.70) and deduce that, to prove the inequality (7.95), it is sufficient to prove

$$\left\|\nabla\mathcal{H}_{\delta,\cdot k} - \sum_{i,j}\left(e_{ij} + \nabla\chi_{ij}\right)\nabla_i\overline{G}_{\delta,jk}\right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\delta}}. \tag{7.96}$$

The rest of the argument of this step is devoted to the proof of (7.96). We first use the explicit formula for the gradient of the two-scale expansion $\nabla\mathcal{H}_{\delta,\cdot k}$ stated in (7.45) and write

$$\left\|\nabla\mathcal{H}_{\delta,\cdot k} - \sum_{i,j}\left(e_{ij} + \nabla\chi_{m,ij}\right)\nabla_i G_{\delta,jk}\right\|_{\underline{L}^2(A_R,\mu_\beta)} \\ \leq \sum_{i,j}\left\|\nabla\nabla_i\overline{G}_{\delta,jk}\chi_{m,ij}\right\|_{\underline{L}^2(A_R,\mu_\beta)} + \left\|\left(\nabla_i\overline{G}_{\delta,jk}\right)\left(\nabla\chi_{m,ij} - \nabla\chi_{ij}\right)\right\|_{\underline{L}^2(A_R,\mu_\beta)}.$$

We then use the regularity estimate (7.26), the quantitative sublinearity of the corrector stated in Proposition 6.28, and Proposition 6.29 to quantify the L^2-norm of the difference between the gradient of finite-volume corrector and the gradient of the infinite-volume corrector. We obtain

$$\left\| \nabla \mathcal{H}_{\delta,\cdot k} - \sum_{i,j} \left(e_{ij} + \nabla \chi_{m,ij}\right) \nabla_i \overline{G}_{\delta,jk} \right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \left(\frac{|A_R|}{|B_{R^{1+\delta}}|} \right)^{\frac{1}{2}} \frac{CR^{1-\alpha}}{R^{d-\varepsilon}} + \left(\frac{|A_R|}{|B_{R^{1+\delta}}|} \right)^{\frac{1}{2}} \frac{CR^{-\alpha}}{R^{d-1-\varepsilon}} \leq \frac{C}{R^{d-1+\gamma_\alpha}}, \tag{7.97}$$

where we have set $\gamma_\alpha := \alpha - \varepsilon - \frac{d\delta}{2} > 0$. Using that the exponent γ_α is larger than the exponent γ_δ completes the proof of the estimate (7.93).

7.4. **Homogenization of the mixed derivative of the Green's matrix.** The objective of this section is to use Proposition 7.6 to prove Theorem 2. We fix a charge $q_1 \in \mathcal{Q}$ and recall the definitions of the maps $\mathcal{U}_{q_1}$ and $\overline{G}_{q_1}$ given in the statement of Theorem 2. The proof is decomposed into three sections and follows the outline of the proof given in Section 7.2.2.

7.4.1. *Preliminary estimates.* In this section, we record some properties pertaining to the functions $\mathcal{U}_{q_1}$ and $\overline{G}_{q_1}$ which are used in the argument.

Proposition 7.8. *There exist an inverse temperature $\beta_0 := \beta_0(d) < 0$ and a constant C_{q_1} which satisfies the estimate $C_{q_1} \leq C \|q_1\|_1^k$, for some $C(d) < \infty$ and $k(d) < \infty$, such that the following statement holds: For each point $y \in \mathbb{Z}^d$ and each integer $k \in \mathbb{N}$, one has the estimates*

$$\|\nabla \mathcal{U}_{q_1}(y,\cdot)\|_{L^\infty(\mu_\beta)} \leq \frac{C_{q_1}}{|y|^{d-\varepsilon}}, \quad \|\mathcal{U}_{q_1}(y,\cdot)\|_{L^\infty(\mu_\beta)} \leq \frac{C_{q_1}}{|y|^{d-1-\varepsilon}} \quad \text{and} \quad \left|\nabla^k \overline{G}_{q_1}(y)\right| \leq \frac{C_{q_1}}{|y|^{d-1+k}}.$$

Proof. The proof is a consequence of the regularity estimates stated in Proposition 3.17 and the identity $q = \mathrm{d}n_q$. □

7.4.2. *Exploiting the symmetry of the Helffer-Sjöstrand operator.* The objective of this section is to use Proposition 7.6 and the symmetry of the Helffer-Sjöstrand operator $\mathcal{L}$ to prove the following estimate

$$\left(R^{-d} \sum_{z \in A_R} \left| \left\langle \mathcal{U}_{q_1}(z,\cdot) \right\rangle_{\mu_\beta} - \overline{G}_{q_1}(z) \right|^2 \right)^{\frac{1}{2}} \leq \frac{C}{R^{d-1+\gamma_\delta}}. \tag{7.98}$$

We start from the formula, for each integer $k \in \left\{1, \ldots, \binom{d}{2}\right\}$,

$$\left\| \mathrm{d}^* \mathcal{G}_{\cdot k} - \sum_{i,j} \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{ij} \right) \nabla_i \overline{G}_{jk} \right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\delta}}, \tag{7.99}$$

which is a direct consequence of Proposition 7.6 since the codifferential is a linear functional of the gradient. Using the estimate (7.99), we deduce that

$$R^{-d} \sum_{x \in A_R} \Big| \left\langle \cos\left(2\pi\left(\phi, q_1(x+\cdot)\right)\right) \left(n_{q_1}(x+\cdot), \mathrm{d}^* \mathcal{G}_{\cdot k}\right) \right\rangle_{\mu_\beta} - \sum_{i,j} \left\langle \cos\left(2\pi\left(\phi, q_1(x+\cdot)\right)\right) \left(n_{q_1}(x+\cdot), \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{ij}\right)\right) \right\rangle_{\mu_\beta} \nabla_i \overline{G}_{jk}(x) \Big| \leq \frac{C_{q_1}}{R^{d-1+\gamma_\delta}}.$$

By the translation invariance of the measure μ_β and the stationarity of the gradient of the infinite-volume corrector, we deduce that

$$\begin{aligned} \sum_{i,j} & \left\langle \cos\left(2\pi\left(\phi, q_1(x+\cdot)\right)\right) \left(n_{q_1}(x+\cdot), \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{ij}\right)\right) \right\rangle_{\mu_\beta} \nabla_i \overline{G}_{jk}(x) \\ & = \sum_{i,j} \left\langle \cos\left(2\pi\left(\phi, q_1\right)\right) \left(n_{q_1}, \left(\mathrm{d}^* l_{e_{ij}} + \mathrm{d}^* \chi_{ij}\right)\right) \right\rangle_{\mu_\beta} \nabla_i \overline{G}_{jk}(x) \\ & = \overline{G}_{q_1}(x). \end{aligned}$$

We now claim that we have the identity, for each point $x \in A_R$,

$$\left\langle \cos\left(2\pi\left(\phi, q_1(x+\cdot)\right)\right) \left(n_{q_1}(x+\cdot), \mathrm{d}^* \mathcal{G}\right) \right\rangle_{\mu_\beta} = \left\langle \mathcal{U}_{q_1(x+\cdot)}(0,\cdot) \right\rangle_{\mu_\beta}.$$

The proof of this result is a consequence of the symmetry of the Helffer-Sjöstrand operator $\mathcal{L}$ and the stationarity of the measure μ_β. We compute

$$\begin{aligned}\left\langle \cos\left(2\pi\left(\phi, q_1(x+\cdot)\right)\right)\left(n_{q_1}\left(x+\cdot\right), \mathrm{d}^*\mathcal{G}\right)\right\rangle_{\mu_\beta} &= \left\langle \left(\cos\left(2\pi\left(\phi, q_1(x+\cdot)\right)\right) q_1\left(x+\cdot\right), \mathcal{G}\right)\right\rangle_{\mu_\beta} \\ &= \left\langle \left(\cos\left(2\pi\left(\phi, q_1(x+\cdot)\right)\right) q_1\left(x+\cdot\right), \mathcal{L}^{-1}\delta_0\right)\right\rangle_{\mu_\beta} \\ &= \left\langle \left(\mathcal{L}^{-1}\cos\left(2\pi\left(\phi, q_1(x+\cdot)\right)\right) q_1\left(x+\cdot\right), \delta_0\right)\right\rangle_{\mu_\beta} \\ &= \left\langle \mathcal{U}_{q_1(x+\cdot)}(0,\cdot)\right\rangle_{\mu_\beta}.\end{aligned}$$

A combination of the four previous displays implies

$$R^{-d}\sum_{x\in A_R}\left|\left\langle \mathcal{U}_{q_1(x+\cdot)}(0,\cdot)\right\rangle_{\mu_\beta} - \overline{G}_{q_1}(x)\right| \leq \frac{C}{R^{d-1+\gamma_\delta}}. \tag{7.100}$$

We then use the translation invariance of the measure μ_β and the definition of the map $\mathcal{U}_{q_1}$ as the solution of the Helffer-Sjöstrand equation (7.2) to write

$$\left\langle \mathcal{U}_{q_1(x+\cdot)}(0,\cdot)\right\rangle_{\mu_\beta} = \left\langle \mathcal{U}_{q_1}(x,\cdot)\right\rangle_{\mu_\beta}. \tag{7.101}$$

Combining the inequality (7.100) with the identity (7.101), we obtain

$$R^{-d}\sum_{x\in A_R}\left|\left\langle \mathcal{U}_{q_1}(x,\cdot)\right\rangle_{\mu_\beta} - \overline{G}_{q_1}(x)\right| \leq \frac{C}{R^{d-1+\gamma_\delta}}. \tag{7.102}$$

We finally upgrade the L^1-inequality stated in (7.102) into an L^2-inequality: by using Proposition 7.8, we write

$$\begin{aligned}&R^{-d}\sum_{x\in A_R}\left|\left\langle \mathcal{U}_{q_1}(x,\cdot)\right\rangle_{\mu_\beta} - \overline{G}_{q_1}(x)\right|^2 \\ &\quad\leq \left(R^{-d}\sum_{x\in A_R}\left|\left\langle \mathcal{U}_{q_1}(x,\cdot)\right\rangle_{\mu_\beta} - \overline{G}_{q_1}(x)\right|\right)\left(\left\|\mathcal{U}_{q_1}(x,\cdot)\right\|_{L^\infty(A_R,\mu_\beta)} + \left\|\overline{G}_{q_1}\right\|_{L^\infty(A_R)}\right) \\ &\quad\leq \frac{C_{q_1}}{R^{d-1+\gamma_\delta}\times R^{d-1-\varepsilon}} \\ &\quad\leq \frac{C_{q_1}}{R^{2d-2+\gamma_\delta}},\end{aligned}$$

where we have used the convention notation described at the beginning of Section 7.3 to absorb the exponent ε into the exponent γ_δ in the third inequality.

7.4.3. *Contraction of the variance of* $\mathcal{U}_{q_1}$. In this section, we prove that the random variable $\mathcal{U}_{q_1}$ contracts around its expectation. To this end, we prove the variance estimate, for each point $z\in\mathbb{Z}^d$,

$$\operatorname{var}\left[\mathcal{U}_{q_1}(z,\cdot)\right] \leq \frac{C_{q_1}}{|z|^{2d-2\varepsilon}}. \tag{7.103}$$

Let us make a comment about the result: since the size of the random variable $\mathcal{U}_{q_1}(z,\cdot)$ is of order $|z|^{1-d}$ (since it behaves like the gradient of a Green's function), we would expect its variance to be of order $|z|^{2-2d}$. The inequality (7.103) asserts that it is in fact of order $|z|^{2\varepsilon-2d}$ which is smaller than the typical size of the random variable $\mathcal{U}_{q_1}(z,\cdot)$ by an algebraic factor: the random variable $\mathcal{U}_{q_1}(z,\cdot)$ concentrate around its expectation.

Once this estimate is established, we can combine it with the estimate (7.98) established in Section 7.4.2 to prove that the map $\mathcal{U}_{q_1}$ is close to the (deterministic) Green's function $\overline{G}_{q_1}$ in the $\underline{L}^2(A_R,\mu_\beta)$-norm: we obtain the inequality

$$\left\|\mathcal{U}_{q_1} - \overline{G}_{q_1}\right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\delta}}. \tag{7.104}$$

We now prove the variance estimate (7.103). We first apply the Brascamp-Lieb inequality and write

$$\operatorname{var}\left[\mathcal{U}_{q_1}(z,\cdot)\right] \leq C\sum_{y,y_1\in\mathbb{Z}^d}\left\|\partial_y\mathcal{U}_{q_1}(z,\cdot)\right\|_{L^2(\mu_\beta)}\frac{C}{|y-y_1|^{d-2}}\left\|\partial_{y_1}\mathcal{U}_{q_1}(z,\cdot)\right\|_{L^2(\mu_\beta)}. \tag{7.105}$$

A consequence of the inequality (7.105) is that, to estimate the variance of the random variable $\mathcal{U}_{q_1}(z,\cdot)$, it is sufficient to understand the behavior of the mapping $y\mapsto\partial_y\mathcal{U}_{q_1}(z,\cdot)$. To this end, we appeal to the second-order

Helffer-Sjöstrand equation: following the arguments developed in Section 5.4, the map $u : (y,z,\phi) \mapsto \partial_y \mathcal{U}_{q_1}(z,\phi)$ is solution of the second-order equation

$$\mathcal{L}_{\mathrm{sec}} u(x,y,\phi) = -\sum_{q \in \mathcal{Q}} 2\pi z(\beta,q) \cos(2\pi(\phi,q)) (\mathcal{U}_{q_1}, q)\, q(x) \otimes q(y) + 2\pi \sin(2\pi(\phi,q_1))\, q_1(x) \otimes q_1(y) \quad \text{in} \quad \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega.$$

The function u can be expressed in terms of the Green's matrix $\mathcal{G}_{\mathrm{sec}}$, and we write, for each triplet $(x,y,\phi) \in \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega$,

$$\begin{aligned} u(x,y,\phi) = \sum_{q \in \mathcal{Q}} 2\pi z(\beta,q) & \sum_{x_1,y_1 \in \mathbb{Z}^d} \mathrm{d}^*_{x_1} \mathrm{d}^*_{y_1} \mathcal{G}_{\mathrm{sec}, \cos(2\pi(\phi,q))(\mathcal{U}_{q_1},q)}(x,y,\phi;x_1,y_1)\, n_q(x_1) \otimes n_q(y_1) \\ & + \sum_{x_1,y_1 \in \mathbb{Z}^d} 2\pi \mathrm{d}^*_{x_1} \mathrm{d}^*_{y_1} \mathcal{G}_{\mathrm{sec}, \sin(2\pi(\phi,q_1))}(x,y,\phi;x_1,y_1)\, n_{q_1}(x_1) \otimes n_{q_1}(y_1). \end{aligned}$$

We use the regularity estimates on the Green's matrix stated in Proposition 5.13 to obtain, for each pair of points $(x,y) \in \mathbb{Z}^d \times \mathbb{Z}^d$,

$$\|u(x,y,\cdot)\|_{L^\infty(\mu_\beta)} \leq \underbrace{C \sum_{q \in \mathcal{Q}} e^{-c\sqrt{\beta}\|q\|_1} \sum_{x_1,y_1 \in \mathbb{Z}^d} \frac{|n_q(x_1)|\,|n_q(y_1)|\, \|(\mathrm{d}^*\mathcal{U}_{q_1}, n_q)\|_{L^\infty(\mu_\beta)}}{|x-x_1|^{2d-\varepsilon} + |y-y_1|^{2d-\varepsilon}}}_{(7.106)-(i)} + \underbrace{\sum_{x_1,y_1 \in \mathbb{Z}^d} \frac{|n_{q_1}(x_1)|\,|n_{q_1}(y_1)|}{|x-x_1|^{2d-\varepsilon} + |y-y_1|^{2d-\varepsilon}}}_{(7.106)-(ii)}. \tag{7.106}$$

We then estimate the two terms (7.106)-(i) and (7.106)-(ii) separately. We first focus on the term (7.106)-(i) and prove the inequality

$$(7.106)-(i) \leq \frac{C_{q_1}}{|x-y|^{d-\varepsilon} \max(|x|,|y|)^{d-1}} \tag{7.107}$$

To prove the estimate (7.107), we first decompose the set of charges $\mathcal{Q}$ according to the following procedure. For each $z \in \mathbb{Z}^d$, we denote by $\mathcal{Q}_z$ the set of charges $q \in \mathcal{Q}$ such that the point z belongs to the support of n_q, i.e., $\mathcal{Q}_z := \{q \in \mathcal{Q} : z \in \operatorname{supp} n_q\}$. We note that we have the equality $\mathcal{Q} := \bigcup_{z \in \mathbb{Z}^d} \mathcal{Q}_z$ but the collection $(\mathcal{Q}_z)_{z \in \mathbb{Z}^d}$ is not a partition of $\mathcal{Q}$. We first prove that, for each point $z \in \mathbb{Z}^d$,

$$\sum_{q \in \mathcal{Q}_z} e^{-c\sqrt{\beta}\|q\|_1} \sum_{x_1,y_1 \in \mathbb{Z}^d} \frac{|n_q(x_1)|\,|n_q(y_1)|\, \|(\mathrm{d}^*\mathcal{U}_{q_1}, n_q)\|_{L^\infty(\mu_\beta)}}{|x-x_1|^{2d-\varepsilon} + |y-y_1|^{2d-\varepsilon}} \leq \frac{C_{q_1}}{(|x-z|^{2d-\varepsilon} + |y-z|^{2d-\varepsilon}) \times |z|^{d-\varepsilon}}. \tag{7.108}$$

To prove the estimate (7.108), we first use Proposition 7.8 to estimate the term $\|(\mathrm{d}^*\mathcal{U}_{q_1}, n_q)\|_{L^\infty(\mu_\beta)}$. We write, for each charge $q \in \mathcal{Q}_z$,

$$\|(\mathrm{d}^*\mathcal{U}_{q_1}, n_q)\|_{L^\infty(\mu_\beta)} \leq \|\nabla \mathcal{U}_{q_1}\|_{L^\infty(\operatorname{supp} n_q, \mu_\beta)} \|n_q\|_{L^1} \leq C_{q,q_1} \sup_{z_1 \in \operatorname{supp} n_q} \frac{1}{|z_1|^{d-\varepsilon}} \leq \frac{C_{q,q_1}}{|z|^{d-\varepsilon}}. \tag{7.109}$$

Putting the inequality (7.109) into the left side of the estimate (7.108), we obtain

$$\begin{aligned} \sum_{q \in \mathcal{Q}_z} e^{-c\sqrt{\beta}\|q\|_1} & \sum_{x_1,y_1 \in \mathbb{Z}^d} \frac{|n_q(x_1)|\,|n_q(y_1)|\, \|(\mathrm{d}^*\mathcal{U}_{q_1}, n_q)\|_{L^\infty(\mu_\beta)}}{|x-x_1|^{2d-\varepsilon} + |y-y_1|^{2d-\varepsilon}} \\ & \leq \sum_{q \in \mathcal{Q}_z} \frac{C e^{-c\sqrt{\beta}\|q\|_1} C_{q,q_1}}{|z|^{d-\varepsilon}} \sum_{x_1,y_1 \in \operatorname{supp} n_q} \frac{1}{|x-x_1|^{2d-\varepsilon} + |y-y_1|^{2d-\varepsilon}}. \end{aligned} \tag{7.110}$$

The term in the right side of (7.110) can be explicitly computed by using the exponential decay of the term $e^{-c\sqrt{\beta}\|q\|_1}$ and we obtain

$$\sum_{q \in \mathcal{Q}_z} e^{-c\sqrt{\beta}\|q\|_1} \sum_{x_1,y_1 \in \mathbb{Z}^d} \frac{|n_q(x_1)|\,|n_q(y_1)|\, \|(\mathrm{d}^*\mathcal{U}_{q_1}, n_q)\|_{L^\infty(\mu_\beta)}}{|x-x_1|^{2d-\varepsilon} + |y-y_1|^{2d-\varepsilon}} \leq \frac{C_{q_1}}{|z|^{d-\varepsilon} \times (|x-z|^{2d-\varepsilon} + |y-z|^{2d-\varepsilon})}.$$

Summing over all the points $z \in \mathbb{Z}^d$, we obtain

$$(7.106)-(i) \leq \sum_{z \in \mathbb{Z}^d} \frac{C_{q_1}}{|z|^{d-\varepsilon} \times (|x-z|^{2d-\varepsilon} + |y-z|^{2d-\varepsilon})}. \tag{7.111}$$

The term on the right side can be explicitly estimated. We omit the details and give the result

$$(7.106) - (i) \leq \frac{C}{|x-y|^d \max(|x|,|y|)^{d-2\varepsilon}}. \tag{7.112}$$

Combining the estimates (7.111) and (7.112) completes the proof of the estimate (7.107).

The term (7.106)-(ii) can also be estimated by an explicit computation which we skip here. We obtain

$$(7.106) - (ii) \leq \frac{C_{q_1}}{|x|^{2d-\varepsilon} + |y|^{2d-\varepsilon}}. \tag{7.113}$$

We then combine the estimates (7.106), (7.107), (7.113) to deduce the inequality, for each pair of points $x, y \in \mathbb{Z}^d$,

$$\|u(x,y,\cdot)\|_{L^\infty(\mu_\beta)} \leq \frac{C_{q_1}}{|x-y|^{d-\varepsilon} \max(|x|,|y|)^{d-\varepsilon}}.$$

We use this inequality to estimate the variance of the random variable $\mathcal{U}_{q_1}(x,\cdot)$ by using the formula (7.105). We obtain

$$\operatorname{var}[\mathcal{U}_{q_1}(z,\cdot)] \leq C \sum_{y,y_1 \in \mathbb{Z}^d} \frac{C_{q_1}}{|z-y|^{d-\varepsilon} \max(|z|,|y|)^{d-1}} \cdot \frac{C}{|y-y_1|^{d-2}} \cdot \frac{C_{q_1}}{|z-y_1|^{d-\varepsilon} \max(|z|,|y_1|)^{d-\varepsilon}}.$$

We use that the terms $\max(|z|,|y_1|)$ and $\max(|z|,|y|)$ are both larger than the value $|z|$ to deduce that

$$\operatorname{var}[\mathcal{U}_{q_1}(z,\cdot)] \leq \frac{C_{q_1}}{|z|^{2d-2\varepsilon}} \sum_{y,y_1 \in \mathbb{Z}^d} \frac{1}{|z-y|^{d-\varepsilon}} \cdot \frac{1}{|y-y_1|^{d-2}} \cdot \frac{1}{|z-y_1|^{d-\varepsilon}} \leq \frac{C_{q_1}}{|z|^{2d-2\varepsilon}}.$$

The proof of the estimate (7.103) is complete.

7.4.4. *Homogenization of the mixed derivative of the Green's matrix.* The objective of this section is to complete the proof of Theorem 2. We fix a radius $R > 1$ and let m be the smallest integer such that the annulus A_R is included in the cube $\square_m$. The proof relies on a two-scale expansion following the outline described in Section 7.2.2. We define the function $\mathcal{H}_{q_1}$ by the formula

$$\mathcal{H}_{q_1} := \overline{G}_{q_1} + \sum_{i,j} \nabla_i \overline{G}_{q_1,j} \chi_{m,ij}. \tag{7.114}$$

We decompose the argument into three steps.

Step 1. In this step, we prove that the $\underline{H}^{-1}(A_R,\mu_\beta)$-norm of the term $\mathcal{L}\mathcal{H}_{q_1}$ is small; more specifically, we prove that there exists an exponent $\gamma_\alpha > 0$ such that one has the estimate

$$\|\mathcal{L}\mathcal{H}_{q_1}\|_{\underline{H}^{-1}(A_R,\mu_\beta)} \leq \frac{C_{q_1}}{R^{d+\gamma_\alpha}}. \tag{7.115}$$

The proof is essentially identical to the argument presented in Section 7.3.2: we use the exact formula for the two-scale expansion $\mathcal{H}_{q_1}$ given in (7.114) to compute the value of $\mathcal{L}\mathcal{H}_{q_1}$ and then use the quantitative properties of the corrector stated in Proposition 6.28 to prove that the $\underline{H}^{-1}(A_R,\mu_\beta)$-norm of the term $\mathcal{L}\mathcal{H}_{q_1}$ satisfies the estimate (7.115). Since the proof is rather long due to the technicalities caused by the specific structure of the operator $\mathcal{L}$ (iterations of the Laplacian, sum over all the charges $q \in \mathcal{Q}$), we do not rewrite it but only point out the main differences:

- We work in the annulus A_R and not in the ball $B_{R^{1+\delta}}$, this difference makes the proof simpler since we do not have to take the additional parameter δ into considerations;
- We can always assume that the diameter of the charge q_1 is smaller than $R/2$, otherwise the constant C_{q_1} is larger than R^k for some large number $k := k(d)$ (since it is allowed to have an algebraic growth in the parameter $\|q_1\|_1$) and the estimate (7.115) is trivial in this situation. Under the assumption $\operatorname{diam} q_1 \leq R/2$, we use the identity $-\nabla \cdot \bar{\mathbf{a}}_\beta \nabla \overline{G}_{q_1} = 0$ in the annulus A_R instead of the identity $-\nabla \cdot \bar{\mathbf{a}}_\beta \nabla \overline{G}_\delta = \rho_\delta$ in the ball $B_{R^{1+\delta}}$;
- We use the regularity estimates on the function $\overline{G}_{q_1}$ stated in Proposition 7.8 instead of the estimates on the Green's function $\overline{G}$ stated in Proposition 7.7. Since the map $\overline{G}_{q_1}$ scales like the gradient of the Green's function (in particular it decays like $|x|^{1-d}$), we obtain an additional factor R in the right side of (7.115) compared to (7.29), i.e., we obtain

$$\|\mathcal{L}\mathcal{H}_{q_1}\|_{\underline{H}^{-1}(A_R,\mu_\beta)} \leq \frac{C_{q_1}}{R^{d+\gamma_\alpha}} \quad \text{instead of} \quad \|\mathcal{L}\mathcal{H}_{\delta,\cdot k} - \rho_{\delta,\cdot k}\|_{\underline{H}^{-1}(A_R,\mu_\beta)} \leq \frac{C}{R^{d-1+\gamma_\alpha}}.$$

Step 2. In this step, we use the main result (7.115) of Substep 3.1 to prove that the gradient of the Green's function $\nabla \mathcal{U}_{q_1}$ is close to the gradient of the two-scale expansion $\nabla \mathcal{H}_{q_1}$ in the $\underline{L}^2(A_R, \mu_\beta)$-norm. We prove the estimate

$$\left\| \nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A_R, \mu_\beta)} \leq \frac{C_{q_1}}{R^{d+\gamma_\delta}}. \tag{7.116}$$

To simplify the rest of the argument, we do not prove the estimate (7.116) directly. We slightly reduce the size of the annulus A_R and define the set A_R^1 to be the annulus $A_R^1 := \left\{ x \in \mathbb{Z}^d \, : \, 1.1R \leq |x| \leq 1.9R \right\}$. We note that we have the inclusion, for each radius $R \geq 1$, $A_R^1 \subseteq A_R$. In this substep, we prove the inequality

$$\left\| \nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A_R^1, \mu_\beta)} \leq \frac{C_{q_1}}{R^{d+\gamma_\delta}}. \tag{7.117}$$

The inequality (7.116) can then be deduced from (7.117) by a covering argument.

The argument is similar to the one presented in Section 7.3.3 except that, instead of making use of the mollifier exponent δ to prove that the H^1-norm is of the difference $(\nabla \mathcal{H}_\delta - \nabla \mathcal{G}_\delta)$ is small, as it was done in the estimates (7.87) and (7.88), we use the main result (7.104) of Section 7.4.3. We first let η be a cutoff function which satisfies the properties:

$$0 \leq \eta \leq 1, \ \operatorname{supp} \eta \subseteq A_R, \ \eta = 1 \text{ in } A_R^1, \ \forall k \in \mathbb{N}, \left| \nabla^k \eta \right| \leq \frac{C}{R^k}. \tag{7.118}$$

We then use the function $\eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right)$ as a test function in the definition of the $\underline{H}^{-1}(A_R, \mu_\beta)$-norm of the inequality (7.115) and use the identity $\mathcal{L} \mathcal{U}_{q_1} = 0$ in the set $A_R \times \Omega$. We obtain

$$\begin{aligned} \frac{1}{R^d} \sum_{x \in A_R} \left\langle \eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \mathcal{L} \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right\rangle_{\mu_\beta} & \leq \left\| \mathcal{L} \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right\|_{\underline{H}^{-1}(B_{R^{1+\delta}}, \mu_\beta)} \left\| \eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right\|_{\underline{H}^1(A_R, \mu_\beta)} \\ & \leq \frac{C}{R^{d+\gamma_\alpha}} \left\| \eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right\|_{\underline{H}^1(A_R, \mu_\beta)}. \end{aligned} \tag{7.119}$$

We then estimate the $\underline{H}^1(A_R, \mu_\beta)$-norm of the function $\mathcal{U}_{q_1} - \mathcal{H}_{q_1}$ with similar arguments as the one presented in the proof of the inequality (7.74), the only difference is that we use the regularity estimates stated in Proposition 7.8 instead of the regularity estimates for the functions $\mathcal{G}_\delta$ and $\mathcal{H}$. We obtain

$$\left\| \eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right\|_{\underline{H}^1(A_R, \mu_\beta)} \leq \left\| \eta \mathcal{U}_{q_1} \right\|_{\underline{H}^1(A_R, \mu_\beta)} + \left\| \eta \mathcal{H}_{q_1} \right\|_{\underline{H}^1(A_R, \mu_\beta)} \leq \frac{C_{q_1}}{R^{d-1-\varepsilon}}. \tag{7.120}$$

For later use, we also note that the same argument yields to the inequality

$$\left\| \nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A_R, \mu_\beta)} \leq \left\| \nabla \mathcal{U}_{q_1} \right\|_{\underline{L}^2(A_R, \mu_\beta)} + \left\| \nabla \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A_R, \mu_\beta)} \leq \frac{C_{q_1}}{R^{d-\varepsilon}}. \tag{7.121}$$

We then combine the inequalities (7.117) and (7.119) and use the ordering $\varepsilon \ll \gamma_\alpha$ to deduce that

$$\frac{1}{R^d} \sum_{x \in A_R} \left\langle \eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \mathcal{L} \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right\rangle_{\mu_\beta} \leq \frac{C_{q_1}}{R^{2d+\gamma_\alpha}}. \tag{7.122}$$

Thus to prove the inequality (7.117), it is sufficient to prove the estimate

$$\left\| \nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A_R^1, \mu_\beta)}^2 \leq \frac{1}{R^d} \sum_{x \in A_R} \left\langle \eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \mathcal{L} \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right\rangle_{\mu_\beta} + \frac{C_{q_1}}{R^{2d+\gamma_\delta}}.$$

First, by definition of the Helffer-Sjöstrand operator $\mathcal{L}$, we have the identity

(7.123)

$$\begin{aligned} \sum_{x \in \mathbb{Z}^d} \left\langle \eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \mathcal{L} \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right\rangle_{\mu_\beta} = & \sum_{x, y \in \mathbb{Z}^d} \eta(x) \left\langle \left(\partial_y \mathcal{U}_{q_1}(x, \cdot) - \partial_y \mathcal{H}_{q_1}(x, \cdot) \right)^2 \right\rangle_{\mu_\beta} \\ & + \frac{1}{2\beta} \sum_{x \in \mathbb{Z}^d} \left\langle \left(\nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1} \right)(x, \cdot) \cdot \nabla \left(\eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right)(x, \cdot) \right\rangle_{\mu_\beta} \\ & + \sum_{q \in \mathcal{Q}} \left\langle \nabla_q \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \cdot \mathbf{a}_q \nabla_q \left(\eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right) \right\rangle_{\mu_\beta} \\ & + \frac{1}{2\beta} \sum_{n \geq 1} \sum_{x \in \mathbb{Z}^d} \frac{1}{\beta^{\frac{n}{2}}} \left\langle \nabla^{n+1} \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right)(x, \cdot) \cdot \nabla^{n+1} \left(\eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right) \right)(x, \cdot) \right\rangle_{\mu_\beta}. \end{aligned}$$

We then estimate the four terms on the right side separately. For the first one, we use that it is non-negative

$$\sum_{x,y\in\mathbb{Z}^d} \eta(x)^2 \left\langle \left(\partial_y \mathcal{U}_{q_1}(x,\cdot) - \partial_y \mathcal{H}_{q_1}(x,\cdot)\right)^2 \right\rangle_{\mu_\beta} \geq 0. \tag{7.124}$$

For the second one, we expand the gradient of the product $\eta^2 (\mathcal{U}_{q_1} - \mathcal{H}_{q_1})$ and use the properties of the function η stated in (7.118) to obtain

$$\begin{aligned} R^{-d} \sum_{x\in\mathbb{Z}^d} \left\langle \left(\nabla \mathcal{U}_{q_1}(x,\cdot) - \nabla \mathcal{H}_{q_1}\right)(x,\cdot) \cdot \nabla \left(\eta \left(\mathcal{U}_{q_1} - \mathcal{H}\right)\right)(x,\cdot) \right\rangle_{\mu_\beta} &\geq c \left\| \eta \left(\nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1}\right) \right\|^2_{\underline{L}^2(A_R,\mu_\beta)} \\ &\quad - \frac{C}{R} \left\| \nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A_R,\mu_\beta)} \left\| \mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A_R,\mu_\beta)}. \end{aligned} \tag{7.125}$$

We then use the inequality (7.104) and the estimate (7.121) and the quantitative sublinearity of the corrector to deduce that

$$\frac{1}{R} \left\| \nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A_R,\mu_\beta)} \left\| \mathcal{U}_{q_1} - \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{1}{R} \cdot \frac{C}{R^{d-\varepsilon}} \cdot \frac{C}{R^{d-1+\gamma_\delta}} \leq \frac{C}{R^{2d+\gamma_\delta}}. \tag{7.126}$$

We then combine the inequalities (7.125) and (7.126) to deduce that

$$R^{-d} \sum_{x\in\mathbb{Z}^d} \left\langle \left(\nabla \mathcal{U}_{q_1}(x,\cdot) - \nabla \mathcal{H}_{q_1}\right)(x,\cdot) \cdot \nabla \left(\eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1}\right)\right)(x,\cdot) \right\rangle_{\mu_\beta} + \frac{C_{q_1}}{R^{2d+\gamma_\delta}} \geq c \left\| \eta \left(\nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1}\right) \right\|^2_{\underline{L}^2(A_R,\mu_\beta)}. \tag{7.127}$$

The two remaining terms in the right side of the estimate (7.123) (involving the iteration of the Laplacian and the sum over the charges) are estimated following the ideas developed in Section 7.3.3 (see (7.90) and (7.91)). We skip the details and write the result:

$$R^{-d} \sum_{q\in\mathcal{Q}} \left\langle \nabla_q \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1}\right) \cdot \mathbf{a}_q \nabla_q \left(\eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1}\right)\right) \right\rangle_{\mu_\beta} + \frac{C_{q_1}}{R^{(2d+\gamma_\delta)}} \geq -Ce^{-c\sqrt{\beta}} \left\| \eta \left(\nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1}\right) \right\|^2_{\underline{L}^2(A_R,\mu_\beta)} \tag{7.128}$$

and

$$R^{-d} \sum_{n\geq 1} \sum_{x\in\mathbb{Z}^d} \frac{1}{\beta^{\frac{n}{2}}} \left\langle \nabla^{n+1} \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1}\right)(x,\cdot) \cdot \nabla^{n+1} \left(\eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1}\right)\right)(x,\cdot) \right\rangle_{\mu_\beta} + \frac{C_{q_1}}{R^{(2d+\gamma_\delta)}} \geq 0. \tag{7.129}$$

We then combine the estimates (7.124), (7.127), (7.128) and (7.129) with the identity (7.123), choose the inverse temperature β large enough so that the right side of (7.128) can be absorbed by the right side of (7.125) and use that the cutoff function η is equal to 1 in the annulus A^1_R. We obtain

$$\left\| \nabla \mathcal{U}_{q_1} - \nabla \mathcal{H}_{q_1} \right\|_{\underline{L}^2(A^1_R,\mu_\beta)} \leq \frac{C}{R^d} \sum_{x\in\mathbb{Z}^d} \left\langle \eta \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1}\right) \mathcal{L} \left(\mathcal{U}_{q_1} - \mathcal{H}_{q_1}\right) \right\rangle_{\mu_\beta} + \frac{C_{q_1}}{R^{d+\gamma_\delta}}. \tag{7.130}$$

We then combine the inequality (7.130) with the estimate (7.122) to complete the proof of (7.117). Step 2 is complete.

Step 3. The conclusion. In this step, we prove the L^2-estimate

$$\left\| \nabla \mathcal{U}_{q_1} - \sum_{i,j} \left(e_{ij} + \nabla \chi_{ij}\right) \nabla \overline{G}_{q_1,j} \right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{C_{q_1}}{R^{d+\gamma_\delta}}. \tag{7.131}$$

In view of the estimate (7.117) proved in Step 2, it is sufficient to prove the inequality

$$\left\| \nabla \mathcal{H}_{q_1} - \sum_{i,j} \left(e_{ij} + \nabla \chi_{ij}\right) \nabla \overline{G}_{q_1,j} \right\|_{\underline{L}^2(A_R,\mu_\beta)} \leq \frac{C_{q_1}}{R^{d+\gamma_\delta}}. \tag{7.132}$$

The proof of (7.132) relies on the regularity estimate on the function $\overline{G}_{q_1}$ stated in Proposition 7.8, the quantitative sublinearity of the corrector stated in Proposition 6.28, and the quantitative estimate for the difference of the finite and infinite-volume gradient of the corrector stated in Proposition 6.29. The argument is identical (and even simpler since we do not have to take into account the parameter δ) to the argument given in Section 7.3.4 so we skip the details. The proof of Step 3, and thus of Theorem 2, is complete.

8. First-order expansion of the two-point function: Technical lemmas

In this section, we present the proofs of the technical lemmas which are used in Section 4 to prove Theorem 1. All the tools used in this section have been introduced in Section 3 except one: The second-order Helffer-Sjöstrand equation introduced in Section 5.4.

Most of the heuristic of the arguments are presented in Section 4 and we refer to it for an overview of the results. As it may be useful to the reader, we record below the tools established in this article which are used in the proofs below:

- In Sections 8.1, 8.2 and 8.3, we study the correlation of random variables; this is achieved by using the Helffer-Sjöstrand representation formula. We need to use the properties of the Green's matrix associated with the Helffer-Sjöstrand operator stated in Proposition 3.17;
- In Section 8.3, we need to study the correlation between a solution of a Helffer-Sjöstrand equation and the random variables X_x and Y_0. To this end, we appeal Helffer-Sjöstrand representation formula and the second-order Helffer-Sjöstrand equation as well as to the properties of the Green's matrix associated with this operator stated in Proposition 5.13;
- Sections 8.4 and 8.5 are devoted to the proofs of some properties of the discrete Green's function on the lattice $\mathbb{Z}^d$; they can be read independently of the rest of the article.

8.1. **Removing the terms $X_{\sin\cos}$, $X_{\cos\cos}$ and $X_{\sin\sin}$.** We recall the definitions of the values $Z_\beta(\sigma)$ and $Z_\beta(0)$ introduced in (3.12), the definitions of the random variables Y_0, X_x, $X_{\sin\cos}$, $X_{\cos\cos}$, $X_{\sin\sin}$ introduced in (4.4) and the identity

$$\frac{Z_\beta(\sigma)}{Z_\beta(0)} = \left\langle Y_0 X_x X_{\sin\cos} X_{\cos\cos} X_{\sin\sin} \right\rangle_{\mu_\beta}. \tag{8.1}$$

Proof of Lemma 4.1. As is explained in Section 4.2, the proof of the lemma is based on the proof of the following estimates

$$\begin{cases} \left\| X_{\sin\cos} - 1 \right\|_{L^\infty} \leq \dfrac{C}{|x|^{d-1}}, \\ \left\| X_{\cos\cos} - 1 \right\|_{L^\infty} \leq \dfrac{C}{|x|^{d-1}}, \\ \operatorname{var}_{\mu_\beta} X_{\sin\sin} \leq \dfrac{C}{|x|^{2d-2}}, \\ \mathbb{E}\left[X_{\sin\sin} \right] = 1 + \dfrac{c}{|x|^{d-2}} + O\left(\dfrac{C}{|x|^{d-1}} \right). \end{cases} \tag{8.2}$$

The fact that (8.2) implies (4.6) is straightforward, and we refer to the long version of this article ([36, Chapter 8, Section 1]) for the details. To prove (8.2), we first focus on the first two inequalities involving the random variables $X_{\sin\cos}$ and $X_{\cos\cos}$. They can be obtained thanks to the following ingredients:

- For each point $y \in \mathbb{Z}^d$ and each charge $q \in \mathcal{Q}_y$, we have the estimate

$$\left| (\nabla G, n_q) \right| \leq \left\| \nabla G \right\|_{L^2(\operatorname{supp} n_q)} \left\| n_q \right\|_2 \leq \frac{C_q}{|y|^{d-1}},$$

  A similar computation shows the estimate $(\nabla G_x, n_q) \leq C_q |y-x|^{1-d}$;
- The standard estimates, for each real number $a \in \mathbb{R}$, $|\sin a| \leq |a|$, $|\cos a - 1| \leq \frac{1}{2}|a|^2$ and the estimate, for each charge $q \in \mathcal{Q}$, $|z(\beta, q)| \leq e^{-c\sqrt{\beta}\|q\|_1}$.

We obtain the inequality

$$\begin{aligned} \left| \sum_{q \in \mathcal{Q}} z(\beta, q) \sin(2\pi(\phi, q)) \sin(2\pi(\nabla G_x, n_q)) \left(\cos(2\pi(\nabla G, n_q)) - 1 \right) \right| & \leq C \sum_{y \in \mathbb{Z}^d} \sum_{q \in \mathcal{Q}_y} \frac{e^{-c\sqrt{\beta}\|q\|_1} C_q}{|y-x|^{d-1}} \frac{1}{|y|^{2d-2}} \\ & \leq C \sum_{y \in \mathbb{Z}^d} \frac{1}{|y-x|^{d-1}} \frac{1}{|y|^{2d-2}} \\ & \leq \frac{C}{|x|^{d-1}}, \end{aligned} \tag{8.3}$$

where we used the exponential decay of the term $e^{-c\sqrt{\beta}\|q\|_1}$ to absorb the algebraic growth of the constant C_q. With a similar strategy, we obtain the two inequalities

$$\begin{aligned} \left| \sum_{q \in \mathcal{Q}} z(\beta, q) \sin(2\pi(\phi, q)) \sin(2\pi(\nabla G, n_q)) (\cos(2\pi(\nabla G_x, q)) - 1) \right| &\leq \frac{C}{|x|^{d-1}}, \\ \left| \sum_{q \in \mathcal{Q}} z(\beta, q) \sin(2\pi(\phi, q)) \frac{1}{2} (\cos(2\pi(\nabla G_x, q)) - 1) (\cos(2\pi(\nabla G, q)) - 1) \right| &\leq \frac{C}{|x|^{d-1}}. \end{aligned} \tag{8.4}$$

We then combine the estimates (8.3) and (8.4) and use that the exponential function is Lipschitz on any bounded intervals of $\mathbb{R}$ to obtain, for each realization of the field $\phi \in \Omega$,

$$|X_{\sin\cos}(\phi) - 1| \leq \frac{C}{|x|^{d-1}} \quad \text{and} \quad |X_{\cos\cos}(\phi) - 1| \leq \frac{C}{|x|^{d-1}}.$$

This result implies the $L^\infty(\mu_\beta)$-estimates stated in (8.2).

There remains to prove the estimates corresponding to the variance and the expectation of the random variable $X_{\sin\sin}$ in (8.2). We first note that a computation similar to the one performed in (8.3) gives the following $L^\infty(\mu_\beta)$-estimate: for each realization of the field $\phi \in \Omega$,

$$\left| \sum_{q \in \mathcal{Q}} z(\beta, q) \cos(2\pi(\phi, q)) \sin(2\pi(\nabla G, n_q)) \sin(2\pi(\nabla G_x, n_q)) \right| \leq C \sum_{y \in \mathbb{Z}^d} \frac{1}{|y - x|^{d-1}} \frac{1}{|y|^{d-1}} \leq \frac{C}{|x|^{d-2}}. \tag{8.5}$$

By the estimate (8.5) and the Taylor expansion of the exponential, we obtain the bound

$$\begin{aligned} &\left| X_{\sin\sin} - 1 - \sum_{q \in \mathcal{Q}} z(\beta, q) \cos(2\pi(\phi, q)) \sin(2\pi(\nabla G, n_q)) \sin(2\pi(\nabla G_x, n_q)) \right| \\ &\quad \leq C \left(\sum_{q \in \mathcal{Q}} z(\beta, q) \cos(2\pi(\phi, q)) \sin(2\pi(\nabla G, n_q)) \sin(2\pi(\nabla G_x, n_q)) \right)^2 \\ &\quad \leq \frac{C}{|x|^{2d-4}}. \end{aligned}$$

Since the dimension d is assumed to be larger than 3, we have the inequality $2d - 4 \geq d - 1$. We deduce that to prove the estimates pertaining to the random variable $X_{\sin\sin}$ in (8.2), it is sufficient to prove the inequality

$$\operatorname{var}\left[\sum_{q \in \mathcal{Q}} z(\beta, q) \cos(2\pi(\phi, q)) \sin(2\pi(\nabla G, n_q)) \sin(2\pi(\nabla G_x, n_q)) \right] \leq \frac{C}{|x|^{2d-2}} \tag{8.6}$$

and the expansion

$$\mathbb{E}\left[\sum_{q \in \mathcal{Q}} z(\beta, q) \cos(2\pi(\phi, q)) \sin(2\pi(\nabla G, n_q)) \sin(2\pi(\nabla G_x, n_q)) \right] = \frac{c}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-1}} \right). \tag{8.7}$$

The estimate (8.6) involving the variance can be estimated by the Helffer-Sjöstrand representation formula and the bounds on the Green's matrix $\mathcal{G}$ stated in Proposition 3.17. We first note that, for each point $y \in \mathbb{Z}^d$,

$$\begin{aligned} &\partial_y \left(\sum_{q \in \mathcal{Q}} z(\beta, q) \cos(2\pi(\phi, q)) \sin(2\pi(\nabla G, n_q)) \sin(2\pi(\nabla G_x, n_q)) \right) \\ &\qquad = -\sum_{q \in \mathcal{Q}} 2\pi z(\beta, q) \sin(2\pi(\phi, q)) \sin(2\pi(\nabla G, n_q)) \sin(2\pi(\nabla G_x, n_q))\, q(y). \end{aligned} \tag{8.8}$$

From the identity (8.8), we deduce that to compute the variance (8.6), one needs to solve the Helffer-Sjöstrand equation

$$\mathcal{L}\mathcal{W}(y, \phi) = -\sum_{q \in \mathcal{Q}} z(\beta, q) \sin(2\pi(\phi, q)) \sin(2\pi(\nabla G, q)) \sin(2\pi(\nabla G_x, q))\, q(y). \tag{8.9}$$

The equation (8.9) can be solved explicitly by using the Green's matrix associated with the Helffer-Sjöstrand operator; we obtain the following formula for the codifferential of $\mathcal{W}$

$$\mathrm{d}^*\mathcal{W}(y, \phi) = -\sum_{q \in \mathcal{Q}} z(\beta, q) \sin(2\pi(\nabla G, n_q)) \sin(2\pi(\nabla G_x, n_q)) \sum_{z \in \operatorname{supp} n_q} \mathrm{d}_y^* \mathrm{d}_z^* \mathcal{G}_{\sin(2\pi(\phi, q))}(y, \phi; z)\, n_q(z).$$

Using the estimate on the Helffer-Sjöstrand Green's matrix proved in Proposition 3.17, and the fact that the codifferential d^* is a linear functional of the gradient, we deduce the estimate, for each point $y \in \mathbb{Z}^d$,

$$\begin{aligned} \|\mathrm{d}^* \mathcal{W}(y,\cdot)\|_{L^\infty(\mu_\beta)} &\leq \sum_{z\in\mathbb{Z}^d} \sum_{q \in \mathcal{Q}_z} \frac{e^{-c\sqrt{\beta}\|q\|_1} C_q}{|z|^{d-1}|z-x|^{d-1}} \frac{1}{|y-z|^{d-\varepsilon}} \\ &\leq \sum_{z\in\mathbb{Z}^d} \frac{C}{|z|^{d-1}|z-x|^{d-1}} \frac{1}{|y-z|^{d-\varepsilon}}. \end{aligned} \tag{8.10}$$

Using the definition of the map $\mathcal{W}$, we apply the Helffer-Sjöstrand representation formula and deduce that

$$\begin{aligned} &\mathrm{var}\left[\sum_{q\in\mathcal{Q}} z(\beta,q) \cos\left(2\pi(\phi,q)\right) \sin\left(2\pi(\nabla G, n_q)\right) \sin\left(2\pi(\nabla G_x, n_q)\right) \right] \\ &\qquad = 4\pi^2 \sum_{y\in\mathbb{Z}^d} \left\langle \left(\sum_{q\in\mathcal{Q}} z(\beta,q) \sin\left(2\pi(\phi,q)\right) \sin\left(2\pi(\nabla G, n_q)\right) \sin\left(2\pi(\nabla G_x, q) n_q(y)\right) \right) \mathrm{d}^* \mathcal{W}(y,\phi) \right\rangle_{\mu_\beta}. \end{aligned}$$

Using the estimates (8.10) and a computation similar to the one performed in (8.3), we deduce that

$$\begin{aligned} &\mathrm{var}\left[\sum_{q\in\mathcal{Q}} z(\beta,q) \cos\left(2\pi(\phi,q)\right) \sin\left(2\pi(\nabla G, n_q)\right) \sin\left(2\pi(\nabla G_x, n_q)\right) \right] \\ &\qquad \leq \sum_{y\in\mathbb{Z}^d} \sum_{q\in\mathcal{Q}_y} \frac{e^{-c\sqrt{\beta}\|q\|_1} C_q}{|y|^{d-1}|x-y|^{d-1}} \|\mathrm{d}^*\mathcal{W}(y,\cdot)\|_{L^\infty(\mu_\beta)} \\ &\qquad \leq C \sum_{y,z\in\mathbb{Z}^d} \frac{1}{|y|^{d-1}|x-y|^{d-1}} \times \frac{1}{|z|^{d-1}|z-x|^{d-1}} \times \frac{1}{|y-z|^{d-\varepsilon}} \\ &\qquad \leq \frac{C}{|x|^{2d-2}}, \end{aligned} \tag{8.11}$$

where we used the results stated in Appendix C in the last line. There only remains to prove the identity (8.7). To this end, we use the ideas and notation presented in Section 4.5.2 and decompose the sum

$$\begin{aligned} &\sum_{q\in\mathcal{Q}} z(\beta,q) \cos\left(2\pi(\phi,q)\right) \sin\left(2\pi(\nabla G, n_q)\right) \sin\left(2\pi(\nabla G_x, n_q)\right) \\ &\qquad = \sum_{[q]\in\mathcal{Q}/\mathbb{Z}^d} z(\beta,q) \sum_{y\in\mathbb{Z}^d} \cos\left(2\pi(\phi, q(y+\cdot))\right) \sin\left(2\pi(\nabla G, n_q(y+\cdot))\right) \sin\left(2\pi(\nabla G_x, n_q(y+\cdot))\right). \end{aligned}$$

Taking the expectation, using the translation invariance of the measure μ_β, we deduce that

$$\begin{aligned} &\mathbb{E}\left[\sum_{q\in\mathcal{Q}} z(\beta,q) \cos\left(2\pi(\phi,q)\right) \sin\left(2\pi(\nabla G, n_q)\right) \sin\left(2\pi(\nabla G_x, n_q)\right) \right] \\ &\qquad = \sum_{[q]\in\mathcal{Q}/\mathbb{Z}^d} z(\beta,q) \mathbb{E}\left[\cos\left(2\pi(\phi,q)\right)\right] \sum_{y\in\mathbb{Z}^d} \sin\left(2\pi(\nabla G(\cdot - y), n_q)\right) \sin\left(2\pi(\nabla G_x(\cdot - y), n_q)\right). \end{aligned} \tag{8.12}$$

Fix an equivalence class $[q] \in \mathcal{Q}/\mathbb{Z}^d$. By using a Taylor expansion of the sine and standard properties of the discrete Green's function G, we obtain the expansion

$$\begin{aligned} \sum_{y\in\mathbb{Z}^d} \sin\left(2\pi(\nabla G(\cdot - y), n_q)\right) \sin\left(2\pi(\nabla G_x(\cdot - y), n_q)\right) &= 4\pi^2 \sum_{y\in\mathbb{Z}^d} \nabla G(y)\cdot(n_q) \times \nabla G_x(y)\cdot(n_q) + O\left(\frac{C_q}{|x|^{d-1}}\right) \\ &= 4\pi^2 \sum_{i,j=1}^d (n_q)_i (n_q)_j \sum_{y\in\mathbb{Z}^d} \nabla_i G(y) \nabla_j G_x(y) + O\left(\frac{C_q}{|x|^{d-1}}\right). \end{aligned} \tag{8.13}$$

Putting this estimate back into (8.12), we deduce that

$$\mathbb{E}\left[\sum_{q\in\mathcal{Q}} z(\beta,q) \cos\left(2\pi(\phi,q)\right) \sin\left(2\pi(\nabla G, n_q)\right) \sin\left(2\pi(\nabla G_x, n_q)\right) \right] = \sum_{i,j=1}^d c_{ij} \sum_{y\in\mathbb{Z}^d} \nabla_i G(y) \nabla_j G_x(y) + O\left(\frac{C}{|x|^{d-1}}\right). \tag{8.14}$$

where the constants c_{ij} are defined by the formulae

$$c_{ij} = 4\pi^2 \sum_{[q] \in \mathcal{Q}/\mathbb{Z}^d} z(\beta, q) \mathbb{E}\left[\cos\left(2\pi(\phi, q)\right)\right] (n_q)_i (n_q)_j .$$

The expansion (8.14) is not exactly (8.7). To complete the argument, we appeal to the symmetry invariance of the model and claim that it implies the identities $c_{ij} = 0$ if $i \neq j$ and $c_{ii} = c_{jj}$ for each pair $(i,j) \in \{1, \ldots, d\}^2$. The proof follows from standard symmetry arguments and we omit it here. Once this result is established, the expansion (8.7) is obtained from (8.14) thanks to an integration by parts and the properties of the discrete Green's function. □

8.2. **Removing the contributions of the cosines.** The goal of this section is to prove Lemma 4.3.

Proof of Lemma 4.3. We start from the Helffer-Sjöstrand representation formula stated in (4.13) and recalled below

$$\operatorname{cov}\left[X_x, Y_0\right] = \sum_{y \in \mathbb{Z}^d} \left\langle \left(\partial_y X_x\right) \mathcal{Y}(y, \cdot) \right\rangle_{\mu_\beta}, \tag{8.15}$$

where $\mathcal{Y} : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ is the solution of the Helffer-Sjöstrand equation, for each pair $(y, \phi) \in \mathbb{Z}^d \times \Omega$,

$$\mathcal{L}\mathcal{Y}(y, \phi) = \partial_y Y_0(\phi). \tag{8.16}$$

Using the definition of the random variables Y_0 and X_x stated in (4.4), we have the identities, for each $y \in \mathbb{Z}^d$,

$$\partial_y Y_0(\phi) = -\left(Q_0(y, \phi) + \frac{1}{2} 2\pi \sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\phi, q)\right) \left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right) q(y) \right) Y_0(\phi) \tag{8.17}$$

and

$$\partial_y X_x(\phi) = -\left(Q_x(y, \phi) + \sum_{q \in \mathcal{Q}} \frac{1}{2} 2\pi z(\beta, q) \sin\left(2\pi(\phi, q)\right) \left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) q(y) \right) X_x(\phi). \tag{8.18}$$

The objective of the proof is to remove the terms involving the cosine in the right sides of the identities (8.17) and (8.18). The proof requires to use the following estimates established in (4.19) and (4.21): for each point $y \in \mathbb{Z}^d$,
(8.19)

$$\left\| n_{Q_x}(y, \cdot) \right\|_{L^\infty(\mu_\beta)} \leq \frac{C}{|y - x|^{d-1}} \quad \text{and} \quad \left| \sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\phi, q)\right) \left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) n_q(y) \right| \leq \frac{C}{|y - x|^{2d-2}},$$

as well as the estimates

$$\left\| n_{Q_0}(y, \cdot) \right\|_{L^\infty(\mu_\beta)} \leq \frac{C}{|y|^{d-1}}, \quad \text{and} \quad \left| \sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\phi, q)\right) \left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right) n_q(y) \right| \leq \frac{C}{|y|^{2d-2}}. \tag{8.20}$$

We split the argument into three steps:

- In Step 1, we prove that the solution of the Helffer-Sjöstrand equation $\mathcal{Y}$ satisfies the upper bound, for each $y \in \mathbb{Z}^d$,

$$\left\| \mathrm{d}^* \mathcal{Y}(y, \cdot) \right\|_{L^2(\mu_\beta)} \leq \frac{C}{|y|^{d-1-\varepsilon}}; \tag{8.21}$$

- In Step 2, we prove that the covariance between the random variables X_x and Y_0 satisfies the expansion

$$\operatorname{cov}\left[X_x, Y_0\right] = \sum_{y \in \mathbb{Z}^d} \left\langle X_x Q_x(y, \cdot) \mathcal{Y}(y, \cdot) \right\rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1-\varepsilon}} \right); \tag{8.22}$$

- In Step 3, we use the symmetry of the Helffer-Sjöstrand operator $\mathcal{L}$ to complete the proof of Lemma 4.3.

Step 1. We first express the function $\mathcal{Y}$ in terms of the Green function associated with the Helffer-Sjöstrand operator $\mathcal{L}$. From the equation (8.16), we deduce the formula for the codifferential of the map $\mathcal{Y}$, for each pair $(y,\phi) \in \mathbb{Z}^d \times \Omega$,

$$\mathrm{d}^* \mathcal{Y}(y,\phi) = 2\pi \sum_{y_1 \in \mathbb{Z}^d} \sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\nabla G, n_q)\right) \mathrm{d}_y^* \mathrm{d}_{y_1}^* \mathcal{G}_{\cos(2\pi(\cdot,q))Y_0}\left(y,\phi; y_1\right) n_q(y_1) \\ + 2\pi \sum_{y_1 \in \mathbb{Z}^d} \sum_{q \in \mathcal{Q}} \frac{1}{2} \left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right) \mathrm{d}_y^* \mathrm{d}_{y_1}^* \mathcal{G}_{\sin(2\pi(\cdot,q))Y_0}(y,\phi; y_1) n_q(y_1).$$

Using the estimate on the Helffer-Sjöstrand Green's matrix proved in Proposition 3.17, that the random variable Y_0 belongs to the space $L^2(\mu_\beta)$, and the Taylor expansions of the sine and cosine, we obtain the inequality

$$\begin{aligned} \left\| \mathrm{d}^* \mathcal{Y}(y,\cdot)\right\|_{L^2(\mu_\beta)} &\leq \sum_{y_1 \in \mathbb{Z}^d} \frac{C}{|y_1|^{d-1}} \left\| \mathrm{d}_y^* \mathrm{d}_{y_1}^* \mathcal{G}_{\cos(2\pi(\cdot,q))Y_0}\left(y,\phi;y_1\right)\right\|_{L^2(\mu_\beta)} \\ &\quad + \sum_{y_1 \in \mathbb{Z}^d} \frac{C}{|y_1|^{2d-2}} \left\| \mathrm{d}_y^* \mathrm{d}_{y_1}^* \mathcal{G}_{\sin 2\pi(\cdot,q)Y_0}\left(y,\phi;y_1\right)\right\|_{L^2(\mu_\beta)} \\ &\leq \sum_{y_1 \in \mathbb{Z}^d} \frac{C}{|y_1|^{d-1}|y-y_1|^{d-\varepsilon}} + \frac{C}{|y_1|^{2d-2}|y-y_1|^{d-\varepsilon}} \\ &\leq \frac{C}{|y|^{d-1-\varepsilon}}. \end{aligned}$$

The proof of Step 1 is complete.

Step 2. By the Helffer-Sjöstrand formula (8.15), we have the identity

$$\begin{aligned} \operatorname{cov}\left[X_x, Y_0\right] &= \sum_{y \in \mathbb{Z}^d} \left\langle \left(\partial_y X_x\right) \mathcal{Y}(y,\cdot)\right\rangle_{\mu_\beta} \\ &= \sum_{y \in \mathbb{Z}^d} \left\langle Q_x(y) X_x \mathcal{Y}(y,\cdot)\right\rangle_{\mu_\beta} - \pi \left\langle \sum_{q \in \mathcal{Q}} z\left(\beta,q\right) \sin\left(2\pi(\phi,q)\right) \left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) n_q(y) X_x \mathrm{d}^* \mathcal{Y}(y,\cdot)\right\rangle_{\mu_\beta}. \end{aligned} \tag{8.23}$$

The objective of this step is to prove that the term involving the cosine in the right side of (8.23) is of lower order; specifically, we prove the estimate (8.24) below. The proof relies on the three following ingredients: the Taylor expansion of the cosine, the $L^2(\mu_\beta)$-estimate $\|X_x\|_{L^2(\mu_\beta)} \leq C$, and the estimate (8.21) proved in Step 1. We obtain

$$\begin{aligned} &\left| \frac{1}{2} \sum_{q \in \mathcal{Q}} z\left(\beta,q\right) \left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) n_q(y) \left\langle \sin\left(2\pi(\phi,q)\right) X_x \mathrm{d}^* \mathcal{Y}(y,\cdot)\right\rangle_{\mu_\beta}\right| \\ &\quad \leq \frac{1}{2} \sum_{q \in \mathcal{Q}} \left| z\left(\beta,q\right) \left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) n_q(y)\right| \left\| \sin\left(2\pi(\phi,q)\right) X_x\right\|_{L^2(\mu_\beta)} \left\| \mathrm{d}^* \mathcal{Y}(y,\cdot)\right\|_{L^2(\mu_\beta)} \\ &\quad \leq \frac{C}{|y-x|^{2d-2}} \cdot \frac{1}{|y|^{d-1-\varepsilon}}. \end{aligned}$$

Summing the inequality over all the points $y \in \mathbb{Z}^d$ and using the results of Appendix C then shows

$$\left| \sum_{y \in \mathbb{Z}^d} \frac{1}{2} \left\langle \sum_{q \in \mathcal{Q}} z\left(\beta,q\right) \sin\left(2\pi(\phi,q)\right) \left(\cos\left(2\pi(\nabla G_x, n_q)\right) - 1\right) n_q(y) X_x \mathrm{d}^* \mathcal{Y}(y,\cdot)\right\rangle_{\mu_\beta}\right| \leq \frac{C}{|x|^{d-1-\varepsilon}}. \tag{8.24}$$

Step 3. The conclusion. We use the main result (8.22) of Step 2 and the symmetry of the Helffer-Sjöstrand operator to complete the proof of Lemma 4.3. By the expansion (8.22), we see that it is sufficient to prove the estimate

$$\sum_{y \in \mathbb{Z}^d} \left\langle Q_x(y,\cdot) X_x \mathcal{V}(y,\cdot)\right\rangle_{\mu_\beta} = \sum_{y \in \mathbb{Z}^d} \left\langle Q_x(y,\cdot) X_x \mathcal{Y}(y,\cdot)\right\rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1-\varepsilon}}\right). \tag{8.25}$$

By the symmetry of the Helffer-Sjöstrand operator, we can write

$$\sum_{y \in \mathbb{Z}^d} \langle Q_x(y,\cdot) X_x \mathcal{Y}(y,\cdot) \rangle_{\mu_\beta} = \sum_{y \in \mathbb{Z}^d} \langle \mathcal{X}_x(y,\cdot) \partial_y Y_0 \rangle_{\mu_\beta}, \tag{8.26}$$

where the mapping $\mathcal{X}_x : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ is the solution of the Helffer-Sjöstrand equation,

$$\mathcal{L}\mathcal{X}_x = Q_x X_x \text{ in } \mathbb{Z}^d \times \Omega. \tag{8.27}$$

The objective of this step is thus to prove the following expansion

$$\sum_{y \in \mathbb{Z}^d} \langle \mathcal{X}_x(y,\cdot) \partial_y Y_0 \rangle_{\mu_\beta} = \sum_{y \in \mathbb{Z}^d} \langle \mathcal{X}_x(y,\cdot) Q_0(y,\cdot) Y_0 \rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1-\varepsilon}}\right). \tag{8.28}$$

The proof is similar to the one written in Steps 1 and 2. With the same arguments as the ones developed in Step 1, one obtains the following upper bound for the function $\mathrm{d}^*\mathcal{X}_x$: for each $y \in \mathbb{Z}^d$,

$$\|\mathrm{d}^*\mathcal{X}_x(y,\cdot)\|_{L^2(\mu_\beta)} \leq \frac{C}{|y-x|^{d-1-\varepsilon}}. \tag{8.29}$$

Using the same arguments as the ones developed in Step 2, we obtain the inequality

$$\left| \sum_{y \in \mathbb{Z}^d} \sum_{q \in \mathcal{Q}} z(\beta,q) \left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right) n_q(y,\phi) \langle \mathrm{d}^*\mathcal{X}_x(y,\phi) \sin\left(2\pi(\phi,q)\right) Y_0(\phi) \rangle_{\mu_\beta} \right| \leq \frac{C}{|x|^{d-1-\varepsilon}}. \tag{8.30}$$

Combining the inequalities (8.29) and (8.30) with the formula (8.17) implies the expansion (8.28). We then use the symmetry of the Helffer-Sjöstrand operator a second time to obtain the identity

$$\sum_{y \in \mathbb{Z}^d} \langle \mathcal{X}_x(y,\cdot) Q_0(y,\cdot) Y_0 \rangle_{\mu_\beta} = \sum_{y \in \mathbb{Z}^d} \langle Q_x(y,\cdot) X_x \mathcal{V}(y,\cdot) \rangle_{\mu_\beta}, \tag{8.31}$$

where the function $\mathcal{V}$ is defined as the solution of the Helffer-Sjöstrand equation (4.23). Combining the identities (8.28), (8.26) and (8.31), we obtain the expansion (8.25). This completes the proof of Step 3 and of Lemma 4.3. □

8.3. **Decoupling the exponentials.** The objective of this section is to remove the exponential terms X_x and Y_0 from the computation. We prove the decorrelation estimate stated in Lemma 4.4. The argument makes use of the bounds on the Green's matrix $\mathcal{G}$ obtained in Proposition 3.17 and on the Green's matrix $\mathcal{G}_{\mathrm{sec},\mathbf{f}}$ associated with the second-order Helffer-Sjöstrand operator proved in Proposition 5.13. Before stating the lemma, we record two estimates which are used in its proof:

- We recall the definition of the random variable $\mathcal{X}_x : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ defined in (8.27) as the solution of the Helffer-Sjöstrand equation, for each $(z,\phi) \in \mathbb{Z}^d \times \Omega$, $\mathcal{L}\mathcal{X}_x(z,\phi) = \partial_z X_x$; by the inequality (8.29), it satisfies the $L^2(\mu_\beta)$-estimates

$$\|\mathcal{X}_x(z,\cdot)\|_{L^2(\mu_\beta)} \leq \frac{C}{|z-x|^{d-2-\varepsilon}}, \quad \text{and} \quad \|\mathrm{d}^*\mathcal{X}_x(z,\cdot)\|_{L^2(\mu_\beta)} \leq \frac{C}{|z-x|^{d-1-\varepsilon}}. \tag{8.32}$$

- The function $\mathcal{V}$ defined in the statement of Lemma 4.3; by the estimate (4.33), it satisfies the estimate

$$\|\mathrm{d}^*\mathcal{V}(z,\cdot)\|_{L^2(\mu_\beta)} \leq \frac{C}{|x|^{d-1-\varepsilon}}. \tag{8.33}$$

Proof of Lemma 4.4. We recall the notation and results introduced in Remarks 4.5, 4.6 and 4.7 which will be used in the proof. We start from the result of Lemma 4.3 which reads

$$\operatorname{cov}[X_x, Y_0] = \sum_{y \in \mathbb{Z}^d} \langle X_x Q_x(y,\cdot) \mathcal{V}(y,\cdot) \rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1-\varepsilon}}\right),$$

where $\mathcal{V}$ is the solution of the Helffer-Sjöstrand equation, for each $(y,\phi) \in \mathbb{Z}^d \times \Omega$,

$$\mathcal{L}\mathcal{V}(y,\phi) = Q_0(y,\phi) Y_0(\phi). \tag{8.34}$$

We split the argument into two steps:

- In Step 1, we prove the decorrelation estimate

$$\sum_{y\in\mathbb{Z}^d}\left\langle X_x Q_x(y,\cdot)\mathcal{V}(y,\cdot)\right\rangle_{\mu_\beta} = \left\langle X_x\right\rangle_{\mu_\beta}\sum_{y\in\mathbb{Z}^d}\left\langle Q_x(y,\cdot)\mathcal{V}(y,\cdot)\right\rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1-\varepsilon}}\right). \tag{8.35}$$

  Let us note that since the measure μ_β is invariant under translations, the value $\left\langle X_x\right\rangle_{\mu_\beta}$ does not depend on the point x.
- In Step 2, we prove the expansion

$$\sum_{y\in\mathbb{Z}^d}\left\langle Q_x(y,\cdot)\mathcal{V}(y,\cdot)\right\rangle_{\mu_\beta} = \left\langle Y_0\right\rangle_{\mu_\beta}\sum_{y\in\mathbb{Z}^d}\left\langle Q_x(y,\cdot)\mathcal{U}(y,\cdot)\right\rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1-\varepsilon}}\right). \tag{8.36}$$

Lemma 4.4 is a consequence of (8.35) and (8.36).

Step 1. The expansion (8.35) can be rewritten in terms of the covariance between the random variables X_x and $Q_x(y)\mathcal{V}(y,\cdot)$; it is equivalent to the expansion

$$\sum_{y\in\mathbb{Z}^d}\operatorname{cov}\left[X_x, Q_x(y,\cdot)\mathcal{V}(y,\cdot)\right] = O\left(\frac{C}{|x|^{d-1-\varepsilon}}\right). \tag{8.37}$$

To prove the expansion (8.37), we apply the Helffer-Sjöstrand representation formula which reads, for each point $y\in\mathbb{Z}^d$,

$$\operatorname{cov}\left[X_x, Q_x(y,\cdot)\mathcal{V}(y,\cdot)\right] = \sum_{z\in\mathbb{Z}^d}\left\langle \mathcal{X}_x(z,\cdot)\partial_z\left(Q_x(y,\cdot)\mathcal{V}(y,\cdot)\right)\right\rangle_{\mu_\beta}. \tag{8.38}$$

Summing over the points $y\in\mathbb{Z}^d$ and performing an integration by parts in the variable y, we deduce that

$$\begin{aligned}\sum_{y\in\mathbb{Z}^d}\operatorname{cov}\left[X_x, Q_x(y,\cdot)\mathcal{V}(y,\cdot)\right] &= \sum_{y,z\in\mathbb{Z}^d}\left\langle \mathcal{X}_x(z,\cdot)\partial_z\left(Q_x(y,\cdot)\mathcal{V}(y,\cdot)\right)\right\rangle_{\mu_\beta} \\ &= \sum_{y,z\in\mathbb{Z}^d}\left\langle \mathcal{X}_x(z,\cdot)\partial_z\left(n_{Q_x}(y,\cdot)\mathrm{d}^*\mathcal{V}(y,\cdot)\right)\right\rangle_{\mu_\beta}.\end{aligned}$$

We split the proof into two substeps:

- In Substep 1.1, we compute the value of $\partial_z\left(n_{Q_x}(y,\cdot)\mathrm{d}^*\mathcal{V}(y,\cdot)\right)$. We prove the identity (8.50) and the inequalities (8.51);
- In Substep 1.2, we deduce the expansion (8.35) from Substep 1.1.

Substep 1.1. We first expand the derivative

$$\partial_z\left(n_{Q_x}(y,\cdot)\mathrm{d}^*\mathcal{V}(y,\cdot)\right) = \underbrace{\left(\partial_z n_{Q_x}(y,\cdot)\right)\mathrm{d}^*\mathcal{V}(y,\cdot)}_{(8.39)-(i)} + \underbrace{n_{Q_x}(y,\cdot)\partial_z\mathrm{d}^*\mathcal{V}(y,\cdot)}_{(8.39)-(ii)}. \tag{8.39}$$

The term (8.39)-(i) can be computed explicitly from the definition of the charge n_{Q_x} and the identity $q = \mathrm{d}n_q$. We obtain

$$\begin{aligned}\left(\partial_z n_{Q_x}(y,\phi)\right)\mathrm{d}^*\mathcal{V}(y,\phi) &= \left(\sum_{q\in\mathcal{Q}}4\pi^2 z(\beta,q)\left(\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\nabla G_x,n_q)\right)\right)n_q(y)\otimes q(z)\right)\mathrm{d}^*\mathcal{V}(y,\phi) \\ &= \mathrm{d}_z\left(\left(\sum_{q\in\mathcal{Q}}4\pi^2 z(\beta,q)\left(\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\nabla G_x,n_q)\right)\right)n_q(y)\otimes n_q(z)\right)\mathrm{d}^*\mathcal{V}(y,\phi)\right).\end{aligned} \tag{8.40}$$

We then estimate the term in the right side of (8.40). To this end, we note that the sum over the charges $q\in\mathcal{Q}$ can be restricted to the set of charges $\mathcal{Q}_{y,z}$ and use the two inequalities: first, $\sum_{q\in\mathcal{Q}_{y,z}}e^{-c\sqrt{\beta}\|q\|_1}\leq e^{-c\sqrt{\beta}|y-z|}$ established in (A.20) and, for each charge $q\in\mathcal{Q}_y$, $\left|\sin\left(2\pi(\nabla G_x,n_q)\right)\right|\leq\frac{C_q}{|x-y|^{d-1}}$. We deduce that

$$\begin{aligned}\left|\sum_{q\in\mathcal{Q}}z(\beta,q)\sin\left(2\pi(\phi,q)\right)\sin\left(2\pi(\nabla G_x,n_q)\right)n_q(y)\otimes n_q(z)\right| &\leq \sum_{q\in\mathcal{Q}_{x,y}}e^{-c\sqrt{\beta}\|q\|_1}\frac{C_q}{|y-x|^{d-1}} \\ &\leq \frac{Ce^{-c\sqrt{\beta}|y-z|}}{|y-x|^{d-1}}.\end{aligned} \tag{8.41}$$

Combining the estimate (8.41) with the inequality (8.33) on the codifferential of the function $\mathcal{V}$, we obtain, for each pair of points $z, y \in \mathbb{Z}^d$,

$$\left\| \left(\sum_{q \in \mathcal{Q}} z(\beta, q) \left(\sin\left(2\pi(\cdot, q)\right) \sin\left(2\pi(\nabla G_x, n_q)\right)\right) n_q(y) \otimes n_q(z) \right) \mathrm{d}^* \mathcal{V}(y, \cdot) \right\|_{L^2(\mu_\beta)} \leq \frac{C e^{-c\sqrt{\beta}|y-z|}}{|y-x|^{d-1} \times |y|^{d-1-\varepsilon}}.$$

We now treat the term (8.39)-(ii). To estimate the $L^2(\mu_\beta)$-norm of the map $\partial_z \mathrm{d}^* \mathcal{V}(y, \phi)$, we start from the definition of the map $\mathcal{V}$ as the solution of the Helffer-Sjöstrand equation (8.34) and apply the derivative ∂_z to both sides of the identity (8.34). Following the arguments developed at the beginning of Section 5.4, we obtain that the map $\mathcal{V}_{\mathrm{sec}} : (y, z, \phi) \to \partial_z \mathcal{V}(y, \phi)$ is the solution of the second-order Helffer-Sjöstrand equation

(8.42)

$$\begin{aligned}
\mathcal{L}_{\mathrm{sec}} \mathcal{V}_{\mathrm{sec}}(y, z, \phi) = & \left(\sum_{q \in \mathcal{Q}} 4\pi^2 z(\beta, q) \left(\sin\left(2\pi(\phi, q)\right) \sin\left(2\pi(\nabla G, n_q)\right)\right) q(y) \otimes q(z) \right) Y_0 \\
& + \sum_{q \in \mathcal{Q}} 2\pi z(\beta, q) \sin\left(2\pi(\phi, q)\right) (\mathrm{d}^* \mathcal{V}, n_q)\, q(y) \otimes q(z) \\
& - Q_0(y, \phi) \otimes \left(Q_0(z, \phi) + \frac{1}{2} 2\pi \sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\phi, q)\right) \left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right) q(z) \right) Y_0.
\end{aligned}$$

We decompose the function $\mathcal{V}_{\mathrm{sec}}$ into three functions, $\mathcal{V}_{\mathrm{sec},1}$, $\mathcal{V}_{\mathrm{sec},2}$ and $\mathcal{V}_{\mathrm{sec},3}$ according to the three terms in the right side of (8.42), i.e.,

(8.43)

$$\begin{cases}
\mathcal{L}_{\mathrm{sec}} \mathcal{V}_{\mathrm{sec},1}(y, z, \phi) = \left(\sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\phi, q)\right) \sin\left(2\pi(\nabla G, n_q)\right) q(y) \otimes q(z) \right) Y_0, \\
\mathcal{L}_{\mathrm{sec}} \mathcal{V}_{\mathrm{sec},2}(y, z, \phi) = \sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\phi, q)\right) (\mathrm{d}^* \mathcal{V}, n_q)\, q(y) \otimes q(z), \\
\mathcal{L}_{\mathrm{sec}} \mathcal{V}_{\mathrm{sec},3}(y, z, \phi) = -Q_0(y, \phi) \otimes \left(Q_0(z, \phi) + \frac{1}{2} \sum_{q \in \mathcal{Q}} z(\beta, q) \sin\left(2\pi(\phi, q)\right) \left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right) q(z) \right) Y_0.
\end{cases}$$

We then estimate the three terms $\mathcal{V}_{\mathrm{sec},1}$, $\mathcal{V}_{\mathrm{sec},2}$ and $\mathcal{V}_{\mathrm{sec},3}$ separately. The first two terms can be estimated by using a strategy similar to the one used in Step 1 of the proof of Lemma 4.3: we use the equations (8.43) to obtain explicit formulae in terms of the Green's matrix $\mathcal{G}_{\mathrm{sec}}$ associated with the second-order Helffer-Sjöstrand equation, and use Proposition 5.13 to estimate them. We omit the technical details which can be found in the long version of this article ([36, Chapter 8, Lemma 3.1]). The results are collected in (8.51) below.

The estimate for the term $\mathcal{V}_{\mathrm{sec},3}$ is more involved. Using a similar strategy with additional technical details: we prove that there exists a map $\mathcal{W}_{\mathrm{sec},3} : \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega \to \mathbb{R}^{d \times \binom{d}{2}}$ which satisfies the identity, for each $(y, z, \phi) \in \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega$,

$$\mathcal{V}_{\mathrm{sec},3}(y, z, \phi) = \mathrm{d}_z \mathcal{W}_{\mathrm{sec},3}(y, z, \phi), \tag{8.44}$$

as well as the upper bounds

$$\left\| \mathcal{W}_{\mathrm{sec},3}(y, z, \cdot) \right\|_{L^2(\mu_\beta)} \leq \frac{C}{|y|^{d-\frac{3}{2}-\varepsilon} \times |z|^{d-\frac{3}{2}-\varepsilon}} \quad \text{and} \quad \left\| \mathrm{d}_y^* \mathcal{W}_{\mathrm{sec},3}(y, z, \cdot) \right\|_{L^2(\mu_\beta)} \leq \frac{C}{|y|^{d-1-\varepsilon} \times |z|^{d-1-\varepsilon}}. \tag{8.45}$$

The strategy to prove the identity (8.44) and the estimate (8.45) is the following. We use the dynamic formulation to solve the Helffer-Sjöstrand equation (8.43), and obtain the identity

$$\begin{aligned}
(8.46) \quad \mathcal{V}_{\mathrm{der},3}(y, z, \phi) = & \sum_{y_1, z_1 \in \mathbb{Z}^d} \int_0^\infty \mathbb{E}_\phi \left[-Y_0(\phi_t) P^{\phi_\cdot}_{\mathrm{der}}(t, y_1, z_1; y, z) Q_0(y_1, \phi_t) \otimes Q_0(z_1, \phi_t) \right] \\
-\pi \sum_{y_1, z_1 \in \mathbb{Z}^d} \sum_{q \in \mathcal{Q}} & z(\beta, q) \left(\cos\left(2\pi(\nabla G, n_q)\right) - 1\right) \int_0^\infty \mathbb{E}_\phi \left[\sin\left(2\pi(\phi_t, q)\right) Y_0(\phi_t) P^{\phi_\cdot}_{\mathrm{der}}(t, y_1, z_1; y, z) Q_0(y_1, \phi_t) \otimes q(z_1) \right],
\end{aligned}$$

where, given a trajectory $(\phi_t)_{t\geq 0}$ of the Langevin dynamics, the map $P^{\phi_\cdot}_{\mathrm{der}}(\cdot,\cdot,\cdot;y,z):(0,\infty)\times\mathbb{Z}^d\times\mathbb{Z}^d\to\mathbb{R}^{\binom{d}{2}^4}$ denotes the solution of the parabolic system of equations,

$$\begin{cases} \partial_t P^{\phi_\cdot}_{\mathrm{der}}(\cdot,\cdot,\cdot;y,z) + \left(\mathcal{L}^{\phi_t}_{\mathrm{spat},x} + \mathcal{L}^{\phi_t}_{\mathrm{spat},y}\right) P^{\phi_\cdot}_{\mathrm{der}}(\cdot,\cdot,\cdot;y,z) = 0 & \text{in } (0,\infty)\times\mathbb{Z}^d\times\mathbb{Z}^d, \\ P^{\phi_\cdot}_{\mathrm{der}}(0,\cdot,\cdot;y,z) = \delta_{(y,z)} & \text{in } \mathbb{Z}^d\times\mathbb{Z}^d. \end{cases}$$

Let us observe that, thanks to the specific structures of the second-order Helffer-Sjöstrand equation (see (5.32)) and of the right-hand side of (8.46), one can factorise this term and obtain

(8.47)
$$\begin{aligned} \mathcal{V}_{\mathrm{der},3}(y,z,\phi) = &- \sum_{q_1,q_2\in\mathcal{Q}} z(\beta,q_1)\, z(\beta,q_2) \sin(2\pi(\nabla G, n_{q_1})) \sin(2\pi(\nabla G, n_{q_2})) \\ &\qquad \times \int_0^\infty \mathbb{E}_\phi\left[\cos(2\pi(\phi_t,q_1))\cos(2\pi(\phi_t,q_2))\, Y_0(\phi_t)(q_1, P^{\phi_\cdot}(t,\cdot;y))\otimes(q_2,P^{\phi_\cdot}(t,\cdot;z))\right] dt \\ &+ \pi \sum_{q_1,q_2\in\mathcal{Q}} z(\beta,q_1)\, z(\beta,q_2) \sin(2\pi(\nabla G, n_{q_1})) \left(\cos(2\pi(\nabla G, n_{q_2})) - 1\right) \\ &\qquad \times \int_0^\infty \mathbb{E}_\phi\left[\cos(2\pi(\phi_t,q_1))\cos(2\pi(\phi_t,q_2))\, Y_0(\phi_t)(q_1, P^{\phi_\cdot}(t,\cdot;y))\otimes(q_2,P^{\phi_\cdot}(t,\cdot;z))\right] dt. \end{aligned}$$

The strategy is then to use the symmetry of the spatial operator $\mathcal{L}^{\phi_t}_{\mathrm{spat}}$, to observe that, for any charge $q\in\mathcal{Q}$ and any time $T>0$, if we let $R_{T,q}:(0,T)\times\mathbb{Z}^d\to\mathbb{R}^{\binom{d}{2}}$ be the solution of the parabolic system of equations (note that we reverse the time in the dynamic ϕ_t and replaced it by ϕ_{T-t})

$$\begin{cases} \partial_t R_{T,q} - \mathcal{L}^{\phi_{T-t}}_{\mathrm{spat}} R_{T,q} = 0 & \text{in } (0,T)\times\mathbb{Z}^d, \\ R_{T,q}(0,\cdot) = q & \text{in } \mathbb{Z}^d, \end{cases}$$

then we have $R_{T,q}(T,y) = (q, P^{\phi_\cdot}(T,\cdot;y))$. We may then consider the solution $S_{T,q}:(0,T)\times\mathbb{Z}^d\to\mathbb{R}^d$ of the parabolic system of equations

$$\begin{cases} \partial_t S_{T,q} - \left(\frac{1}{2\beta}\Delta - \frac{1}{2\beta}\sum_{n\geq 1}\frac{1}{\beta^{\frac{n}{2}}}(-\Delta)^{n+1}\right) S_{T,q} = -\sum_{q\in\mathcal{Q}} z(\beta,q)\cos(2\pi(\phi_{T-\cdot},q))\,(\nabla_q R_{T,q}) n_q & \text{in } (0,\infty)\times\mathbb{Z}^d, \\ S^{\phi_\cdot}_{T,q}(0,\cdot) = n_q & \text{in } \mathbb{Z}^d, \end{cases}$$

and define $Q^{\phi_\cdot}_q(T,y) := S_{T,q}(T,y)$. Using the estimate on the heat-kernel in dynamic environment stated in Proposition 3.10 and the Duhamel principle, one can prove the following results, for each point $y\in\mathbb{Z}^d$, and each time $t\geq 1$,

$$P^{\phi_\cdot}_q(t,y) = \mathrm{d}Q^{\phi_\cdot}_q(t,y) \quad\text{and}\quad \left|Q^{\phi_\cdot}_q(t,y)\right| \leq C_q t^{\varepsilon}\Phi_C(t, y-y_1), \tag{8.48}$$

where y_1 is a point which lies in the support of the charge n_q. We then define $\mathcal{W}_{\mathrm{der},3}$ by the formula

$$\begin{aligned} \mathcal{W}_{\mathrm{der},3}(y,z,\phi) = &- \sum_{q_1,q_2\in\mathcal{Q}} z(\beta,q_1)\, z(\beta,q_2) \sin(2\pi(\nabla G, n_{q_1})) \sin(2\pi(\nabla G, n_{q_2})) \\ &\qquad \times \int_0^\infty \mathbb{E}_\phi\left[\cos(2\pi(\phi_t,q_1))\cos(2\pi(\phi_t,q_2))\, Y_0(\phi_t)(q_1, P^{\phi_\cdot}(t,y))\otimes Q^{\phi_\cdot}_{q_2}(t,z)\right] dt \\ &+ \frac{1}{2}2\pi \sum_{q_1,q_2\in\mathcal{Q}} z(\beta,q_1)\, z(\beta,q_2) \sin(2\pi(\nabla G, n_{q_1})) \left(\cos(2\pi(\nabla G, n_{q_2})) - 1\right) \\ &\qquad \times \int_0^\infty \mathbb{E}_\phi\left[\cos(2\pi(\phi_t,q_1))\cos(2\pi(\phi_t,q_2))\, Y_0(\phi_t)(q_1, P^{\phi_\cdot}(t,y))\otimes Q^{\phi_\cdot}_{q_2}(t,z)\right] dt. \end{aligned} \tag{8.49}$$

We can verify the equality (8.44) by an explicit (and straightforward) computation making use of the identities (8.47) and (8.48). The bounds (8.45) can be verified by using the explicit formula (8.49), the bound on the heat kernel in the dynamic environment (Proposition 3.10), and the bounds on the function Q_q stated in (8.48). We omit the details which can be found in the long version of this article ([36, Chapter 8, Lemma 3.1]).

Conclusion of Substep 1.1. We have the identity, for each pair of points $(y,z) \in \mathbb{Z}^d$,

$$\begin{aligned} \partial_z \left(n_{Q_x}(y,\cdot) \mathrm{d}^* \mathcal{V}(y,\cdot) \right) &= \left(\partial_z n_{Q_x}(y,\cdot) \right) \mathrm{d}^* \mathcal{V}(y,\cdot) + n_{Q_x}(y,\cdot) \partial_z \mathrm{d}^* \mathcal{V}(y,\cdot) \\ &= \left(\partial_z n_{Q_x}(y,\cdot) \right) \mathrm{d}^* \mathcal{V}(y,\cdot) + n_{Q_x}(y,\cdot) \mathrm{d}_y^* \mathcal{V}_{\mathrm{sec},1}(y,z,\cdot) \\ &\quad + n_{Q_x}(y,\cdot) \mathrm{d}_y^* \mathcal{V}_{\mathrm{sec},2}(y,z,\cdot) + \mathrm{d}_z \left(n_{Q_x}(y,\cdot) \mathrm{d}_y^* \mathcal{W}_{\mathrm{sec},3}(y,z,\cdot) \right), \end{aligned} \tag{8.50}$$

with the estimates

$$\begin{cases} \left\| \left(\partial_z n_{Q_x}(y,\cdot) \right) \mathrm{d}^* \mathcal{V}(y,\cdot) \right\|_{L^2(\mu_\beta)} \leq \dfrac{C e^{-c\sqrt{\beta}|y-z|}}{|y-x|^{d-1} \times |y|^{d-1-\varepsilon}}, \\ \left\| n_{Q_x}(y,\cdot) \mathrm{d}_y^* \mathcal{V}_{\mathrm{sec},1}(y,z,\cdot) \right\|_{L^2(\mu_\beta)} \leq \dfrac{C}{|x-y|^{d-1-\varepsilon} \times |z-y|^{d+1-\varepsilon} \times \max\left(|y|,|z|\right)^{d-1-\varepsilon}}, \\ \left\| n_{Q_x}(y,\cdot) \mathrm{d}_y^* \mathcal{V}_{\mathrm{sec},2}(y,z,\cdot) \right\|_{L^2(\mu_\beta)} \leq \dfrac{C}{|x-y|^{d-1-\varepsilon} \times |z-y|^{d+1-\varepsilon} \times \max\left(|y|,|z|\right)^{d-1-\varepsilon}}, \\ \left\| n_{Q_x}(y,\cdot) \mathrm{d}_y^* \mathcal{W}_{\mathrm{sec},3}(y,z,\cdot) \right\|_{L^2(\mu_\beta)} \leq \dfrac{C}{|y-x|^{d-1}|y|^{d-1-\varepsilon}|z|^{d-1-\varepsilon}}. \end{cases} \tag{8.51}$$

Substep 1.2. We prove the covariance estimate (8.37). By the Helffer-Sjöstrand representation formula we have, for each point $y \in \mathbb{Z}^d$,

$$\mathrm{cov}\left[X_x, Q_x(y,\cdot) \mathcal{V}(y,\cdot) \right] = \sum_{z \in \mathbb{Z}^d} \left\langle \mathcal{X}_x(z,\cdot) \partial_z \left(Q_x(y,\cdot) \mathcal{V}(y,\cdot) \right) \right\rangle_{\mu_\beta}. \tag{8.52}$$

Using the formula (8.52), we write

$$\begin{aligned} \sum_{y \in \mathbb{Z}^d} \mathrm{cov}\left[X_x, Q_x(y,\cdot) \mathcal{V}(y,\cdot) \right] &= \sum_{y,z \in \mathbb{Z}^d} \left\langle \mathcal{X}_x(z,\cdot) \partial_z \left(Q_x(y,\cdot) \mathcal{V}(y,\cdot) \right) \right\rangle_{\mu_\beta} \\ &= \sum_{y,z \in \mathbb{Z}^d} \left\langle \mathcal{X}_x(z,\cdot) \partial_z \left(n_{Q_x}(y,\cdot) \mathrm{d}^* \mathcal{V}(y,\cdot) \right) \right\rangle_{\mu_\beta}. \end{aligned} \tag{8.53}$$

We combine the identities (8.50) and (8.53), and obtain

$$\begin{aligned} &\sum_{y \in \mathbb{Z}^d} \mathrm{cov}\left[X_x, Q_x(y,\cdot) \mathcal{V}(y,\cdot) \right] \\ &\quad = \sum_{y,z \in \mathbb{Z}^d} \left\langle \mathcal{X}_x(z,\cdot) \left(\partial_z n_{Q_x}(y) \right) \mathrm{d}^* \mathcal{V}(y,\cdot) \right\rangle_{\mu_\beta} + \sum_{y,z \in \mathbb{Z}^d} \left\langle \mathcal{X}_x(z,\cdot) n_{Q_x}(y) \mathrm{d}_y^* \mathcal{V}_{\mathrm{sec},1}(y,z,\cdot) \right\rangle_{\mu_\beta} \\ &\quad\quad + \sum_{y,z \in \mathbb{Z}^d} \left\langle \mathcal{X}_x(z,\cdot) n_{Q_x}(y) \mathrm{d}_y^* \mathcal{V}_{\mathrm{sec},2}(y,z,\cdot) \right\rangle_{\mu_\beta} + \sum_{y,z \in \mathbb{Z}^d} \left\langle \mathrm{d}^* \mathcal{X}_x(z,\cdot) \left(n_{Q_x}(y) \mathrm{d}_y^* \mathcal{W}_{\mathrm{sec},3}(y,z,\cdot) \right) \right\rangle_{\mu_\beta}. \end{aligned}$$

We use the estimates (8.32) on the function $\mathcal{X}_x$ and the estimates (8.51). We obtain

$$\begin{aligned} \left| \sum_{y \in \mathbb{Z}^d} \mathrm{cov}\left[X_x, Q_x(y,\cdot) \mathcal{V}(y,\cdot) \right] \right| &\leq \sum_{y,z \in \mathbb{Z}^d} \frac{C}{|z-x|^{d-2-\varepsilon}} \frac{e^{-c\sqrt{\beta}|y-z|}}{|y-x|^{d-1} \times |y|^{d-1-\varepsilon}} \\ &\quad + \sum_{y,z \in \mathbb{Z}^d} \frac{C}{|z-x|^{d-2-\varepsilon}} \frac{1}{|x-y|^{d-1-\varepsilon} \times |z-y|^{d+1-\varepsilon} \times \max\left(|y|,|z|\right)^{d-1-\varepsilon}} \\ &\quad + \sum_{y,z \in \mathbb{Z}^d} \frac{C}{|z-x|^{d-1-\varepsilon}} \frac{1}{|y-x|^{d-1}|y|^{d-1-\varepsilon}|z|^{d-1-\varepsilon}}. \end{aligned} \tag{8.54}$$

The term in the right side can be estimated by an explicit computation. We skip the details here and obtain the expansion (8.37). Step 1 is complete.

Step 2. To complete the proof of Lemma 4.4, there remains to prove the expansion (8.36). The strategy of the proof relies on the symmetry of the Helffer-Sjöstrand operator $\mathcal{L}$; if we let $\mathcal{U}_x$ the solution of the equation $\mathcal{L}\mathcal{U}_x = Q_x$ in $\mathbb{Z}^d \times \Omega$, then we have the identities

$$\sum_{y \in \mathbb{Z}^d} \left\langle Q_x(y,\cdot) \mathcal{V}(y,\cdot) \right\rangle_{\mu_\beta} = \sum_{y \in \mathbb{Z}^d} \left\langle \mathcal{U}_x\left(y,\cdot\right) Q_0(y,\cdot) Y_0 \right\rangle_{\mu_\beta}$$

and

$$\sum_{y \in \mathbb{Z}^d} \left\langle Q_x(y,\cdot) \mathcal{U}(y,\cdot) \right\rangle_{\mu_\beta} = \sum_{y \in \mathbb{Z}^d} \left\langle \mathcal{U}_x(y,\cdot) Q_0(y,\cdot) \right\rangle_{\mu_\beta}.$$

Using these identities, we see that the expansion (8.36) is equivalent to

$$\sum_{y\in\mathbb{Z}^d} \left\langle \mathcal{U}_x\left(y,\cdot\right) Q_0(y,\cdot) Y_0 \right\rangle_{\mu_\beta} = \left\langle Y_0\right\rangle_{\mu_\beta} \sum_{y\in\mathbb{Z}^d} \left\langle \mathcal{U}_x\left(y,\cdot\right) Q_0(y,\cdot)\right\rangle_{\mu_\beta} + O\left(\frac{C}{|x|^{d-1-\varepsilon}}\right).$$

The proof of this result is similar to the proof written in Step 1, and is in fact simpler since we do not have to treat the term $\mathcal{V}_{\mathrm{sec},3}$ in (8.43); we omit the details. □

8.4. **Using the symmetry and rotation invariance of the dual Villain model.** This section is devoted to the proof of some properties of the discrete convolution of the discrete Green's function on the lattice $\mathbb{Z}^d$. We recall the definition of the group H of the lattice-preserving maps introduced in Section 2.1.

Lemma 8.1. *Fix four integers $j, j_1, k, k_1 \in \{1, \ldots, d\}$ and let $F : \mathbb{Z}^d \to \mathbb{R}$ be the function*

$$F_{j,k,j_1,k_1}(x) := \sum_{y,\kappa\in\mathbb{Z}^d} \nabla_j G(y) \nabla_k G(x-y-\kappa) \nabla_{j_1}\nabla_{k_1} G(\kappa).$$

Then, if we let $J_{j,k,j_1,k_1} : \mathbb{R}^d \setminus \{0\} \to \mathbb{R}$ be the $(2-d)$-homogeneous map whose Fourier transform is given by the formula $\widehat{J_{i,j,k,l}}(\xi) = \xi_i\xi_j\xi_k\xi_l \left|\xi\right|^{-6}$. Then for any $\varepsilon > 0$, one has the identity

$$F_{j,k,j_1,k_1}(x) = J_{j,k,j_1,k_1}(x) + O\left(\frac{1}{|x|^{d-1-\varepsilon}}\right).$$

A direct consequence of the previous lemma is the corollary stated below.

Corollary 8.2. *Fix two integers $j, j_1 \in \{1, \ldots, d\}$ and let $F_{j,j_1} : \mathbb{Z}^d \to \mathbb{R}$ be the map*

$$F_{j,j_1}(x) = \sum_{y\in\mathbb{Z}^d} \nabla_j G(y) \nabla_{j_1} G(x-y),$$

then, for any $\varepsilon > 0$, one has the identity

$$F_{j,j_1}(x) = J_{j,j_1}(x) + O\left(\frac{1}{|x|^{d-1-\varepsilon}}\right),$$

where the map J_{j,j_1} is $(2-d)$-homogeneous and its Fourier transform is given by the formula $\widehat{J_{i,j}}(\xi) = \xi_i\xi_j \left|\xi\right|^{-4}$.

The proofs of this lemma and this corollary follow standard arguments; we refer to the long version of this article ([36, Chapter 8, Section 4]) for the details.

The following proposition is used in the proof of Theorem 1. It asserts that if a linear combination of the maps $F_{i,j,k,l}$ and $F_{i,j}$, with a specific structure given by the problem considered in this article, is invariant under the group H lattice preserving maps, then it must satisfy the expansion given by (8.56).

Proposition 8.3. *Assume that there exist coefficients $(c_{ij})_{1\le i,j\le d}$ and $(K_{ij})_{1\le i,j\le d}$, an exponent $\alpha > 0$ and a map U which is invariant under the group H of the lattice-preserving maps such that*

$$U(x) = \sum_{i,j,k,l=1}^{d} c_{ij}c_{kl} F_{i,j,k,l}(x) + \sum_{i,j=1}^{d} K_{ij} F_{i,j}(x) + O\left(\frac{C}{|x|^{d-2+\alpha}}\right), \tag{8.55}$$

then there exists a constant $c \in \mathbb{R}$ such that

$$U(x) = \frac{c}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\alpha}}\right). \tag{8.56}$$

Proof. Applying Lemma 8.1 and Corollary 8.2, the expansion (8.55) can be rewritten

$$U(x) = \sum_{i,j,k,l=1}^{d} c_{ij}c_{kl} J_{i,j,k,l}(x) + \sum_{i,j=1}^{d} K_{ij} J_{i,j}(x) + O\left(\frac{C}{|x|^{d-2+\alpha}}\right).$$

Using that the maps $J_{i,j,k,l}$ and $J_{i,j}$ are $(2-d)$-homogeneous, we see that the assumption that U is invariant under the lattice-preserving maps implies that the same property holds for the function $\sum_{i,j,k,l=1}^{d} c_{ij}c_{kl}J_{i,j,k,l} + \sum_{i,j=1}^{d} K_{ij}J_{i,j}$: for each $h \in H$ and each $x \in \mathbb{Z}^d \setminus \{0\}$, one has

$$\sum_{i,j,k,l=1}^{d} c_{ij}c_{kl} J_{i,j,k,l}(h(x)) + \sum_{i,j=1}^{d} K_{ij} J_{i,j}(h(x)) = \sum_{i,j,k,l=1}^{d} c_{ij}c_{kl} J_{i,j,k,l}(x) + \sum_{i,j=1}^{d} K_{ij} J_{i,j}(x). \tag{8.57}$$

Using the homogeneity of the maps $J_{i,j,k,l}$ and $J_{i,j}$, the result can be extended to each point of $\mathbb{R}^d \setminus \{0\}$. Let us denote by P the homogeneous polynomial of degree 4

$$P(\xi) = \left(\sum_{i,j=1}^{d} c_{ij} \xi_i \xi_j \right)^2 + |\xi|^2 \sum_{i,j=1}^{d} K_{ij} \xi_i \xi_j, \tag{8.58}$$

so that the Fourier transform of the map $\sum_{i,j,k,l=1}^{d} c_{ij} c_{kl} J_{i,j,k,l} + \sum_{i,j=1}^{d} K_{ij} J_{i,j}$ is equal to the function $\xi \mapsto P(\xi) |\xi|^{-6}$.

Taking the Fourier transform on both sides of the identity (8.57), we obtain the identity, for any $\xi \in \mathbb{R}^d$, and any $h \in H$,

$$P(h(\xi)) = P(\xi). \tag{8.59}$$

Using the definition (8.58), the property (8.59) and an explicit computation which we omit here, we prove that there exists a coefficient $a \in \mathbb{R}$ such that

$$P(\xi) = a \left(\sum_{i=1}^{d} \xi_i^2 \right)^2 = a |\xi|^4. \tag{8.60}$$

The equality (8.60) implies that the Fourier transform of the map $\sum_{i,j,k,l=1}^{d} c_{ij} c_{kl} J_{i,j,k,l} + \sum_{i,j=1}^{d} K_{ij} J_{i,j}$ is equal to $a|\xi|^{-2}$, which implies, by taking the inverse Fourier transform, that there exists a constant c such that, for any $x \in \mathbb{R}^d \setminus \{0\}$,

$$\sum_{i,j,k,l=1}^{d} c_{ij} c_{kl} J_{i,j,k,l}(x) + \sum_{i,j=1}^{d} K_{ij} J_{i,j}(x) = \frac{c}{|x|^{d-2}}. \tag{8.61}$$

Combining the identity (8.61) with the expansion (8.55), we have obtained

$$U(x) = \frac{c}{|x|^{d-2}} + O\left(\frac{C}{|x|^{d-2+\alpha}} \right).$$

The proof of Proposition 8.3 is complete. □

8.5. **Treating the error term $\mathcal{E}_{q_1,q_2}$.** This section is devoted to the treatment the error term $\mathcal{E}_{q_1,q_2}$ used in the proof of Theorem 1.

Proposition 8.4. *Fix two exponents $\gamma, \varepsilon \in (0,1]$ such that $\varepsilon \leq \frac{\gamma}{4(d-2)}$, and two charges $q_1, q_2 \in \mathcal{Q}$. Let $\mathcal{E}_{q_1,q_2} : \mathbb{Z}^d \to \mathbb{R}$ be a function which satisfies the pointwise and L^1-estimates, for each point $\kappa \in \mathbb{Z}^d$ and each radius $R \geq 1$,*

$$|\mathcal{E}_{q_1,q_2}(\kappa)| \leq \frac{C}{|\kappa|^{d-\varepsilon}} \text{ and } \sum_{\kappa \in B_{2R} \setminus B_R} |\mathcal{E}_{q_1,q_2}(\kappa)| \leq C R^{-\gamma}. \tag{8.62}$$

Then, the constant $K_{q_1,q_2} := 4\pi^2 \sum_{\kappa \in \mathbb{Z}^d} \mathcal{E}_{q_1,q_2}(\kappa)$ is well-defined in the sense that the sum converges absolutely, and one has the expansion

$$\begin{aligned} 4\pi^2 \sum_{z_2, \kappa \in \mathbb{Z}^d} & \nabla G(z_2) \cdot (n_{q_2}) \nabla G_x(z_2 + \kappa) \cdot (n_{q_1}) \mathcal{E}_{q_1,q_2}(\kappa) \\ & = K_{q_1,q_2} \sum_{z_2 \in \mathbb{Z}^d} \nabla G(z_2) \cdot (n_{q_2}) \nabla G(z_2 - x) \cdot (n_{q_1}) + O\left(\frac{C_{q_1,q_2}}{|x|^{d-2+\frac{\gamma}{4(d-2)}}} \right). \end{aligned} \tag{8.63}$$

Proof. The proof of this lemma relies on (elementary) considerations about the discrete Green's function, we refer to the long version of this article ([36, Chapter 8, Section 5]) for the details. □

Appendix A. List of notation and preliminary results

A.1. Notation and assumptions.

A.1.1. *General notation and assumptions.* We work on the Euclidean lattice $\mathbb{Z}^d$ in dimension $d \geq 3$. We denote by $|\cdot|$ the standard Euclidean norm on the lattice $\mathbb{Z}^d$. We say that two points $x, y \in \mathbb{Z}^d$ are neighbors, and denote it by $x \sim y$, if $|x-y| = 1$. We denote by $e_1, \ldots, e_k$ the canonical basis of $\mathbb{R}^d$.

Given a subset $U \subseteq \mathbb{Z}^d$, we define its interior U° and its boundary ∂U by the formulae

$$U^\circ := \{x \in U : x \sim y \implies y \in U\} \quad \text{and} \quad \partial U := U \setminus U^\circ.$$

If the subset $U \subseteq \mathbb{Z}^d$ is finite, we denote by $|U|$ its cardinality and refer to this quantity as the volume of U. We denote by $\operatorname{diam} U$ the diameter of U defined by the formula $\operatorname{diam} U := \sup_{x,y \in U} |x-y|$. Given a point $x \in \mathbb{Z}^d$ and a radius $r > 0$, we denote by $B(x,r)$ the discrete euclidean ball of center x and radius r. We frequently use the notation B_r to mean $B(0,r)$. We also define the annulus $A_R := B_{2R} \setminus B_R$.

A discrete cube $\square$ of $\mathbb{Z}^d$ is a subset of the form

$$\square := x + [-N, N]^d \cap \mathbb{Z}^d \text{ with } x \in \mathbb{Z}^d \text{ and } N \in \mathbb{N}. \tag{A.1}$$

We refer to the point x as the center of the cube $\square$, and to the integer $2N+1$ as its length. We denote by $\square_L := [-N,N]^d \cap \mathbb{Z}^d$. Given a parameter $r > 0$, we use the nonstandard convention of denoting by $r\square$ the cube

$$r\square := x + [-rN, rN]^d \cap \mathbb{Z}^d.$$

If $\square$ is the cube given by (A.1), then we define the trimmed cube $\square^-$ by the formula

$$\square^- := x + \left(-\frac{N}{2} + \frac{\sqrt{N}}{10}, \frac{N}{2} - \frac{\sqrt{N}}{10}\right)^d \cap \mathbb{Z}^d. \tag{A.2}$$

Given three real numbers $X, Y \in \mathbb{R}$ and $\kappa \in [0, \infty)$, we write

$$X = Y + O(\kappa) \quad \text{if and only if} \quad |X - Y| \leq \kappa.$$

For each integer $i \in \{1, \ldots, d\}$, we denote by h_i the reflection of the lattice $\mathbb{Z}^d$ with respect to the hyperplane $\left\{z \in \mathbb{Z}^d : z_i = 0\right\}$, i.e.,

$$h_i := \begin{cases} \mathbb{Z}^d \to \mathbb{Z}^d \\ (z_1, \ldots, z_d) \mapsto (z_1, \ldots, -z_i, \ldots, z_d). \end{cases}$$

For each pair of integers $i, j \in \{1, \ldots, d\}$ with $i < j$, we denote by h_{ij} the map

$$h_{ij} := \begin{cases} \mathbb{Z}^d \to \mathbb{Z}^d \\ (z_1, \ldots, z_d) \mapsto (z_1, \ldots, z_j, \ldots, z_i, \ldots, z_d). \end{cases}$$

We define H the group of lattice preserving transformation to be the group of linear maps generated by the collections of functions $(h_i)_{1 \leq i \leq d}$ and $(h_{ij})_{1 \leq i < j \leq d}$ with respect to the composition law.

We frequently consider functions defined from $\mathbb{Z}^d$ and valued in $\mathbb{R}$ of the form $x \to |x|^{-k}$. We implicitly extend these functions at the point $x = 0$ by the value 1 so that they are defined on the entire lattice $\mathbb{Z}^d$.

A.1.2. *Notation for vector-valued functions.* For each integer $k \in \mathbb{N}$, we let $\mathcal{F}\left(\mathbb{Z}^d, \mathbb{R}^k\right)$ be the set of functions defined on $\mathbb{Z}^d$ and taking values in $\mathbb{R}^k$. Given a function $g \in \mathcal{F}\left(\mathbb{Z}^d, \mathbb{R}^k\right)$, we denote by $g_1, \ldots, g_k$ its components on the canonical basis of $\mathbb{R}^k$ and write $g = (g_1, \ldots, g_k)$. We define the support of the function g to be the set

$$\operatorname{supp} g := \left\{x \in \mathbb{Z}^d : g(x) \neq 0\right\}.$$

For each integer $i \in \{1, \ldots, d\}$, we define its discrete i-th derivative $\nabla_i g : \mathbb{Z}^d \to \mathbb{R}^k$ by the formula, for each $x \in \mathbb{Z}^d$,

$$\nabla_i g(x) := g(x + e_i) - g(x),$$

and its gradient $\nabla g : \mathbb{Z}^d \to \mathbb{R}^{d \times k}$ by the formula

$$\nabla g(x) = (\nabla_i g(x))_{1 \leq i \leq d} = (\nabla_i g_j(x))_{1 \leq i \leq d, 1 \leq j \leq k}.$$

We denote by ∇_i^* the adjoint gradient defined by the formula $\nabla_i^* g(x) = g(x - e_i) - g(x)$ and the adjoint gradient

$$\nabla^* g(x) = (\nabla_i^* g(x))_{1 \leq i \leq d} = (\nabla_i^* g_j(x))_{1 \leq i \leq d, 1 \leq j \leq k}.$$

We define similarly the divergence, for any function $g : \mathbb{Z}^d \to \mathbb{R}^d$,

$$\nabla \cdot g(x) = - \sum_{i=1}^{d} \nabla_i^* g(x).$$

We extend this definition to a more general class of vector-valued functions as follows: For an integer $k \in \mathbb{N}$, and a function $g = (g_{ij})_{1 \leq i \leq d, 1 \leq j \leq k} : \mathbb{Z}^d \to \mathbb{R}^{d \times k}$, we define $\nabla \cdot g : \mathbb{Z}^d \to \mathbb{R}^k$ by the identity

$$\nabla \cdot g(x) = \left(\sum_{i=1}^{d} g_{i,1}(x) - g_{i,1}(x - e_i), \ldots, \sum_{i=1}^{d} g_{i,k}(x) - g_{i,k}(x - e_i) \right).$$

The Laplacian is then defined by the identity $\Delta = \nabla \cdot \nabla$ and is equivalently given by the explicit formula: for any $g : \mathbb{Z}^d \to \mathbb{R}^k$,

$$\Delta g(x) = \sum_{y \sim x} \left(g(y) - g(x) \right). \tag{A.3}$$

We will consider higher order derivatives as follows: For each integer $n \in \mathbb{N}$, we denote by $\nabla^n g : \mathbb{Z}^d \to \mathbb{R}^{nd \times k}$ as follows

$$\nabla^n g(x) := \left(\nabla_{i_1} \cdots \nabla_{i_n} g(x) \right)_{1 \leq i_1, \ldots, i_n \leq d}.$$

We will also denote by $\Delta^n = \Delta \circ \ldots \circ \Delta$ the Laplacian operator iterated n-times. We note that these discrete operators have range n and $2n$ respectively, i.e., given a point $x \in \mathbb{Z}^d$ and a function $u : \mathbb{Z}^d \to \mathbb{R}^k$ one can compute the value of $\nabla^n u(x)$ (resp. $\Delta^n u(x)$) by knowing only the values of u inside the ball $B(x, n)$ (resp. $B(x, 2n)$).

For each function $g : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}^k$, we denote by ∇_x the discrete gradient with respect to the first variable and by and ∇_y the discrete gradient with respect to the second variable. Formally, for each point $(x, y) \in \mathbb{Z}^d \times \mathbb{Z}^d$, we write

$$\nabla_x g(x, y) = \left(g_j(x + e_i, y) - g_j(x, y) \right)_{1 \leq i \leq d, 1 \leq j \leq k} \text{ and } \nabla_y g(x, y) = \left(g_j(x, y + e_i) - g_j(x, y) \right)_{1 \leq i \leq d, 1 \leq j \leq k}.$$

We similarly define the i-th derivatives $\nabla_{i,x}$ and $\nabla_{i,y}$ and the Laplacians Δ_x and Δ_y, and the higher-order derivatives $\nabla_{i,x}^n$ and $\nabla_{i,y}^n$ with respect to the first and second variables.

Given a vector $p \in \mathbb{R}^{d \times k}$, we write $p = (p_1, \ldots, p_k)$ where the components $p_1, \ldots, p_k$ belong to the space $\mathbb{R}^d$. We denote by l_p the affine function defined by the formula

$$l_p := \begin{cases} \mathbb{Z}^d \to \mathbb{R}^k, \\ x \mapsto (p_1 \cdot x, \ldots, p_k \cdot x). \end{cases} \tag{A.4}$$

This notation will be frequently used in the case $k = \binom{d}{2}$. For $i, j \in \{1, \ldots, d\} \times \{1, \ldots, \binom{d}{2}\}$, we will use the notation l_{ij} for the affine functions

$$l_{ij} := \begin{cases} \mathbb{R}^d \to \mathbb{R}^{\binom{d}{2}}, \\ x \mapsto (0, \ldots, 0, x \cdot e_i, 0, \ldots, 0), \end{cases}$$

where the term $x \cdot e_i$ appears in the j-th position.

A.1.3. *Notation for matrix-valued functions.* A tool frequently used in this article is the notion of fundamental solution for system of elliptic equations (and in particular for the Helffer-Sjöstrand equation), which requires to introduce matrix-valued function. Given an pair of integers $k, l \in \mathbb{N}$, we may identify the vector space $\mathbb{R}^{k \times l}$ with the space of $(k \times l)$-matrices with real coefficients. Given a map $F : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}^{k \times l}$, we denote its components by $(F_{ij})_{1 \leq i \leq k, 1 \leq j \leq l}$. For each integer $i \in \{1, \ldots, k\}$, we denote by $F_{i\cdot}$ the map

$$F_{i\cdot} : \begin{cases} \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}^l, \\ (x, y) \mapsto (F_{ij}(x, y))_{1 \leq j \leq l}. \end{cases}$$

We similarly define the map $F_{\cdot j} : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}^k$ for each integer $j \in \{1, \ldots, l\}$. As in Section A.1.2, we define the gradients $\nabla_x F : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}^{(d \times k) \times l}$ and $\nabla_y F : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}^{k \times (d \times l)}$ with respect to the first and second variables.

A.1.4. *Scalar and matrix product.* We present in this section an index of the intrinsic scalar products used in the article.

Given two functions $f, g : \mathbb{Z}^d \to \mathbb{R}^k$ and a point $x \in \mathbb{Z}^d$, we will use the notation

$$\begin{cases} f(x)g(x) = \sum_{i=1}^{k} f_i(x) g_i(x) \\ \nabla f(x) \nabla g(x) = \sum_{i=1}^{d} \sum_{j=1}^{k} \nabla_i f_j(x) \nabla_i g_j(x) \\ \nabla f(x) g(x) = \left(\sum_{j=1}^{k} \nabla_i f_j(x) g_j(x) \right)_{1 \le i \le d}. \end{cases}$$

Given a matrix-valued function $F : \mathbb{Z}^d \to \mathbb{R}^{k \times l}$, two functions $f : \mathbb{Z}^d \to \mathbb{R}^k$ and $g : \mathbb{Z}^d \to \mathbb{R}^l$, and $x, y \in \mathbb{Z}^d$, we denote by

$$\begin{cases} F(x,y) f(x) = \left(\sum_{i=1}^{k} F_{ij}(x,y) f_i(x) \right)_{1 \le j \le l} \\ F(x,y) g(y) = \left(\sum_{j=1}^{l} F_{ij}(x,y) g_j(y) \right)_{1 \le i \le k} \\ f(x) F(x,y) g(y) = \sum_{i=1}^{k} \sum_{j=1}^{l} F_{ij}(x,y) f_i(x) g_j(y). \end{cases}$$

Similarly, for $f : \mathbb{Z}^d \to \mathbb{R}^{d \times k}$ and $g : \mathbb{Z}^d \to \mathbb{R}^{d \times l}$, we write

$$\begin{cases} \nabla_x F(x,y) f(x) = \left(\sum_{i_1=1}^{d} \sum_{i=1}^{k} \nabla_{x,i_1} F_{ij}(x,y) f_{i_1 i}(x) \right)_{1 \le j \le l} \\ \nabla_y F(x,y) g(y) = \left(\sum_{j_1=1}^{d} \sum_{j=1}^{l} \nabla_{y,j_1} F_{ij}(x,y) g_{j_1 j}(y) \right)_{1 \le i \le k} \\ f(x) \nabla_x \nabla_y F(x,y) g(y) = \sum_{i_1=1}^{d} \sum_{j_1=1}^{d} \sum_{i=1}^{k} \sum_{j=1}^{l} \nabla_{x,i_1} \nabla_{y,j_1} F_{ij}(x,y) f_{i_1 i}(x) g_{j_1 j}(y). \end{cases}$$

As it will be useful when dealing with the second-order Helffer-Sjöstrand equation, we extend these definitions to functions defined on $\mathbb{Z}^{2d}$ and $\mathbb{Z}^{2d} \times \mathbb{Z}^{2d}$ (see also Section A.2.2).

Given $f : \mathbb{Z}^d \to \mathbb{R}^k$ and $h : \mathbb{Z}^d \to \mathbb{R}^l$, we denote by $f \otimes g : \mathbb{Z}^d \times \mathbb{Z}^d \to \mathbb{R}^{k \times l}$ the tensor product between the two functions f and g; it is defined by the formula, for each $x \in \mathbb{Z}^d$,

$$f \otimes g(x) := \left(f_i(x) g_j(x) \right)_{1 \le i \le k, 1 \le j \le l}. \tag{A.5}$$

This notation allows to expand gradients of products of functions: for each function $u : \mathbb{Z}^d \to \mathbb{R}$, one has

$$\nabla (ug)(x) = \nabla u \otimes g(x) + u(x) \nabla g(x). \tag{A.6}$$

For $x, y \in \mathbb{Z}^d$, we will also use the notation

$$f(x) \otimes g(y) := \left(f_i(x) g_j(y) \right)_{1 \le i \le k, 1 \le j \le l}.$$

A.1.5. *Norms and functional spaces.* We define the L^2-scalar product $(\cdot, \cdot)$ according to the formula

$$(f, g) = \sum_{x \in \mathbb{Z}^d} f(x) g(x), \tag{A.7}$$

We restrict this scalar product to a set $U \subseteq \mathbb{Z}^d$ and define, for any pair of functions $f, g : U \to \mathbb{R}^k$,

$$(f, g)_U := \sum_{x \in U} f(x) g(x). \tag{A.8}$$

Given a bounded subset $U \subseteq \mathbb{Z}^d$, we define the average of g over the set U by the formula

$$(g)_U := \frac{1}{|U|} \sum_{x \in U} g(x) \in \mathbb{R}^k.$$

For each subset $U \subseteq \mathbb{Z}^d$, we define the $L^\infty(U)$-norm

$$\|g\|_{L^\infty(U)} := \sup_{x\in U} |g(x)|.$$

where the notation $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^k$. Given a bounded subset $U \subseteq \mathbb{Z}^d$, we denote by $\underline{L}^p(U)$ the normalized norms

$$\|g\|_{\underline{L}^p(U)} := \left(\frac{1}{|U|} \sum_{x\in\mathbb{Z}^d} |g(x)|^p \right)^{\frac{1}{p}}.$$

We introduce the normalized Sobolev norms $\underline{H}^1(U)$ and $\underline{H}^{-1}(U)$ by the formulae

$$\|g\|_{\underline{H}^1(U)} := \frac{1}{(\operatorname{diam} U)} \|g\|_{\underline{L}^2(U)} + \|\nabla g\|_{\underline{L}^2(U)} \quad \text{and} \quad \|g\|_{\underline{H}^{-1}(U)} := \left\{ (f,g)_U \,:\, f : U \to \mathbb{R}^k,\, \|f\|_{\underline{H}^1(U)} \leq 1 \right\}.$$

We denote by $H_0^1(U)$ the set of functions from U to $\mathbb{R}^k$ which are equal to 0 outside the set U (by analogy to the Sobolev space). We implicitly extend the functions of $H_0^1(U)$ by the value 0 to the entire lattice $\mathbb{Z}^d$.

A.1.6. *Notation for the parabolic problem.* In Section 5, we study the solutions of parabolic equations. We introduce in this section a few definitions and notation pertaining to this setting. For $s > 0$ and $t \in \mathbb{R}$, we define the time intervals $I_s := (-s, 0]$ and $I_s(t) := (-s+t, t]$. Given a point $x \in \mathbb{Z}^d$ and a radius $r > 0$, we denote the parabolic cylinder by $Q_r(t,x) := I_{r^2}(t) \times B(x,r)$ (where $B(x,r)$ is the discrete ball). To simplify the notation, we write Q_r to mean $Q_r(0,0)$.

A.1.7. *Notation for Gibbs measures.* We let Ω be the set of vector-valued functions $\phi : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2}}$. Given a cube $\square \subseteq \mathbb{Z}^d$, we let $\Omega_0(\square)$ be the set of vector-valued functions $\phi : \square \to \mathbb{R}^{\binom{d}{2}}$ such that $\phi = 0$ on $\partial\square$. Given $z \in \mathbb{Z}^d$, we define $\tau_z : \Omega \to \Omega$ to be the translation: $\tau_z\phi(\cdot) = \phi(z + \cdot)$.

Given an inverse temperature $\beta > 0$, a probability measure μ_β on Ω and measurable function $X : \Omega \to \mathbb{R}$ which is either nonnegative or integrable with respect to the measure μ_β, we denote its expectation and variance by $\langle X \rangle_{\mu_\beta}$ and $\operatorname{var}_{\mu_\beta}[X]$. We define the $L^2(\mu_\beta)$-norm of the random variable X according to the formula

$$\|X\|_{L^2(\mu_\beta)} := \left(\int_\Omega |X(\phi)|^2 \mu_\beta(d\phi) \right)^{\frac{1}{2}}.$$

For each point $x \in \mathbb{Z}^d$ and each integer $i \in \left\{1, \ldots, \binom{d}{2}\right\}$, we let $\omega_{x,i}$ be the function

$$\omega_{x,i}(y) := \begin{cases} e_i & \text{if } x = y \\ 0 & \text{if } x \neq y, \end{cases}$$

where $\left(e_1, \ldots, e_{\binom{d}{2}}\right)$ is the canonical basis of $\mathbb{R}^{\binom{d}{2}}$. We define the differential operators $\partial_{x,i}$ and ∂_x by the formulae

$$\partial_{x,i} u(\phi) := \lim_{h\to 0} \frac{u(\phi + h\omega_{x,i}) - u(\phi)}{h} \quad \text{and} \quad \partial_x u(\phi) = \left(\partial_{x,1} u, \ldots, \partial_{x,\binom{d}{2}} u \right).$$

We let $C^\infty_{\mathrm{loc}}(\Omega)$ be the set of smooth, local and compactly supported functions of the set Ω. We define the space $H^1(\mu_\beta)$ to be the closure of the space $C^\infty_{\mathrm{loc}}(\Omega)$ with respect to the norm (rescaled with respect to the inverse temperature β)

$$\|u\|_{H^1(\mu_\beta)} := \|u\|_{L^2(\mu_\beta)} + \left(\beta \sum_{x\in\mathbb{Z}^d} \|\partial_x u\|^2_{L^2(\mu_\beta)} \right)^{\frac{1}{2}}.$$

For any subset $U \subseteq \mathbb{Z}^d$, we let $L^2(U, \mu_\beta)$ to be the set of measurable functions $u : \mathbb{Z}^d \times \Omega \to \mathbb{R}^k$ which satisfy

$$\|u\|_{L^2(U,\mu_\beta)} := \left(\sum_{x\in U} \|u(x,\cdot)\|^2_{L^2(\mu_\beta)} \right)^{\frac{1}{2}} < \infty.$$

When the set U is finite, we define the normalized $\underline{L}^2(U, \mu_\beta)$ semi-norm by the formula

$$\|u\|_{\underline{L}^2(U,\mu_\beta)} := \left(\frac{1}{|U|} \sum_{x\in U} \|u(x,\cdot)\|^2_{L^2(\mu_\beta)} \right)^{\frac{1}{2}}, \tag{A.9}$$

as well as the space and space-field averages for: $u : \mathbb{Z}^d \times \Omega \to \mathbb{R}^{\binom{d}{2}}$ and $\phi \in \Omega$,

$$(u)_U(\phi) = \frac{1}{|U|}\sum_{x\in U} u(x,\phi) \quad \text{and} \quad (u)_{U,\mu_\beta} = \frac{1}{|U|}\sum_{x\in U} \langle u(x,\cdot)\rangle_{\mu_\beta}.$$

We define the norm $H^1(U,\mu_\beta)$ by the formula

$$\|u\|_{H^1(U,\mu_\beta)} := \left(\sum_{x\in U} \|u(x,\cdot)\|^2_{H^1(\mu_\beta)} + \|\nabla u\|^2_{L^2(U,\mu_\beta)}\right)^{\frac12},$$

as well as the normalized $\underline{H}^1(U,\mu_\beta)$-norm

$$\|u\|^2_{\underline{H}^1(U,\mu_\beta)} := \frac{1}{(\operatorname{diam} U)^2|U|}\sum_{x\in U}\|u(x,\cdot)\|^2_{L^2(\mu_\beta)} + \frac{\beta}{|U|}\sum_{y\in\mathbb{Z}^d}\sum_{x\in U}\|\partial_y u(x,\cdot)\|^2_{L^2(\mu_\beta)} + \frac{1}{|U|}\|\nabla u\|^2_{L^2(U,\mu_\beta)}.$$

We define the subset $H^1_0(U,\mu_\beta)$ to be the subset of functions of $H^1(U,\mu_\beta)$ which are equal to 0 on the boundary $\partial U \times \Omega$. We implicitly extend these functions by the value 0 to the space $\mathbb{Z}^d$. In particular, we always think of elements of $H^1_0(U,\mu_\beta)$ as functions defined on the entire space. We introduce the seminorm

$$[\![u]\!]^2_{\underline{H}^1(U,\mu_\beta)} := \frac{\beta}{|U|}\sum_{x\in U, y\in\mathbb{Z}^d}\|\partial_y u(x,\cdot)\|^2_{L^2(\mu_\beta)} + \frac{1}{|U|}\|\nabla u\|^2_{L^2(U,\mu_\beta)}.$$

We define the $\underline{H}^{-1}(U,\mu)$-norm by the formula

$$\|u\|_{\underline{H}^{-1}(U,\mu_\beta)} := \sup\left\{\frac{1}{|U|}\sum_{x\in U}\langle u(x,\cdot)v(x,\cdot)\rangle_{\mu_\beta} \, : \, v\in H^1_0(U,\mu_\beta),\ \|v\|_{\underline{H}^1(U,\mu_\beta)} \le 1\right\}.$$

We next state a Poincaré inequality for $H^1(U,\mu_\beta)$. We give two statements, one for functions which vanish on the boundary of U and another for zero-mean functions in the case U is a cube.

Lemma A.1 (Poincaré inequality for $H^1(U,\mu_\beta)$)**.** *Let $\square_L$ be a cube of size L. There exists $C(d,\beta)<\infty$ such that:*

(i) *For every subset $U\subseteq\square_L$ and $w\in H^1_0(U,\mu_\beta)$,*

$$\|w\|_{L^2(U,\mu_\beta)} \le CL\,[\![w]\!]_{H^1(U,\mu_\beta)}. \tag{A.10}$$

(ii) *For every cube $\square'\subseteq\square_L$ and $w\in H^1(\square',\mu_\beta)$,*

$$\left\|w-(w)_{\square'}\right\|_{L^2(\square',\mu_\beta)} \le CL\,[\![w]\!]_{H^1(\square',\mu_\beta)}. \tag{A.11}$$

Proof. The results are obtained by applying the standard Poincaré's inequalities for each realization of the field $\phi\in\Omega$, and then integrating over the fields. □

A.2. **Discrete differential forms.** For each integer $k\in\{1,\ldots,d\}$, a k-cell of the lattice $\mathbb{Z}^d$ is a set of the form, for a subset $\{i_1,\ldots,i_k\}\subseteq\{1,\ldots,d\}$, and a point $x\in\mathbb{Z}^d$,

$$\left\{x+\sum_{l=1}^k \lambda_l e_{i_l}\in\mathbb{R}^d \, : \, 0\le\lambda_1,\ldots,\lambda_k\le 1\right\}.$$

We equip the set of k-cells with an orientation induced by the canonical orientation of the lattice $\mathbb{Z}^d$ and denote by $\Lambda^k(\mathbb{Z}^d)$ the set of oriented k-cells of the lattice $\mathbb{Z}^d$. Given a k-cell c_k, we denote by ∂c_k the boundary of the cell; it can be decomposed into a disjoint union of $(k-1)$-cells. The values $k=0,1,2$ are of specific interest to us; they correspond to the set of vertices, edges and faces of the lattice $\mathbb{Z}^d$. We will denote these spaces by $V(\mathbb{Z}^d)$, $E(\mathbb{Z}^d)$ and $F(\mathbb{Z}^d)$ respectively.

A.2.1. *Definitions and basic properties.* Given an integer $k\in\{0,\ldots,d\}$, we denote by $\Lambda^k(\mathbb{Z}^d)$ the set of oriented k-cells of the hypercubic lattice $\mathbb{Z}^d$.

For each k-cell c_k, we denote by c_k^{-1} the same k-cell as c_k with reverse orientation and by ∂c_k the boundary this cell. A k-form u is a mapping from $\Lambda^k(\square)$ to $\mathbb{R}$ such that $u\left(c_k^{-1}\right) = -u\left(c_k\right)$.

Given a k-form u, we define its exterior derivative $\mathrm{d}u$ according to the formula, for each oriented $(k+1)$-cell c_{k+1},

$$\mathrm{d}u\left(c_{k+1}\right) = \sum_{c_k\subseteq\partial c_{k+1}} u(c_k), \tag{A.12}$$

where the orientation of the face c_k is given by the orientation of the $(k+1)$-cell c_{k+1}; we set the convention $\mathrm{d}u = 0$ for any d-form u. We define the codifferential d^* according to the formula, for each $(k-1)$-cell c_{k-1} and each k-form $u : \Lambda^k(\square) \to \mathbb{R}$,

$$\mathrm{d}^* u\left(c_{k-1}\right) := \sum_{\partial c_k \ni c_{k-1}} u(c_k). \tag{A.13}$$

Clearly, $\mathrm{d}u$ is a $(k+1)$-form and d^*u is a $(k-1)$-form; we set $\mathrm{d}^*u = 0$ for any 0-form u. One also verifies the properties, for each k-form $u : \Lambda^k(\square) \to \mathbb{R}$, $\mathrm{dd}u = 0$ and $\mathrm{d}^*\mathrm{d}^*u = 0$. For arbitrary k-forms $u, v : \Lambda^k(\mathbb{Z}^d) \to \mathbb{R}$ with finite support, we define the scalar product $(\cdot,\cdot)$ by the formula

$$(u,v) = \sum_{c_k \in \Lambda^k(\mathbb{Z}^d)} u(c_k)v(c_k). \tag{A.14}$$

We may restrict the scalar product $(\cdot,\cdot)$ to forms which are only defined in a cube $\square$; we denote the corresponding scalar product by $(\cdot,\cdot)_\square$. It is defined by the formula, for each pair of k-forms $u, v : \Lambda^k(\square) \to \mathbb{R}$,

$$(u,v) = \sum_{c_k \in \Lambda^k(\square)} u(c_k)v(c_k).$$

The codifferential d^* is the formal adjoint of the exterior derivative d with respect to this scalar product: Given a k-form $u : \Lambda^k(\mathbb{Z}^d) \to \mathbb{R}$ and a $(k+1)$-form $v : \Lambda^{k+1}(\mathbb{Z}^d) \to \mathbb{R}$ with finite supports, one has the identity

$$(\mathrm{d}u, v) = (u, \mathrm{d}^* v)\,. \tag{A.15}$$

For an integer $k \in \{0, \ldots, d-1\}$ and a cube $\square \subseteq \mathbb{Z}^d$, we define the tangential boundary of the cube $\partial_{k,\mathbf{t}}\square$ to be the set of all the k-cells which are included in the boundary of the cube $\square$. Given a k-form $u : \Lambda^k(\square) \to \mathbb{R}$, we define its tangential trace $\mathbf{t}u$ to be the restriction of the form u to the set $\partial_{k,\mathbf{t}}\square$. One has the formula, for each k-form $u : \Lambda^k(\square) \to \mathbb{R}$ such that $\mathbf{t}u = 0$ and each $(k+1)$-form $v : \Lambda^k(\square) \to \mathbb{R}$,

$$(\mathrm{d}u, v)_\square = (u, \mathrm{d}^* v)_\square\,.$$

We will need the classical Poincaré lemma in the discrete setting. In the continuous setting, a proof of this result can be found in [79], in the discrete setting in [27, Lemma 2.2].

Lemma A.2 (Poincaré)**.** *Let $\square \subseteq \mathbb{Z}^d$ be a cube of the lattice $\mathbb{Z}^d$ of sidelength R and k be an integer in the set $\{1, \ldots, d-1\}$. For each k-form $f : \Lambda^k(\square) \to \mathbb{R}$ such that $\mathrm{d}f = 0$ and $\mathbf{t}f = 0$ on the tangential boundary $\partial_{k,\mathbf{t}}\square$, there exists a $(k-1)$-form $u : \Lambda^{k-1}(\square) \to \mathbb{R}$ such that $\mathbf{t}u = 0$ on the tangential boundary $\partial_{k,\mathbf{t}}\square$ and $\mathrm{d}u = f$ in the cube $\square$. Additionally, one can choose the form u such that*

$$\|u\|_{L^2(\square)} \leq CR\,\|f\|_{L^2(\square)}\,.$$

An important role is played by the set of integer-valued, compactly supported forms q which satisfy $\mathrm{d}q = 0$ and have connected support. We denote by $\mathcal{Q}$ the set of these forms, i.e.,

$$\mathcal{Q} := \left\{ q : \mathbb{Z}^d \to \mathbb{Z} \,:\, |\mathrm{supp}\, q| < \infty,\ \mathrm{supp}\, q \text{ is connected and } \mathrm{d}q = 0 \right\}. \tag{A.16}$$

We may restrict our considerations to the charges of $\mathcal{Q}$ whose support is included in a cube $\square \subseteq \mathbb{Z}^d$; to this end, we introduce the notation

$$\mathcal{Q}_\square := \left\{ q : \mathbb{Z}^d \to \mathbb{Z} \,:\, \mathrm{supp}\, q \subseteq \square,\ \mathrm{supp}\, q \text{ is connected and } \mathrm{d}q = 0 \right\}.$$

We will need to use the following version of Lemma A.2 for the forms of the set $\mathcal{Q}$, for which we refer to [27, Lemma 2.2].

Lemma A.3 (Poincaré for integer-valued forms)**.** *Let k be an integer of the set $\{1, \ldots, d-1\}$ and q be a k-form with values in $\mathbb{Z}$ such that $\mathrm{d}q = 0$, then there exists a $(k-1)$-form n_q with values in $\mathbb{Z}$ such that $q = \mathrm{d}n_q$. Moreover, n_q can be chosen such that $\mathrm{supp}\, n_q$ is contained in the smallest hypercube containing the support of q and such that*

$$\|n_q\|_{L^\infty} \leq C\,\|q\|_1\,.$$

As it is useful in the article, we record a series of inequalities satisfied by the charges $q \in \mathcal{Q}$,

$$\left\{ \begin{aligned} \|q\|_{L^\infty} &\leq \|q\|_1 \,, \\ \operatorname{diam} q \leq |\operatorname{supp} q| &\leq \|q\|_1 \,, \\ \operatorname{diam} n_q &\leq C \|q\|_1 \,, \\ |\operatorname{supp} n_q| &\leq C \|q\|_1^d \,, \\ \|n_q\|_{L^1} \leq |\operatorname{supp} n_q| \, \|n_q\|_{L^\infty} &\leq C \|q\|_1^{d+1} \,, \\ \|n_q\|_{L^2} \leq \|n_q\|_{L^1}^{\frac{1}{2}} \|n_q\|_{L^\infty}^{\frac{1}{2}} &\leq C \|q\|_1^{\frac{d}{2}+1} \,. \end{aligned} \right. \tag{A.17}$$

The proofs of the first two inequalities is a consequence of

$$\|q\|_{L^\infty} = \sup_{x \in \mathbb{Z}^d} |q(x)| \leq \sum_{x \in \mathbb{Z}^d} |q(x)| = \|q\|_1$$

and, using that the charge q is integer-valued,

$$\operatorname{diam} q \leq |\operatorname{supp} q| = \sum_{x \in \mathbb{Z}^d} \mathbb{1}_{\{q(x) \neq 0\}} \leq \sum_{x \in \mathbb{Z}^d} |q(x)| \leq \|q\|_1$$

For the third inequality, we note that the sidelength of the smallest integer containing the support of q is smaller than $(C \operatorname{diam} q)$ (for some constant C depending only on the dimension). Since the form n_q is supported in this hypercube, we have $\operatorname{diam} n_q \leq C \operatorname{diam} q \leq C \|q\|_1$. Similarly the cardinality of the support of n_q is smaller than the cardinality of the hypercube, and thus $\operatorname{diam} n_q \leq (C \operatorname{diam} q)^d \leq C \|q\|_1^d$. The last two inequalities are obtained by combining the previous results with interpolation arguments.

Given $x, y \in \mathbb{Z}^d \times \mathbb{Z}^d$, we denote by $\mathcal{Q}_x$ and $\mathcal{Q}_{x,y}$ the set of charges $q \in \mathcal{Q}$ such that the point x and the points x, y belong to the support of n_q respectively, i.e.,

$$\mathcal{Q}_x := \{ q \in \mathcal{Q} \, : \, x \in \operatorname{supp} n_q \} \quad \text{and} \quad \mathcal{Q}_{x,y} := \{ q \in \mathcal{Q} \, : \, x \in \operatorname{supp} n_q \text{ and } y \in \operatorname{supp} n_q \} \,. \tag{A.18}$$

We also define, for any $\square \subseteq \mathbb{Z}^d$,

$$\mathcal{Q}_{\square,x} := \{ q \in \mathcal{Q}_\square \, : \, x \in \operatorname{supp} n_q \} \quad \text{and} \quad \mathcal{Q}_{\square,x,y} := \{ q \in \mathcal{Q}_\square \, : \, x \in \operatorname{supp} n_q \text{ and } y \in \operatorname{supp} n_q \} \,. \tag{A.19}$$

We also record two inequalities involving the sum of charges: for each pair of points $(x,y) \in \mathbb{Z}^d$, each integer $k \in \mathbb{N}$, and each constant $c > 0$, there exists $\beta_0 > 0$ such that for any $\beta \geq \beta_0$,

$$\left\{ \begin{aligned} \sum_{q \in \mathcal{Q}_x} \|q\|_1^k \, e^{-c\sqrt{\beta}\|q\|_1} &\leq C e^{-c_0 \sqrt{\beta}}, \\ \sum_{q \in \mathcal{Q}_{x,y}} \|q\|_1^k \, e^{-c\sqrt{\beta}\|q\|_1} &\leq C e^{-c_0 \sqrt{\beta}|x-y|}, \end{aligned} \right. \tag{A.20}$$

where the constants C, c_0 depend on k, c and the dimension d. To prove the first inequality, we first absorb the polynomial factor by writing

$$\|q\|_1^k \, e^{-c\sqrt{\beta}\|q\|_1} \leq C e^{-c'\sqrt{\beta}\|q\|_1}$$

for some constant $c' \in (0, c)$. We then decompose over the supports of the charges. To this end, let us denote by $\mathcal{A}_x$ the set of the finite connected subsets of $\mathbb{Z}^d$ containing the vertex x. We then write

$$\sum_{q \in \mathcal{Q}_x} e^{-c\sqrt{\beta}\|q\|_1} = \sum_{X \in \mathcal{A}_x} \sum_{\substack{q \in \mathcal{Q}_\square \\ \operatorname{supp} q = X}} e^{-c'\sqrt{\beta}\|q\|_1} = \sum_{X \in \mathcal{A}_x} \sum_{\substack{q \in \mathcal{Q}_\square \\ \operatorname{supp} q = X}} \left(\prod_{x \in X} e^{-c'\sqrt{\beta}|q(x)|} \right).$$

Exchanging the sum and the product, we see that

$$\sum_{\substack{q \in \mathcal{Q}_\square \\ \operatorname{supp} q = X}} \left(\prod_{x \in X} e^{-c'\sqrt{\beta}|q(x)|} \right) \leq \prod_{x \in X} \left(\sum_{q(x)=1}^{\infty} e^{-c'\sqrt{\beta}|q(x)|} \right) = \left(\frac{e^{-c\sqrt{\beta}}}{1 - e^{-c\sqrt{\beta}}} \right)^{|X|}.$$

We thus obtain

$$\sum_{q \in \mathcal{Q}_x} e^{-c\sqrt{\beta}\|q\|_1} \leq \sum_{X \in \mathcal{A}_x} e^{-c\sqrt{\beta}|X|} = \sum_{n=1}^{\infty} \left| \{ X \in \mathcal{A}_x \, : \, |X| = n \} \right| e^{-c\sqrt{\beta} n}.$$

We next note that the number of connected components of the lattice containing a vertex $x \in \mathbb{Z}^d$ and of size $n \in \mathbb{N}$ grows exponentially fast in n, i.e.,

$$|\{X \in \mathcal{A}_x : |X| = n\}| \leq e^{Cn}.$$

Choosing the inverse temperature β large enough (i.e., such that $c\sqrt{\beta} \geq C$), we deduce that

$$\sum_{q \in \mathcal{Q}_x} e^{-c\sqrt{\beta}\|q\|_1} \leq e^{-c_0\sqrt{\beta}}.$$

The proof of the second estimate of (A.20) can be deduced from the first one by noting that, any charge $q \in \mathcal{Q}_{x,y}$ has a diameter larger than $|x-y|$ (and thus $\|q\|_1 \geq |x-y|$ since the charge q is assumed to have connected support) and that $\mathcal{Q}_{x,y} \subseteq \mathcal{Q}_x$. We thus write (increasing the value of β if necessary)

$$\begin{aligned}
\sum_{q \in \mathcal{Q}_{x,y}} \|q\|_1^k e^{-c\sqrt{\beta}\|q\|_1} &= e^{-\frac{c}{2}\sqrt{\beta}|x-y|} \sum_{q \in \mathcal{Q}_{x,y}} \|q\|_1^k e^{-\frac{c}{2}\sqrt{\beta}\|q\|_1} \\
&\leq e^{-\frac{c}{2}\sqrt{\beta}|x-y|} \sum_{q \in \mathcal{Q}_{x}} \|q\|_1^k e^{-\frac{c}{2}\sqrt{\beta}\|q\|_1} \\
&\leq e^{-\frac{c}{2}\sqrt{\beta}|x-y|} e^{-c_0\sqrt{\beta}} \\
&\leq C e^{-c_0\sqrt{\beta}|x-y|}.
\end{aligned}$$

A.2.2. *Differential forms as vector-valued functions.* Given a subset $I = (i_1, \ldots, i_k) \subseteq \{1, \ldots, d\}$ of cardinality k. We denote by $\Lambda^k_I(\mathbb{Z}^d)$ the set of oriented k-cells of the hypercubic lattice $\mathbb{Z}^d$ which are parallel to the vectors $(e_{i_1}, \ldots, e_{i_k})$. This set can be characterized as follows: if we let c_I be the k-cell defined by the formula

$$c_I := \left\{ \sum_{l=1}^k \lambda_l e_{i_l} \in \mathbb{R}^d : 0 \leq \lambda_1, \ldots, \lambda_k \leq 1 \right\},$$

then we have

$$\Lambda^k_I(\mathbb{Z}^d) = \left\{ x + c_I : x \in \mathbb{Z}^d \right\}. \tag{A.21}$$

The identity (A.21) allows to identify the vector space of k-forms to the vector space of functions defined on $\mathbb{Z}^d$ and valued in $\mathbb{R}^{\binom{d}{k}}$ according the procedure described below. Note that there are $\binom{d}{k}$ subsets of $\{1, \ldots, d\}$ of cardinality k and consider an arbitrary enumeration $I_1, \ldots, I_{\binom{d}{k}}$ of these sets. To each k-form $\hat{u} : \Lambda^k(\mathbb{Z}^d) \to \mathbb{R}$, we can associate a vector-valued function $u : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{k}}$ defined by the formula, for each point $x \in \mathbb{Z}^d$,

$$u(x) = \left(\hat{u}\left(x + c_{I_1}\right), \ldots, \hat{u}\left(x + c_{I_{\binom{d}{k}}}\right) \right). \tag{A.22}$$

This identification is enforced in most of the article; in fact, except in Section 3.1, we always work with vector-valued functions instead of differential forms. We use the identification (A.22) to extend the formalism described in Section A.1 to differential forms; we may for instance refer to the gradient of a form, or the Laplacian of a form etc. Reciprocally, we extend the formalism described in Section A.2.1 to vector-valued functions; given a function $u : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{k}}$, we may refer to the exterior derivative, the codifferential and the tangential trace of the function u, which we still denote by $\mathrm{d}u$, d^*u and $\mathbf{t}u$ respectively. We note that the two definitions of the scalar products (A.7) for vector valued functions and (A.14) for differential forms coincide through the identification (A.22).

From the definition of the exterior derivative d and the codifferential d^* given in (A.12) and (A.13) and the identification (A.22), one sees that the differential operators d and d^* are linear functionals of the gradient ∇: for each integer $k \in \{1, \ldots, d\}$, there exist linear maps $L_{k,\mathrm{d}} : \mathbb{R}^{d \times \binom{d}{k}} \to \mathbb{R}^{\binom{d}{k+1}}$ and $L_{k,\mathrm{d}^*} : \mathbb{R}^{d \times \binom{d}{k}} \to \mathbb{R}^{\binom{d}{k-1}}$ such that, for each function $u : \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{k}}$, and each point $x \in \mathbb{Z}^d$,

$$\mathrm{d}u(x) = L_{k,\mathrm{d}}\left(\nabla u(x)\right) \quad \text{and} \quad \mathrm{d}^*u(x) = L_{k,\mathrm{d}^*}\left(\nabla^* u(x)\right). \tag{A.23}$$

Using that linear maps on finite dimensional vector spaces are continuous, we obtain the estimates, for each point $x \in \mathbb{Z}^d$,

$$|\mathrm{d}u(x)| \leq C\,|\nabla u(x)| \quad \text{and} \quad |\mathrm{d}^*u(x)| \leq C\,|\nabla^* u(x)|,$$

for some constant C depending only on the dimension d.

This article frequently deals with functions defined on the space $\mathbb{Z}^d \times \Omega \times \mathbb{Z}^d$ (resp. $\mathbb{Z}^{2d} \times \Omega \times \mathbb{Z}^{2d}$) and valued in $\mathbb{R}^{\binom{d}{2} \times \binom{d}{2}}$ (resp. $\mathbb{R}^{\binom{d}{2}^2 \times \binom{d}{2}^2}$) since these maps correspond to the fundamental solutions of the Helffer-Sjöstrand operator (resp. second-order Helffer-Sjöstrand operator) associated with the dual Villain model.

Given a map $F : \mathbb{Z}^d \times \Omega \times \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2} \times \binom{d}{2}}$, we denote by

$$\mathrm{d}_x F : \mathbb{Z}^d \times \Omega \times \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2} \times \binom{d}{3}}, \qquad \mathrm{d}_y F : \mathbb{Z}^d \times \Omega \times \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{3} \times \binom{d}{2}},$$

the exterior derivative with respect to the first and second variable, and by

$$\mathrm{d}_x^* F : \mathbb{Z}^d \times \Omega \times \mathbb{Z}^d \to \mathbb{R}^{d \times \binom{d}{2}} \qquad \mathrm{d}_y^* F : \mathbb{Z}^d \times \Omega \times \mathbb{Z}^d \to \mathbb{R}^{\binom{d}{2} \times d},$$

the codifferential with respect to the first and second variable respectively. They are defined by the formulae, for each triplet $(x, y, \phi) \in \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega$, and each integer $k \in \left\{1, \ldots, \binom{d}{2}\right\}$,

$$(\mathrm{d}_x F(x, \phi, y))_{\cdot k} = L_{2,\mathrm{d}} \left(\nabla_x F_{\cdot k}(x, \phi, y)\right), \qquad (\mathrm{d}_y F(x, \phi, y))_{k \cdot} = L_{2,\mathrm{d}} \left(\nabla_y F_{k \cdot}(x, \phi, y)\right)$$

and

$$(\mathrm{d}_x^* F(x, y, \phi))_{\cdot k} = L_{2,\mathrm{d}^*} \left(\nabla_x^* F_{\cdot k}(x, y, \phi)\right), \qquad \left(\mathrm{d}_y^* F(x, y, \phi)\right)_{k \cdot} = L_{2,\mathrm{d}^*} \left(\nabla_y^* F_{k \cdot}(x, y, \phi)\right).$$

Similarly, given a function $F : \mathbb{Z}^{2d} \times \Omega \times \mathbb{Z}^{2d} \to \mathbb{R}^{\binom{d}{2}^2 \times \binom{d}{2}^2}$, we define, for each $(x, y, \phi, x_1, y_1) \in \mathbb{Z}^d \times \mathbb{Z}^d \times \Omega \times \mathbb{Z}^d \times \mathbb{Z}^d$, each field $\phi \in \Omega$ and each triplet of integers $i, j, k \in \left\{1, \ldots, \binom{d}{2}\right\}$

$$\begin{cases} (\mathrm{d}_x F(x, x_1, \phi, y, y_1))_{\cdot ijk} = L_{2,\mathrm{d}} \left(\nabla_x F_{\cdot ijk}(x, x_1, \phi, y, y_1)\right), \\ (\mathrm{d}_{x_1} F(x, x_1, \phi, y, y_1))_{i \cdot jk} = L_{2,\mathrm{d}} \left(\nabla_{x_1} F_{i \cdot jk}(x, x_1, \phi, y, y_1)\right), \\ (\mathrm{d}_y F(x, x_1, \phi, y, y_1))_{ij \cdot k} = L_{2,\mathrm{d}} \left(\nabla_y F_{ij \cdot k}(x, x_1, \phi, y, y_1)\right), \\ (\mathrm{d}_{y_1} F(x, x_1, \phi, y, y_1))_{ijk \cdot} = L_{2,\mathrm{d}} \left(\nabla_{y_1} F_{ijk \cdot}(x, x_1, \phi, y, y_1)\right), \end{cases}$$

and similarly

$$\begin{cases} (\mathrm{d}_x^* F(x, x_1, \phi, y, y_1))_{\cdot ijk} = L_{2,\mathrm{d}^*} \left(\nabla_x^* F_{\cdot ijk}(x, x_1, \phi, y, y_1)\right), \\ \left(\mathrm{d}_{x_1}^* F(x, x_1, \phi, y, y_1)\right)_{i \cdot jk} = L_{2,\mathrm{d}^*} \left(\nabla_{x_1}^* F_{i \cdot jk}(x, x_1, \phi, y, y_1)\right), \\ \left(\mathrm{d}_y^* F(x, x_1, \phi, y, y_1)\right)_{ij \cdot k} = L_{2,\mathrm{d}^*} \left(\nabla_y^* F_{ij \cdot k}(x, x_1, \phi, y, y_1)\right), \\ \left(\mathrm{d}_{y_1}^* F(x, x_1, \phi, y, y_1)\right)_{ijk \cdot} = L_{2,\mathrm{d}^*} \left(\nabla_{y_1}^* F_{ijk \cdot}(x, x_1, \phi, y, y_1)\right). \end{cases}$$

We extend these definitions so that we can consider *mixed derivatives*; for instance, we may use the notation $\mathrm{d}_y^* \mathrm{d}_x^* F$ (or any other combination of exterior derivatives and codifferentials). It is clear that as long as the derivatives involve different variables, they commute: we have for instance $\mathrm{d}_y^* \mathrm{d}_x^* F = \mathrm{d}_x^* \mathrm{d}_y^* F$.

We complete this section by recording the Gaffney-Friedrichs inequality which provides an upper bound on the L^2-norm of the gradient of a form in terms of the L^2-norm of its exterior derivative and the codifferential assuming that the tangential trace of the form vanishes.

Proposition A.4 (Gaffney-Friedrichs inequality for cubes)**.** *Let $\square$ be a cube of $\mathbb{Z}^d$. Then there exists a constant $C := C(d) < \infty$ such that for each k-form $u : \Lambda^k(\square) \to \mathbb{R}$ with vanishing tangential trace, we have*

$$\|\nabla u\|_{\underline{L}^2(\square)} \leq C \left(\|\mathrm{d}u\|_{\underline{L}^2(\square)} + \|\mathrm{d}^* u\|_{\underline{L}^2(\square)}\right).$$

The proof of the continuous version of this inequality can be found in [53, 46] or in the monograph [86, Proposition 2.2.3]. We complete this section by proving the solvability of a boundary value problem involving discrete differential forms used in Section 3.1.

Proposition A.5. *For any integer $k \in \{1, \ldots, d-1\}$, any cube $\square \in \mathbb{Z}^d$, and any k-form $q := (q_1, \ldots, q_{\binom{d}{k}}) : \square \to \mathbb{R}^{\binom{d}{k}}$ such that $\mathrm{d}q = 0$ in the cube $\square$ and $\mathbf{t}q = 0$ on the boundary $\partial\square$, there exists a unique solution to the boundary value problem*

$$\begin{cases} \mathrm{d}\mathrm{d}^* w = q \ \text{in } \square, \\ \mathrm{d}w = 0 \ \text{in } \square, \\ \mathbf{t}w = 0 \ \text{on } \partial\square, \\ \mathbf{t}\mathrm{d}^* w = 0 \ \text{on } \partial\,\square. \end{cases} \tag{A.24}$$

If we denote by $w_1, \ldots, w_{\binom{d}{k}}$ the coordinates of the map w, then they solve the following boundary value problem: for each $i \in \left\{1, \ldots, \binom{d}{k}\right\}$, if we denote by $\partial_{I_i} \square$ the subset of faces of the boundary $\partial \square$ which are parallel to the cell c_{I_i}, then we have

$$\begin{cases} -\Delta w_i = q_i \ \text{in } \square, \\ \quad w_i = 0 \ \text{in } \partial_{I_i} \square, \\ \nabla w_i \cdot \mathbf{n} = 0 \ \text{on } \partial \square \setminus \partial_{I_i} \square \,. \end{cases} \tag{A.25}$$

Remark A.6. The boundary condition (A.25) is a combination of the Dirichlet and Neumann boundary conditions: given an integer $i \in \{1, \ldots, \binom{d}{k}\}$, we assign Dirichlet boundary condition to the faces which are parallel to the cell c_{I_i}, and Neumann boundary condition to the faces which are orthogonal to the cell c_{I_i}.

Proof. The boundary value problem (A.24) admits a variational formulation which can be used to prove existence and uniqueness of the solutions. We first define the set of k-forms

$$C_0^k(\square) := \left\{ u : \square \to \mathbb{R}^{\binom{d}{k}} \, : \, \mathrm{d}u = 0 \text{ in } \square \text{ and } \mathbf{t}u = 0 \,\text{on}\, \partial \square \right\}.$$

We then define the energy functional $J_q : C_0^k(\square) \to \mathbb{R}$ according to the formula

$$J_q(u) := \frac{1}{2} \left\| \mathrm{d}^* u \right\|_{L^2(\square)} - (q, u)_{\square} \,.$$

To prove the solvability of the problem (A.24), we prove that there exists unique minimizer to the variational problem

$$\inf_{u \in C_0^k(\square)} J(u).$$

We first use that, by Lemma A.2, there exists a $(k-1)$-form $n_q : \square \to \mathbb{R}^{\binom{d}{k-1}}$ such that $\mathbf{t} n_q = 0$ on $\partial \square$ and $\mathrm{d} n_q = q$ in the cube $\square$. We then perform an integration by parts to write

$$J_q(u) = \frac{1}{2} \left\| \mathrm{d}^* u \right\|_{L^2(\square)} - (n_q, \mathrm{d}^* u)_{\square} \,.$$

The technique then follows the standard strategy of the calculus of variations. The energy functional J_q is bounded from below and we consider a minimizing sequence $(w_n)_{n \in \mathbb{N}}$. It is clear that the norms $\left\| \mathrm{d}^* w_n \right\|_{L^2(\square)}$ are uniformly bounded in $n \in \mathbb{N}$. Using that $\mathrm{d} w_n = 0$ and the Gaffney-Friedrich inequality stated in Proposition A.4, we obtain that the norms $\left\| \nabla w_n \right\|_{L^2(\square)}$ and $\left\| w_n \right\|_{L^2(\square)}$ are uniformly bounded in n. We can thus extract a subsequence which converges in the discrete space $L^2(\square)$ and verify that the limit is solution to the problem (A.24). The uniqueness is a consequence of the uniform convexity of the functional J_q.

To prove (A.25), note that the condition $\mathrm{d}w = 0$ and the identity $\mathrm{d}\delta + \delta \mathrm{d} = -\Delta$ imply that $-\Delta w = q$ in the cube $\square$. Using the definition of the Laplacian for vector-valued function (stated in (A.3)), we have that for each integer $i \in \{1, \ldots, \binom{d}{k}\}$, $-\Delta w_i = q_i$ in the cube $\square$. The boundary condition $\mathbf{t} w = 0$ implies that w_i is equal to 0 on each face which is parallel to the cell c_{I_i}; the condition $\mathbf{t} \mathrm{d}^* w = 0$ implies that the function w_i satisfies a Neumann boundary condition on the faces of the boundary $\partial \square$ which are orthogonal to the cell c_{I_i}. $\square$

Appendix B. Multiscale Poincaré inequality

Proposition B.1 (Multiscale Poincaré inequality)**.** *There exists a constant $C := C(d)$ such that for each cube integer $n \in \mathbb{N}$, the following statements hold:*

(1) *For each function $f \in L^2(\square_n, \mu_\beta)$,*

$$\left\| f - (f)_{\square_n} \right\|^2_{\underline{H}^{-1}(\square_n, \mu_\beta)} \leq C \left\| f \right\|^2_{\underline{L}^2(\square_n, \mu_\beta)} + C 3^n \sum_{m=0}^{n} \frac{3^m}{|\mathcal{Z}_{m,n}|} \sum_{z \in \mathcal{Z}_{m,n}} \left\langle \left(\frac{1}{|z + \square_m|} \sum_{x \in z + \square_m} f(x, \cdot) \right)^2 \right\rangle_{\mu_\beta};$$

(2) *For any function $f \in L^2(\square_n, \mu_\beta)$, one has the estimate*

$$\left\| f - (f)_{\square} \right\|^2_{\underline{L}^2(\square_n, \mu_\beta)} \leq C \left\| \nabla f \right\|^2_{\underline{L}^2(\square_n, \mu_\beta)} + C 3^n \sum_{m=0}^{n} \frac{3^m}{|\mathcal{Z}_{m,n}|} \sum_{z \in \mathcal{Z}_{m,n}} \left\langle \left(\frac{1}{|z + \square_m|} \sum_{x \in z + \square_m} \nabla f(x, \cdot) \right)^2 \right\rangle_{\mu_\beta};$$

(3) *for each function $f \in L^2(\square_n, \mu_\beta)$ such that $f = 0$ on the boundary of the cube $\square_n$*

$$\left\| f \right\|^2_{\underline{L}^2(\square_n, \mu_\beta)} \leq C \left\| \nabla f \right\|^2_{\underline{L}^2(\square_n, \mu_\beta)} + C 3^n \sum_{m=0}^{n} \frac{3^m}{|\mathcal{Z}_{m,n}|} \sum_{z \in \mathcal{Z}_{m,n}} \left\langle \left(\frac{1}{|z + \square_m|} \sum_{x \in z + \square_m} \nabla f(x, \cdot) \right)^2 \right\rangle_{\mu_\beta}.$$

Proof. The proof is an almost immediate application of the multiscale Poincaré inequality proved in [8, Proposition 1.7 and Lemma 1.8]. We only treat the inequality (1); the other two estimates are similar. We consider a field $\phi \in \Omega$ and apply [8, Proposition 1.7 and Lemma 1.8] and a Cauchy-Schwarz inequality to the map $x \to f(x, \phi)$ (with a fixed field ϕ). We obtain

$$\left\| f(\cdot, \phi) - (f(\cdot, \phi))_{\square_n} \right\|^2_{\underline{H}^{-1}(\square_n)} \leq C \left\| f(\cdot, \phi) \right\|^2_{\underline{L}^2(\square_n)} + C 3^n \sum_{m=0}^{n} \frac{3^m}{|\mathcal{Z}_{m,n}|} \sum_{z \in \mathcal{Z}_{m,n}} \left(\frac{1}{|z + \square_m|} \sum_{x \in z + \square_m} f(x, \phi) \right)^2.$$

Taking the expectation with respect to the field ϕ gives

$$\left\langle \left\| f - (f)_{\square_n} \right\|^2_{\underline{H}^{-1}(\square_n)} \right\rangle_{\mu_\beta} \leq C \left\| f \right\|^2_{\underline{L}^2(\square_n)} + C 3^n \sum_{m=0}^{n} \frac{3^m}{|\mathcal{Z}_{m,n}|} \sum_{z \in \mathcal{Z}_{m,n}} \left\langle \left(\frac{1}{|z + \square_m|} \sum_{x \in z + \square_m} f(x, \cdot) \right)^2 \right\rangle_{\mu_\beta}.$$

We complete the proof by using the estimate

$$\left\| f - (f)_{\square_n} \right\|^2_{\underline{H}^{-1}(\square_n, \mu_\beta)} \leq \left\langle \left\| f - (f)_{\square_n} \right\|^2_{\underline{H}^{-1}(\square_n)} \right\rangle_{\mu_\beta},$$

which is a direct consequence of the definitions of the $\underline{H}^{-1}(\square)$ and $\underline{H}^{-1}(\square, \mu_\beta)$-norms stated in Appendix A. □

Appendix C. Basic estimates on discrete convolutions

The objective of this appendix is to collect estimates on some discrete convolutions of functions decaying algebraically fast at infinity. These estimates are used in various places in the article and are elementary; their proof can be found in the long version of this article [36, Appendix C].

Proposition C.1. *Given a pair of exponents $\alpha, \beta > 0$ such that $\alpha + \beta > d$, a small exponent $\varepsilon > 0$, and a point $x \in \mathbb{Z}^d$, then:*

(i) *If $\alpha \in (0, d)$ and $\beta \in (0, d)$,*

$$\sum_{y \in \mathbb{Z}^d} \frac{1}{|y|^\alpha} \frac{1}{|x-y|^\beta} \leq \frac{C}{|x|^{\alpha+\beta-d}};$$

(ii) *If $\alpha = d$ and $\beta \in (0, d]$,*

$$\sum_{y \in \mathbb{Z}^d} \frac{1}{|y|^\alpha} \frac{1}{|x-y|^\beta} \leq \frac{C \ln |x|}{|x|^\beta};$$

(iii) *If $\alpha > d$ and $\beta \in (0, \infty)$,*

$$\sum_{y \in \mathbb{Z}^d} \frac{1}{|y|^\alpha} \frac{1}{|x-y|^\beta} \leq \frac{C}{|x|^{\min(\alpha,\beta)}};$$

(iv) *One has the estimate*

$$\sum_{z_1, z_2 \in \mathbb{Z}^d} \frac{1}{|x-z_1|^d} \frac{1}{|z_1 - z_2|^{d-\varepsilon}} \frac{1}{|z_2|^{d-1}} \leq \frac{C \ln |x|}{|x|^{d-1-\varepsilon}};$$

(v) *One has the estimate*

$$\sum_{z_1, z_2 \in \mathbb{Z}^d} \frac{1}{|x-z_1|^{d-1}} \frac{1}{|z_1 - z_2|^{d-\varepsilon}} \frac{1}{|z_2|^d} \leq \frac{C \ln |x|}{|x|^{d-1-\varepsilon}};$$

and equivalently

$$\sum_{y, z \in \mathbb{Z}^d} \frac{1}{|y|^{d-1}|x-y|^{d-1}} \frac{1}{|z|^{d-1}|z-x|^{d-1}} \frac{1}{|y-z|^{d-\varepsilon}} \leq \frac{C}{|x|^{2d-2}};$$

(vii) *For each exponent $\alpha > d$,*

$$\sum_{y \in \mathbb{Z}^d} \frac{1}{|y|^\alpha + |y-x|^\alpha} \leq \frac{C}{|x|^{\alpha-d}},$$

where the constant C depends on the parameters α and d. A variation of the proof gives the following generalization of (C.1): for every cube $\square \subseteq \mathbb{Z}^d$ of center 0 and sidelength $R \geq 1$, and every point $y \in \mathbb{Z}^d$,

$$\sum_{y_0 \in \mathbb{Z}^d \setminus \square} \frac{1}{|y_0|^\alpha + |y_0 - y|^\alpha} \leq \frac{C}{\max(R, |y|)^{\alpha-d}}.$$

References

[1] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.

[2] S. Adams, S. Buchholz, R. Kotecký, and S. Müller. Cauchy-born rule from microscopic models with non-convex potentials. *arXiv preprint arXiv:1910.13564*, 2019.

[3] S. Adams, R. Kotecký, and S. Müller. Strict convexity of the surface tension for non-convex potentials. *arXiv preprint arXiv:1606.09541*, 2016.

[4] M. Aizenman, M. Harel, R. Peled, and J. Shapiro. Depinning in the integer-valued Gaussian field and the BKT phase of the 2D Villain model. *arXiv preprint arXiv:2110.09498*, 2021.

[5] R. Alicandro and M. Cicalese. Variational analysis of the asymptotics of the XY model. *Arch. Rational Mech. Anal.*, 192(3):501–536, 2009.

[6] S. Armstrong, T. Kuusi, and J.-C. Mourrat. Mesoscopic higher regularity and subadditivity in elliptic homogenization. *Comm. Math. Phys.*, 347(2):315–361, 2016.

[7] S. Armstrong, T. Kuusi, and J.-C. Mourrat. The additive structure of elliptic homogenization. *Invent. Math.*, 208(3):999–1154, 2017.

[8] S. Armstrong, T. Kuusi, and J.-C. Mourrat. *Quantitative stochastic homogenization and large-scale regularity*, volume 352 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Cham, 2019.

[9] S. Armstrong and W. Wu. C^2 regularity of the surface tension for the $\nabla\phi$ interface model. *Communications on Pure and Applied Mathematics*, 75(2):349–421, 2022.

[10] 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.
[11] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bull. Amer. Math. Soc.*, 73:890–896, 1967.
[12] M. Avellaneda and F.-H. Lin. Compactness methods in the theory of homogenization. *Comm. Pure Appl. Math.*, 40(6):803–847, 1987.
[13] M. Avellaneda and F.-H. Lin. L^p bounds on singular integrals in homogenization. *Comm. Pure Appl. Math.*, 44(8-9):897–910, 1991.
[14] T. Balaban, J. Imbrie, and A. Jaffe. The mass gap for Higgs models on a unit lattice. *Ann. Physics*, 158(2):281–319, 1984.
[15] T. Balaban and M. O'Carroll. Low temperature properties for correlation functions in classical N-vector spin models. *Comm. Math. Phys.*, 199(3):493–520, 1999.
[16] R. Bauerschmidt. Ferromagnetic spin systems. *lecture notes available at http://www.statslab.cam.ac.uk/ rb812/doc/ spin.pdf.*
[17] R. Bauerschmidt, J. Park, and P.-F. Rodriguez. The Discrete Gaussian model, I. Renormalisation group flow at high temperature. *arXiv preprint arXiv:2202.02286*, 2022.
[18] R. Bauerschmidt, J. Park, and P.-F. Rodriguez. The Discrete Gaussian model, II. Infinite-volume scaling limit at high temperature. *arXiv preprint arXiv:2202.02287*, 2022.
[19] P. Bella and A. Giunti. Green's function for elliptic systems: moment bounds. *Netw. Heterog. Media*, 13(1):155–176, 2018.
[20] P. Bella, A. Giunti, and F. Otto. Quantitative stochastic homogenization: local control of homogenization error through corrector. In *Mathematics and materials*, volume 23 of *IAS/Park City Math. Ser.*, pages 301–327. Amer. Math. Soc., Providence, RI, 2017.
[21] M. Biskup and R. Kotecký. Phase coexistence of gradient Gibbs states. *Probab. Theory and Related Fields*, 139(1):1–39, 2007.
[22] M. Biskup and H. Spohn. Scaling limit for a class of gradient fields with nonconvex potentials. *Ann. Probab.*, 39(1):224–251, 2011.
[23] S. G. Bobkov and M. Ledoux. From Brunn-Minkowski to Brascamp-Lieb and to logarithmic Sobolev inequalities. *Geom. Funct. Anal.*, 10(5):1028–1052, 2000.
[24] H. J. Brascamp and E. H. Lieb. On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *J. Functional Analysis*, 22(4):366–389, 1976.
[25] J. Bricmont, J. R. Fontaine, and L. J. Landau. On the uniqueness of the equilibrium state for plane rotators. *Comm. Math. Phys.*, 56(3):281–296, 1977.
[26] J. Bricmont, J.-R. Fontaine, J. L. Lebowitz, E. H. Lieb, and T. Spencer. Lattice systems with a continuous symmetry. III. Low temperature asymptotic expansion for the plane rotator model. *Comm. Math. Phys.*, 78(4):545–566, 1980/81.
[27] S. Chatterjee. Wilson loops in Ising lattice gauge theory. *Comm. Math. Phys.*, 377(1):307–340, 2020.
[28] J. G. Conlon, A. Giunti, and F. Otto. Green's function for elliptic systems: existence and Delmotte-Deuschel bounds. *Calc. Var. Partial Differential Equations*, 56(6):Art. 163, 51, 2017.
[29] J. G. Conlon and T. Spencer. A strong central limit theorem for a class of random surfaces. *Comm. Math. Phys.*, 325(1):1–15, 2014.
[30] C. Cotar and J.-D. Deuschel. Decay of covariances, uniqueness of ergodic component and scaling limit for a class of $\nabla\phi$ systems with non-convex potential. *Ann. Inst. Henri Poincaré Probab. Stat.*, 48(3):819–853, 2012.
[31] C. Cotar, J.-D. Deuschel, and S. Müller. Strict convexity of the free energy for a class of non-convex gradient models. *Comm. Math. Phys.*, 286(1):359–376, 2009.
[32] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization. *Ann. Mat. Pura Appl. (4)*, 144:347–389, 1986.
[33] G. Dal Maso and L. Modica. Nonlinear stochastic homogenization and ergodic theory. *J. Reine Angew. Math.*, 368:28–42, 1986.
[34] P. Dario. Quantitative homogenization of the disordered $\nabla\phi$ model. *Electron. J. Probab.*, 24:Paper No. 90, 99, 2019.
[35] P. Dario. Quantitative homogenization of differential forms. *Ann. Inst. Henri Poincaré Probab. Stat.*, 57(2):1157–1202, 2021.
[36] P. Dario and W. Wu. Massless phases for the Villain model in $d \geq 3$. *arXiv preprint arXiv:2002.02946*, 2020.
[37] E. B. Davies. Explicit constants for Gaussian upper bounds on heat kernels. *Amer. J. Math.*, 109(2):319–333, 1987.
[38] E. B. Davies. Heat kernel bounds for second order elliptic operators on Riemannian manifolds. *Amer. J. Math.*, 109(3):545–569, 1987.
[39] E. De Giorgi. Sulla differenziabilità e l'analiticità delle estremali degli integrali multipli regolari. *Mem. Accad. Sci. Torino. Cl. Sci. Fis. Mat. Nat. (3)*, 3:25–43, 1957.
[40] E. De Giorgi. Un esempio di estremali discontinue per un problema variazionale di tipo ellittico. *Boll. Un. Mat. Ital. (4)*, 1:135–137, 1968.
[41] T. Delmotte and J.-D. Deuschel. On estimating the derivatives of symmetric diffusions in stationary random environment, with applications to $\nabla\phi$ interface model. *Probab. Theory Related Fields*, 133(3):358–390, 2005.
[42] F. J. Dyson. General theory of spin-wave interactions. *Phys. Rev. (2)*, 102:1217–1230, 1956.
[43] E. B. Fabes and D. W. Stroock. A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash. *Arch. Rational Mech. Anal.*, 96(4):327–338, 1986.
[44] P. Falco. Kosterlitz-Thouless transition line for the two dimensional Coulomb gas. *Comm. Math. Phys.*, 312(2):559–609, 2012.
[45] P. Falco. Critical exponents of the two dimensional Coulomb gas at the Berezinskii-Kosterlitz-Thouless transition. *arXiv preprint arXiv:1311.2237*, 2013.
[46] K. O. Friedrichs. Differential forms on Riemannian manifolds. *Comm. Pure Appl. Math.*, 8:551–590, 1955.
[47] J. Fröhlich, B. Simon, and T. Spencer. Infrared bounds, phase transitions and continuous symmetry breaking. *Comm. Math. Phys.*, 50(1):79–95, 1976.
[48] J. Fröhlich and T. Spencer. Kosterlitz-Thouless transition in the two-dimensional plane rotator and Coulomb gas. *Phys. Rev. Lett.*, 46(15):1006–1009, 1981.

[49] J. Fröhlich and T. Spencer. On the statistical mechanics of classical Coulomb and dipole gases. *J. Stat. Phys.*, 24(4):617–701, 1981.
[50] J. Fröhlich and T. Spencer. Massless phases and symmetry restoration in abelian gauge theories and spin systems. *Comm. Math. Phys.*, 83(3):411–454, 1982.
[51] T. Funaki. Stochastic interface models. In *Lectures on probability theory and statistics*, volume 1869 of *Lecture Notes in Math.*, pages 103–274. Springer, Berlin, 2005.
[52] T. Funaki and H. Spohn. Motion by mean curvature from the Ginzburg-Landau $\nabla\phi$ interface model. *Comm. Math. Phys.*, 185(1):1–36, 1997.
[53] M. P. Gaffney. The harmonic operator for exterior differential forms. *Proc. Nat. Acad. Sci. U.S.A.*, 37:48–50, 1951.
[54] G. Giacomin, S. Olla, and H. Spohn. Equilibrium fluctuations for $\nabla\phi$ interface model. *Ann. Probab.*, 29(3):1138–1172, 2001.
[55] M. Giaquinta. *Multiple integrals in the calculus of variations and nonlinear elliptic systems*, volume 105 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1983.
[56] 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.
[57] J. Ginibre. General formulation of Griffiths' inequalities. *Comm. Math. Phys.*, 16:310–328, 1970.
[58] 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.
[59] A. Gloria, S. Neukamm, and F. Otto. A regularity theory for random elliptic operators. *Milan Journal of Mathematics*, 88(1):99–170, 2020.
[60] A. Gloria and F. Otto. An optimal variance estimate in stochastic homogenization of discrete elliptic equations. *Ann. Probab.*, 39(3):779–856, 2011.
[61] A. Gloria and F. Otto. An optimal error estimate in stochastic homogenization of discrete elliptic equations. *Ann. Appl. Probab.*, 22(1):1–28, 2012.
[62] A. Gloria and F. Otto. The corrector in stochastic homogenization: optimal rates, stochastic integrability, and fluctuations. *arXiv preprint arXiv:1510.08290*, 2015.
[63] A. Gloria and F. Otto. Quantitative results on the corrector equation in stochastic homogenization. *J. Eur. Math. Soc. (JEMS)*, 19(11):3489–3548, 2017.
[64] M. Guo, G. Papanicolaou, and S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. *Comm. Math. Phys.*, 118(1):31–59, 1988.
[65] A. H. Guth. Existence proof of a nonconfining phase in four-dimensional U(1) lattice gauge theory. *Phys. Rev. D (3)*, 21(8):2291–2307, 1980.
[66] Q. Han and F. Lin. *Elliptic partial differential equations*, volume 1 of *Courant Lecture Notes in Mathematics*. Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, second edition, 2011.
[67] B. Helffer and J. Sjöstrand. On the correlation for Kac-like models in the convex case. *J. Statist. Phys.*, 74(1-2):349–409, 1994.
[68] S. Hofmann and S. Kim. Gaussian estimates for fundamental solutions to certain parabolic systems. *Publ. Mat.*, 48(2):481–496, 2004.
[69] T. Kennedy and C. King. Spontaneous symmetry breakdown in the abelian Higgs model. *Comm. Math. Phys.*, 104(2):327–347, 1986.
[70] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.
[71] J. M. Kosterlitz and D. J. Thouless. Ordering, metastability and phase transitions in two-dimensional systems. *Journal of Physics C: Solid State Physics*, 6(7):1181, 1973.
[72] R. Koteckỳ and S. Luckhaus. Nonlinear elastic free energies and gradient Young-Gibbs measures. *Comm. Math. Phys.*, 326(3):887–917, 2014.
[73] S. M. Kozlov. Averaging of differential operators with almost periodic rapidly oscillating coefficients. *Mat. Sb. (N.S.)*, 107(149)(2):199–217, 317, 1978.
[74] P. Lammers. Height function delocalisation on cubic planar graphs. *Probab. Theory Related Fields*, 182(1):531–550, 2022.
[75] G. F. Lawler and V. Limic. *Random walk: a modern introduction*, volume 123 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
[76] O. A. McBryan and T. Spencer. On the decay of correlations in SO(n)-symmetric ferromagnets. *Comm. Math. Phys.*, 53(3):299–302, 1977.
[77] N. D. Mermin and H. Wagner. Absence of ferromagnetism or antiferromagnetism in one-or two-dimensional isotropic Heisenberg models. *Physical Review Letters*, 17(22):1133, 1966.
[78] A. Messager, S. Miracle-Sole, and C. Pfister. Correlation inequalities and uniqueness of the equilibrium state for the plane rotator ferromagnetic model. *Comm. Math. Phys.*, 58(1):19–29, 1978.
[79] D. Mitrea, M. Mitrea, and S. Monniaux. The Poisson problem for the exterior derivative operator with Dirichlet boundary condition in nonsmooth domains. *Commun. Pure Appl. Anal.*, 7(6):1295–1333, 2008.
[80] J.-C. Mourrat and F. Otto. Correlation structure of the corrector in stochastic homogenization. *Ann. Probab.*, 44(5):3207–3233, 2016.
[81] A. Naddaf and T. Spencer. On homogenization and scaling limit of some gradient perturbations of a massless free field. *Comm. Math. Phys.*, 183(1):55–84, 1997.
[82] A. Naddaf and T. Spencer. Estimates on the variance of some homogenization problems, 1998, unpublished preprint.
[83] J. Nash. Continuity of solutions of parabolic and elliptic equations. *Amer. J. Math.*, 80:931–954, 1958.
[84] C. M. Newman and W. Wu. Lee–Yang property and Gaussian multiplicative chaos. *Comm. Math. Phys.*, 369(1):153–170, 2019.

[85] 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.
[86] G. Schwarz. *Hodge Decomposition-A method for solving boundary value problems*. Springer, 2006.
[87] S. Sheffield. Random surfaces. *Astérisque*, (304):vi+175, 2005.
[88] J. Sjöstrand. Correlation asymptotics and Witten Laplacians. *Algebra i Analiz*, 8(1):160–191, 1996.
[89] D. van Engelenburg and M. Lis. An elementary proof of phase transition in the planar XY model. *arXiv preprint arXiv:2110.09465*, 2021.
[90] J. Villain. Theory of one-and two-dimensional magnets with an easy magnetization plane. II. The planar, classical, two-dimensional magnet. *Journal de Physique*, 36(6):581–590, 1975.
[91] V. V. Yurinskiĭ. Averaging of symmetric diffusion in a random medium. *Sibirsk. Mat. Zh.*, 27(4):167–180, 215, 1986.