



About a nonideal Rayleigh gas mixture model

CONVERGENCE OF THE CORRELATION FUNCTIONS

—
April 2026

Florent Fougères



SAPIENZA
UNIVERSITÀ DI ROMA

1 Introduction: model and results

1.1 Mathematical context

The Boltzmann equation provides a statistical model for rarefied gas dynamics. This kinetic representation may be thought about as an intermediate mesoscopic step between microscopic and macroscopic scales, and might thus be used to derive fluid mechanics equations from Newton classical equations (see for example [4, 18, 26]). This model both allows numerous quantitative simulations [31, 16] and qualitative understanding of the intrinsic statistical behaviour of fluid matter. Indeed, introducing the *entropy* of the system as a Lyapunov functional of the system, Boltzmann showed that the density of particles irreversibly converges towards an equilibrium on large time scales. This equilibrium is called the Maxwellian state: the gas occupies uniformly the available space, and the velocities of the particles get distributed according to the Maxwell Gaussian distribution (1.6), only depending on the temperature of the system.

The rigorous derivation of this equation from microscopic Newton equations has eventually been proved mathematically in 1975 by Oscar Erasmus Lanford III [20, 21], in the case of hard sphere interactions. The method that Lanford used is very rigid in the tracking of particles' trajectories, hindering to prove his result for time scales larger than very small times, when only about a fifth of particles have collided. The major obstruction to large time scales is the correlation occurring between colliding particles: the quantitative estimates on the system's chaoticity deteriorate over time, making it impossible to deal with recollisions of particles.

Nonetheless, in a setting close to thermodynamic equilibrium (the Maxwellian state), the statistical stability of the dynamics guarantees a certain amount of chaos over large times, allowing a very strong control of the correlations. In the Rayleigh gas model, describing the behaviour of a small fraction of tagged particles near equilibrium [27], a proof relying on the same ideas as Lanford's yielded the derivation *on large time scales* of a linear version of the Boltzmann equation, called the linear Rayleigh–Boltzmann equation. This is the work of Henk van Beijeren, Lanford, Joel Louis Lebowitz and Herbert Spohn in 1980 [29], later completed with applications to color-changing boundary conditions [22].

A few decades later, in 2013, Isabelle Gallagher, Laure Saint-Raymond and Benjamin Texier reopened the work of Lanford and his former student, Francis Gordon King [19], generalizing their proof to the case of compactly supported potentials, and providing in parallel precise estimates on the convergence rate in Lanford's theorem. Eventually, these estimates led in 2016 to an article by Thierry Bodineau, Gallagher and Saint-Raymond on the convergence rate in the linear case, and its dependence on the long time scaling, so as to infer Brownian hydrodynamic limits [5]. These estimates on the convergence rates have been improved in 2024 in our previous paper [12], using the adaptive time cutting that is also used in the present work.

In the present work, we introduce a generalized model for the nonideal Rayleigh gas, where the number of tagged particles is not fixed anymore and goes to infinity with the total number of particles, yet in a small fraction so as to provide the linear Rayleigh–Boltzmann equation at the limit. This model relies on a grand-canonical mixture model, and is necessary to define the *empirical measure of the tagged particle*, whose first order convergence is given in Corollary 1.1. We study the first properties of our new Rayleigh gas mixture model on large time scales, providing quantitative convergence results for its correlation functions in Theorem 1. We extend the existing results to all the correlation functions, and not only the first one, exhibiting the big combinatorial factors that appear for high correlation functions, stemming from the time cutting method that we use. To perform this study, we harness the adaptive time cutting introduced in [12] to improve the quantitative convergence rates, even improving it to gain on the power of the time of validity. Performing an analysis on large time scales imposes a condition on the tagged particles' number, which is merely a technical threshold and is not required on small times.

Further statistical refinements of convergence are studied in the companion paper [13] thanks to the study of its *cumulants*, exploring in particular the phase transition around the low-density scaling.

This first introducing Section 1 defines our Rayleigh gas mixture model, with the grand canonical ensemble for tagged particles, and the main results are stated in Section 1.5. The proof of the convergence in large times of the correlation functions is given in Section 2, using the adaptive time cutting method of [5, 12].

Two appendixes follow to expose technical results, about the cumulants of the exclusion (Appendix A), and the partition function combinatorics (Appendix B).

1.2 Microscopic hard sphere model

Microscopically, we consider exactly the hard sphere model, which resembles a perfectly elastic d -dimensional billiards. The state of a gas of N particles is completely determined by the positions (in the d -dimensional torus \mathbb{T}^d) and the velocities of every particle, represented by the vector

$$\underline{z}_N = (z_1, \dots, z_N) \doteq (\underline{x}_N, \underline{v}_N) \in \mathcal{D}^N \doteq (\mathbb{T}^d \times \mathbb{R}^d)^N.$$

The *hard sphere* model consists in an exclusion condition, which states that two particles cannot get closer than a certain diameter $\varepsilon > 0$: the positions have to belong to the *hard sphere exclusion set*

$$\mathcal{X}_N^\varepsilon \doteq \{\underline{x}_N \in \mathbb{T}^{dN}; \forall i \neq j, d(x_i, x_j) > \varepsilon\}, \quad (1.1)$$

where $d(\cdot, \cdot)$ denotes the distance on the torus; and hence the state of the gas \underline{z}_N must belong to the open domain $\mathcal{D}_N^\varepsilon \doteq \mathcal{X}_N^\varepsilon \times \mathbb{R}^{dN}$.

Within this set, the particles' dynamics is given by the Newton equations for uniform line movement, while on the boundary of $\mathcal{D}_N^\varepsilon$, for at least two particles (let us say i and j) we have that $|x_i - x_j| = \varepsilon$: they collide. Then, if the scalar product $(x_i - x_j) \cdot (v_i - v_j)$ is positive, it means that the particles are exiting the collision in uniform line movement, but otherwise they are entering the collision and must scatter according to the following system giving the post-collisional velocities (v_i', v_j') :

$$\begin{cases} v_i' = v_i - \left\langle v_i - v_j, \frac{x_i - x_j}{\varepsilon} \right\rangle \frac{x_i - x_j}{\varepsilon} \\ v_j' = v_j + \left\langle v_i - v_j, \frac{x_i - x_j}{\varepsilon} \right\rangle \frac{x_i - x_j}{\varepsilon}. \end{cases} \quad (1.2)$$

In the hard sphere case, the interaction is taken instantaneous and elastic: the system (1.2) stems from the preservation of momentum and kinetic energy. We choose to parametrize this scattering system with the following angle

$$\omega = \frac{x_i - x_j}{\varepsilon} \in \mathbb{S}^{d-1}. \quad (1.3)$$

This dynamics is well-defined up to a zero measure set of initial configurations, in which infinite amounts of collisions might happen in finite times, along with collisions between more than two particles at a time. This result was proved by Roger Keith Alexander [1], and might also be found in [15]. Some other models use non-instantaneous scattering governed by potentials of interaction that can be short-range [15] or long-range [10, 3]. The review by Cédric Villani [30] gives a global overview of collisional kinetic theory.

1.3 Statistical description and linear Rayleigh–Boltzmann equation

As the number of particles N gets large, this hard sphere dynamics becomes very difficult to compute, especially because it is very chaotic, so that we choose to describe the gas statistically. This paper

is dedicated to the study of a gas of identical particles, yet divided in two distinguishable parts, represented by tags $\underline{\ell}_N = (\ell_1, \dots, \ell_N) \in \{0, 1\}^N$: the tag $\ell = 0$ will be attributed to particles initially distributed at thermodynamic equilibrium, and the tag $\ell = 1$ to ‘tagged’ particles initially perturbed from that equilibrium.

At fixed $N \in \mathbb{N}$ and $\varepsilon > 0$, we consider $W_N^\varepsilon(t, \underline{z}_N, \underline{\ell}_N)$ the *canonical* probability density of presence of particles with tags $\underline{\ell}_N$ on the phase space $\mathcal{D}_N^\varepsilon$ at time $t \geq 0$: by exchangeability, it is invariant by permutation among particles with identical tags. The microscopic dynamics provides the Liouville transport equation for W_N^ε within $\mathcal{D}_N^\varepsilon$

$$\partial_t W_N^\varepsilon + \underline{v}_N \cdot \nabla_{\underline{x}_N} W_N^\varepsilon = 0. \quad (1.4)$$

The solutions to this equation are provided by the method of characteristics and expressed in terms of the initial distribution $W_N^\varepsilon(0)$ and of the free transport going *back in time*, in such a way that we need the following boundary condition to pursue transporting back the density when two particles emerge from a collision:

$$d(x_i, x_j) = \varepsilon \text{ and } \langle x_i - x_j, v_i - v_j \rangle > 0 \quad \Rightarrow \quad W_N^\varepsilon(\underline{z}_N) \doteq W_N^\varepsilon(\underline{z}_N^*), \quad (1.5)$$

where $\underline{z}_N^* = (z_1, \dots, x_i, v_i^*, \dots, x_j, v_j^*, \dots, z_N)$ denotes the *pre-collisional* state associated to \underline{z}_N . Note that by reversibility, the formulas for (v_i^*, v_j^*) are the same as for (v_i', v_j') .

In the next section, we will present the grand canonical model, which randomizes the number N of particles to a random variable \mathcal{N} , the expectancy of which is tuned by a parameter μ called the *chemical potential*. The kinetic limit that we consider is called the *low density limit*, or Boltzmann–Grad limit, and consists in letting this chemical potential μ go to infinity while keeping a constant mean free path $\mu^{-1}\varepsilon^{1-d} = 1$, so that the particles’ diameter ε goes to 0. In this limit, assuming initial chaos, the first marginal of the density usually converges to the solution to the Boltzmann equation [20, 15, 9].

This solution to the Boltzmann equation, studied by Boltzmann and Maxwell, when well defined, relaxes in large times to an equilibrium called the *Maxwell state* and defined [7] as

$$M_\beta(x, v) \doteq \left(\frac{\beta}{2\pi}\right)^{d/2} \exp\left(-\frac{\beta}{2}|v|^2\right). \quad (1.6)$$

The parameter β stands for the inverse temperature of the system, tuning its intensive (kinetic) energy. It appears that the density $M_\beta^{\otimes N} \mathbb{1}_{\mathcal{D}_N^\varepsilon}$ is an equilibrium of the microscopic hard sphere dynamics, and in this paper we consider specific initial conditions that are close to this thermodynamic equilibrium, to retrieve a *linear version* of the Boltzmann equation, whose theory is much simpler and hence might be derived for long time scales. More precisely, we consider the *nonideal Rayleigh gas* model [27]: a subset of particles are tagged, breaking the particles’ exchangeability. In the case of a single tagged particle, the derivation of this linear equation has been shown [5] when choosing the following perturbation of equilibrium as initial state

$$W_N^\varepsilon(0, \underline{z}_N) = \frac{\mathbb{1}_{\mathcal{X}_N^\varepsilon}(\underline{x}_N)}{\mathcal{Z}_N^{\varepsilon, c}} \rho(x_1) M_\beta^{\otimes N}(\underline{z}_N),$$

for $\rho \in \mathcal{C}(\mathbb{T}^d)$ a continuous space perturbation on the torus, and $\mathcal{Z}_N^{\varepsilon, c}$ a normalization constant. Indeed, in the Boltzmann–Grad limit, the first marginal of W_N^ε behaves like the solution $g \doteq M_\beta \varphi$ to the linear *Rayleigh–Boltzmann equation* [5] with initial condition ρ :

$$\begin{cases} \partial_t \varphi + v \cdot \nabla_x \varphi &= \int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}^d} [\varphi(v^*) - \varphi(v)] M_\beta(v_c) \langle \omega, v_c - v \rangle_+ dv_c d\omega, \\ \varphi(0, x, v) &= \rho(x). \end{cases} \quad (1.7)$$

This linear equation (1.7) is globally well-posed in the velocity-weighted space $\mathbb{L}_x^\infty \mathbb{L}_v^\infty(M_{\beta/2})$ (see the companion paper [13]), and allows to derive the linear heat equation in the hydrodynamic limit [5].

Some partial results exist for this same model with long-range interactions instead of hard sphere collisions [10, 3], yet the complete derivation of the Rayleigh–Boltzmann equation for general potentials is still an open problem. Other ways to derive the linear Rayleigh–Boltzmann equation for long time scales are the *ideal* Rayleigh gas model, in which the particles at equilibrium do not interact among themselves [11, 23, 24], and the Lorentz gas model, which consists in letting a tagged particle evolve in a frozen random background [28, 8, 17].

1.4 Grand canonical framework for the mixture

As introduced in the previous section, we want to study the behaviour of an arbitrarily large subset of tagged particles perturbed away from equilibrium, whose density will follow a linear version of the Boltzmann equation. This derivation has been studied in the case of a finite set [5, 12], but to provide extended statistical results on this gas we hereafter take its size to infinity, yet remaining a tiny fraction of the gas. To *preserve a symmetric structure* on the objects that we consider, we work in the grand canonical ensemble and introduce an additional tagging variable indicating to which set each particle belongs. This approach to describe a gas mixture is different from the canonical one used by Ioakeim Ampatzoglou, Joseph K. Miller and Nataša Pavlović in their article deriving a mixed Boltzmann equation [2]; indeed their description is made in the canonical ensemble, at fixed numbers of particles of each kind.

The particles at equilibrium are taken in the usual low density (Boltzmann–Grad) limit, whereas the tagged perturbed particles only occupy a tiny fraction of the gas, smaller than the Boltzmann–Grad density: otherwise indeed they would behave like a classical Boltzmann dilute gas, satisfying the non-linear Boltzmann equation. This all boils down to the following scaling,

$$\mu \varepsilon^{d-1} = 1 \quad \text{and} \quad 1 \ll \lambda \ll \mu, \quad (S_{\varepsilon, \mu, \lambda})$$

where $\mu > 0$ corresponds to the chemical potential of the particles at equilibrium, and $\lambda > 0$ to that of tagged particles. The notation $\lambda \ll \mu$ simply means that $\lambda \mu^{-1}$ goes to 0 as λ and μ go to infinity. Formally, **the particles at initial equilibrium will be tagged with a 0, and the initially perturbed ‘tagged’ particles will be tagged with a 1**. The tags of all particles hence form a vector $\underline{\ell}_n \in \Lambda_n \doteq \{0, 1\}^n$, identified to the corresponding subset $\underline{\ell}_n \subset \llbracket 1, n \rrbracket$, with the following notation

$$|\underline{\ell}_n| = \|\underline{\ell}_n\|_1 = |\{i \leq n, \ell_i = 1\}| \quad \text{and} \quad \varphi_0^{\otimes \underline{\ell}_n}(\underline{z}_{\underline{\ell}_n}) = \prod_{\substack{i \leq n \\ \ell_i = 1}} \varphi_0(z_i).$$

Moreover, instead of ρ a perturbation happening in space only, as in [5], we will hereafter consider an initial perturbation $\varphi_0 \in \mathbb{L}_x^\infty \mathbb{L}_v^\infty(M_{\beta/2})$ also happening according to the velocities. Now, the mixed grand canonical ensemble consists in relaxing the number of particles, weighting it with a mixed Poisson law depending on the number of tagged particles. More precisely, at fixed λ , and $\mu = \varepsilon^{1-d}$, we take the following weighted canonical initial densities

$$W_n^\varepsilon(0, \underline{z}_n, \underline{\ell}_n) \doteq \frac{\lambda^{|\underline{\ell}_n|} \mu^{n-|\underline{\ell}_n|}}{\mathcal{Z}_\mu} M_\beta^{\otimes n}(\underline{v}_n) \varphi_0^{\otimes \underline{\ell}_n}(\underline{z}_{\underline{\ell}_n}) \mathbb{1}_{\mathcal{X}_n^\varepsilon}(\underline{x}_n) \quad (1.8)$$

driven by the Liouville equation (1.4), for a normalizing constant \mathcal{Z}_μ that we will adjust soon (1.9). We introduce the following *correlation functions*, based on the marginals of the canonical densities. Note that we also take the marginal according to the tags, which corresponds to a sum over $\underline{\ell}_p^* \in \Lambda_p$,

and we denote $\tilde{\ell}_{n+p} \doteq (\ell_n, \ell_p^*)$, $\tilde{z}_{n+p} \doteq (z_n, z_p^*)$:

$$\begin{aligned} F_n^\varepsilon(t, \tilde{z}_n, \ell_n) &= \frac{1}{\mu^{n-|\ell_n|} \lambda^{|\ell_n|}} \sum_{p \geq 0} \sum_{\ell_p^* \in \Lambda_p} \frac{1}{p!} \int W_{n+p}^\varepsilon(t, \tilde{z}_{n+p}, \tilde{\ell}_{n+p}) d\tilde{z}_p^* \\ &\doteq \frac{1}{\mu^{n-|\ell_n|} \lambda^{|\ell_n|}} \sum_{p \geq 0} \frac{1}{p!} W_{n+p}^{\varepsilon, (n)}(t, \tilde{z}_n, \ell_n), \end{aligned}$$

where the normalizing (grand canonical) partition function is hence defined as follows

$$\mathcal{Z}_\mu \doteq \sum_{p \geq 0} \sum_{\ell_p \in \Lambda_p} \frac{\lambda^{|\ell_p|} \mu^{p-|\ell_p|}}{p!} \int M_\beta^{\otimes p}(v_p) \varphi_0^{\otimes \ell_p}(\tilde{z}_{\ell_p}) \mathbf{1}_{\mathcal{X}_p^\varepsilon}(x_p) d\tilde{z}_p. \quad (1.9)$$

In the previous paper [5] presenting the situation where only one particle was perturbed, thanks to the invariance by translation of the system, the partition function was not depending on the perturbation φ_0 . Here, the perturbed particles are correlated one with another, preventing us from using the same argument, which is why the initial perturbation appears in the formula above. Note that the velocities of the non-perturbed particles (in the complementary of ℓ_p) may be integrated using the fact that the equilibrium M_β is of integral 1. We still keep this formulation for symmetry reasons; a more precise study of this object is made in Appendix B.

Our probabilistic study is based on the random variables $(Z_{\varepsilon, i}^{[t]}, L_i)_{1 \leq i \leq \mathcal{N}}$, giving the states at time t , and the tags of the particles. The probability density of the initial state $(Z_{\varepsilon, i}^{[0]}, L_i)$ is given above, and the evolution at time t is a deterministic piecewise affine function of the initial state. The correlation functions are in fact defined such that for any observable $H_n \in \mathcal{C}_c^\infty(\mathcal{D}^n \times \Lambda_n)$, we have

$$\begin{aligned} \mathbb{E} \left[\sum_{1 \leq i_k \neq i_j \leq \mathcal{N}} H_n(Z_{\varepsilon, i_1}^{[t]}, L_{i_1}, \dots, Z_{\varepsilon, i_n}^{[t]}, L_{i_n}) \right] &= \mathbb{E} \left[\delta_{\mathcal{N} \geq n} \frac{\mathcal{N}!}{(\mathcal{N} - n)!} H_n(Z_{\varepsilon, 1}^{[t]}, L_1, \dots, Z_{\varepsilon, n}^{[t]}, L_n) \right] \\ &= \sum_{p=n}^{\infty} \frac{1}{p!} \frac{p!}{(p-n)!} \sum_{\ell_p \in \Lambda_p} \int_{\mathcal{D}_p^\varepsilon} W_p^\varepsilon(t, \tilde{z}_p, \ell_p) H_n(\tilde{z}_n, \ell_n) d\tilde{z}_p \\ &= \sum_{\ell_n \in \Lambda_n} \mu^{n-|\ell_n|} \lambda^{|\ell_n|} \int_{\mathcal{D}_n^\varepsilon} F_n^\varepsilon(t, \tilde{z}_n, \ell_n) H_n(\tilde{z}_n, \ell_n) d\tilde{z}_n. \end{aligned} \quad (1.10)$$

By the usual BBGKY arguments [14, 15] the canonical densities (1.8) satisfy the following hierarchy

$$\begin{aligned} \partial_t W_N^{\varepsilon, (n)} + \underline{v}_n \cdot \nabla_{\underline{x}_n} W_N^{\varepsilon, (n)} &= (N-n) \varepsilon^{d-1} \sum_{\ell=0}^1 \sum_{i=1}^n \int \langle \omega, v_c - v_i \rangle W_N^{\varepsilon, (n+1)}(\tilde{z}_n, \ell_n, x_i + \varepsilon \omega, v_c, \ell) d\omega dv_c \\ &\doteq (N-n) \varepsilon^{d-1} \left(\mathcal{C}_n^{(0)} W_N^{\varepsilon, (n+1)} + \mathcal{C}_n^{(1)} W_N^{\varepsilon, (n+1)} \right). \end{aligned} \quad (1.11)$$

The two operators we hence define are called *collision operators* and represent the action of a $(n+1)$ -th particle (tagged 0 or tagged 1) on the average distribution of n particles. To better understand these operators, we rewrite them using the boundary condition (1.5) when $\langle \omega, v_{n+1} - v_i \rangle > 0$, and the change of variable $\omega \mapsto -\omega$ otherwise:

$$\begin{aligned} \mathcal{C}_n^\ell W_N^{\varepsilon, (n+1)} &= \sum_{i=1}^n \int d\omega dv_{n+1} \langle \omega, v_{n+1} - v_i \rangle \times \\ &\quad \left[W_N^{\varepsilon, (n+1)}(\tilde{z}_n^*, \ell_n, x_i + \varepsilon \omega, v_{n+1}^*, \ell) - W_N^{\varepsilon, (n+1)}(\tilde{z}_n, \ell_n, x_i - \varepsilon \omega, v_{n+1}, \ell) \right]. \end{aligned} \quad (1.12)$$

Morally, we look at the influence of a $(n + 1)$ -th particle—with tag ℓ —on the dynamics, colliding with one of the n existing ones with angle ω and velocity v_{n+1} , whence the name collision operators. The *cross section* $\langle \omega, v_{n+1} - v_i \rangle_+$ weights the likelihood of such a collision. As a consequence, the correlation functions satisfy

$$\begin{aligned} \partial_t F_n^\varepsilon + \underline{v}_n \cdot \nabla_{\underline{x}_n} F_n^\varepsilon &= \frac{1}{\mu^{n-|\ell_n|} \lambda^{|\ell_n|}} \sum_{k \geq 0} \frac{k \varepsilon^{d-1}}{k!} \left(\mathcal{C}_n^{(0)} W_{n+k}^{\varepsilon, (n+1)} + \mathcal{C}_n^{(1)} W_{n+k}^{\varepsilon, (n+1)} \right) \\ &= \frac{\mu \varepsilon^{d-1}}{\mu^{n+1-|\ell_n|} \lambda^{|\ell_n|}} \sum_{k \geq 0} \frac{1}{k!} \mathcal{C}_n^{(0)} W_{n+k+1}^{\varepsilon, (n+1)} + \frac{\lambda \varepsilon^{d-1}}{\mu^{n-|\ell_n|} \lambda^{|\ell_n|+1}} \sum_{k \geq 0} \frac{\varepsilon^{d-1}}{k!} \mathcal{C}_n^{(1)} W_{n+k+1}^{\varepsilon, (n+1)}, \end{aligned}$$

which eventually leads to the following hierarchy in our mixed Boltzmann–Grad scaling $(\mathcal{S}_{\varepsilon, \mu, \lambda})$

$$\partial_t F_n^\varepsilon + \underline{v}_n \cdot \nabla_{\underline{x}_n} F_n^\varepsilon = \mathcal{C}_n^{(0)} F_{n+1}^\varepsilon + \frac{\lambda}{\mu} \mathcal{C}_n^{(1)} F_{n+1}^\varepsilon. \quad (1.13)$$

The formal limit of the hierarchy above (1.13) with initial conditions (1.8), presented in Section 2.1, is satisfied by the family $(M_\beta^{\otimes n} \varphi^{\otimes \ell_n})_{n \geq 1}$, where φ is the solution of the Rayleigh–Boltzmann equation (1.7) with initial data φ_0 . This result is formalized in the following section, with a quantitative convergence rate.

1.5 Main results: long-time convergence of the correlation functions

The following theorem provides a convergence rate of the mixed correlation functions to the solutions of the Rayleigh–Boltzmann equation, which is a generalization of [12, 5] with a time scale of validity improved by a power 1/4 thanks to a more precise computation (see the proof of Proposition 2.5). We extend the result to all the correlation functions (not only the first one), yet at the cost of a bad constant n^{cn} stemming from the time cutting method we use. This constant, which did not appear in the convergence of the first marginal, is due to an accumulation of errors at each time step of our cutting. Note finally that we use here the adaptive time cutting introduced in [12], improving greatly the convergence rate compared with [5].

Theorem 1 (Convergence of the correlation functions). *For some sets $\Delta_n^\varepsilon \subset \mathcal{D}^n$ whose measure goes to 0 with ε , there exists a constant c_β depending only on the temperature and the dimension such that, for any $\alpha \in (0, 3/4)$, as long as*

$$t \lesssim (\log |c_\beta \log \varepsilon|)^{\frac{3}{4}-\alpha} \quad \text{and} \quad \lambda \lesssim |\log \varepsilon|^{1-\alpha}, \quad (1.14)$$

and for ε small enough, one has the following convergence rate of the correlation functions to the linear Rayleigh–Boltzmann solutions in our mixed low density scaling $(\mathcal{S}_{\varepsilon, \mu, \lambda})$, for a constant $c > 0$;

$$\|F_n^\varepsilon - M_\beta^{\otimes n} \varphi^{\otimes \ell_n}\|_{\mathbb{L}^\infty([0, t] \times \mathcal{D}^n \setminus \Delta_n^\varepsilon)} \leq n^{cn} \exp\left(-c_\beta |\log \varepsilon|^{1-\alpha}\right).$$

The notation $t \lesssim (\log |c_\beta \log \varepsilon|)^{\frac{3}{4}-\alpha}$ means that, for a good constant $c > 0$ depending only on the dimension d and the inverse temperature β , one has

$$t \leq c (\log |c_\beta \log \varepsilon|)^{\frac{3}{4}-\alpha}.$$

The proof of this theorem is the subject of Section 2. It is close to the proof presented in [12], but in the mixed grand canonical framework and for all the correlation functions instead of the first marginal only.

This theorem provides a first corollary on the statistical behaviour of the gas. So as to state this result, for any observable $H \in \mathcal{C}_c^\infty(\mathcal{D} \times \Lambda_1)$ we define the following random variables: the *empirical measure of all particles*

$$\pi_t^\varepsilon[H] \doteq \frac{1}{\mu} \sum_{i=1}^{\mathcal{N}} H(Z_{\varepsilon,i}^{[t]}, L_i), \quad (1.15)$$

and the *empirical measure of tagged particles*

$$\tilde{\pi}_t^\varepsilon[H] \doteq \frac{1}{\lambda} \sum_{i=1}^{\mathcal{N}} H(Z_{\varepsilon,i}^{[t]}, L_i) \mathbb{1}_{L_i=1}. \quad (1.16)$$

Thanks to Theorem 1, one can deduce the convergence of these empirical measures, hence providing a law of large numbers for the hard sphere dynamics. This result is given in the space \mathbf{L}^2 of *square-integrable random variables*, endowed with the norm $\mathbb{E} [|\cdot|^2]$.

Corollary 1.1 (Law of large numbers for the dynamics). *The empirical measures converge as random variables in \mathbf{L}^2 , the non-tagged particles towards an equilibrium state, and the tagged ones to a state described by the linear Rayleigh–Boltzmann equation (1.7), in the following way*

$$\pi_t^\varepsilon[H] \xrightarrow[\varepsilon \rightarrow 0]{\mathbf{L}^2} \int M_\beta(v) H(z, 0) dz, \quad (1.17)$$

and

$$\tilde{\pi}_t^\varepsilon[H] \xrightarrow[\varepsilon \rightarrow 0]{\mathbf{L}^2} \int M_\beta(v) \varphi(t, z) H(z, 1) dz. \quad (1.18)$$

Proof. To show that the random variable $\pi_t^\varepsilon[H]$ converges in \mathbf{L}^2 to a deterministic limit $a \in \mathbb{R}$, writing

$$\pi_t^\varepsilon[H] - a = \pi_t^\varepsilon[H] - \mathbb{E}[\pi_t^\varepsilon[H]] + \mathbb{E}[\pi_t^\varepsilon[H]] - a,$$

it is enough to show that $\mathbb{E}[\pi_t^\varepsilon[H]] \xrightarrow[\varepsilon \rightarrow 0]{} a$ and $\mathbb{E} [|\pi_t^\varepsilon[H] - \mathbb{E}[\pi_t^\varepsilon[H]]|^2] \xrightarrow[\varepsilon \rightarrow 0]{} 0$. Using formula (1.10), one can write

$$\mathbb{E} \left[\frac{1}{\mu} \sum_{i=1}^{\mathcal{N}} H(Z_{\varepsilon,i}^{[t]}, L_i) \right] = \int F_1^\varepsilon(t, z_1, 0) H(z_1, 0) dz_1 + \frac{\lambda}{\mu} \int F_1^\varepsilon(t, z_1, 1) H(z_1, 1) dz_1,$$

so that the expectancies converge by Theorem 1 above. For concision, we show the fact that the variance vanishes in the tagged case, denoting $h \doteq H(\cdot, 1)$. The equilibrium case is treated in a similar though simpler way. Let us compute, once again by formula (1.10),

$$\begin{aligned} \mathbb{E} [|\tilde{\pi}_t^\varepsilon - \mathbb{E}[\tilde{\pi}_t^\varepsilon[H]]|^2] &= \frac{1}{\lambda^2} \mathbb{E} \left[\sum_{i,j=1}^{\mathcal{N}} h(Z_{\varepsilon,i}^{[t]}) h(Z_{\varepsilon,j}^{[t]}) \right] - \frac{2}{\lambda} \mathbb{E} \left[\sum_{i=1}^{\mathcal{N}} h(Z_{\varepsilon,i}^{[t]}) \int F_1^\varepsilon(1) h \right] + \left(\int F_1^\varepsilon(1) h \right)^2 \\ &= \int F_2^\varepsilon(1, 1) h^{\otimes 2} + \frac{1}{\lambda} \int F_1^\varepsilon(1) h^2 - 2 \left(\int F_1^\varepsilon(1) h \right)^2 + \left(\int F_1^\varepsilon(1) h \right)^2 \\ &\xrightarrow[\varepsilon \rightarrow 0]{} \int M_\beta^{\otimes 2} \varphi^{\otimes 2} + 0 - \left(\int M_\beta \varphi \right)^2 = 0, \end{aligned}$$

thanks to the convergence of the correlation functions (Theorem 1). The other convergence (1.17) follows similarly in our scaling, using the fact that $\frac{\lambda}{\mu}$ goes to 0. \square

2 Proof of the convergence of the correlation functions

We give in this section the proof of Theorem 1.

2.1 Pseudo-trajectories and strategy of proof

Let us choose to denote $p_\mu \doteq \frac{\lambda}{\mu}$ the fraction of initially perturbed particles, so that iterating Duhamel formula as in [15] or [5], we can write the Dyson expansion

$$F_n^\varepsilon(t) = \sum_{k \geq 0} \sum_{\underline{\ell}_k^* \in \Lambda_k} p_\mu^{|\underline{\ell}_k^*|} Q_{n, \underline{\ell}_k^*}(t) F_{n+k}^\varepsilon(0), \quad (2.1)$$

developing the choice of the encountered tags $\underline{\ell}_k^* \doteq (\tilde{\ell}_{n+1}, \dots, \tilde{\ell}_{n+k})$, with the successive-collision operators defined as

$$Q_{n, \underline{\ell}_k^*}(t) \doteq \int_{T_k(t)} \Theta_n(t - t_1) \mathcal{C}_n^{\tilde{\ell}_{n+1}} \Theta_{n+1}(t_1 - t_2) \dots \mathcal{C}_{n+k-1}^{\tilde{\ell}_{n+k}} \Theta_{n+k}(t_k) dt_k, \quad (2.2)$$

where $\Theta_n(\tau)$ denotes the transport semi-group operator in $\mathcal{D}_n^\varepsilon$ with specular reflections, for a time τ . The collision times are integrated over

$$T_k(t) \doteq \left\{ \underline{t}_k \mid 0 \doteq t_{k+1} \leq t_k \leq \dots \leq t_1 \leq t_0 \doteq t \right\}. \quad (2.3)$$

The main idea of the proof, coming from Lanford's original paper [20], is to use a coupling between this expansion and its limit version, implying imaginary histories of the particles, among those that eventually lead to the state z_n at time t . These histories, called *pseudo-trajectories*, are non-physical trajectories that—in a way—allow to extend the method of characteristics for the successive-collision operators.

Indeed, the transport operators appearing in (2.2) correspond to following the characteristics of free transport, with specular reflections: taking the first operator $\Theta_n(t - t_1)$ of a functional amounts to considering this functional at time t_1 , in a state $z_n^{[t_1]}$ given by the backwards hard sphere dynamics.

Then, the first collision operator (1.12) writes

$$\mathcal{C}_n^\ell F_{n+1}^\varepsilon = \sum_{i=1}^n \sum_{s_1=\pm 1} s_1 \int d\omega_1 dv_{n+1} \langle \omega_1, v_{n+1} - v_i \rangle_+ F_{n+1}^\varepsilon(z_n^{(s_1)}, \underline{\ell}_n, x_i + s_1 \varepsilon \omega_1, v_{n+1}^{(s_1)}, \ell),$$

where $z_n^{(+1)} = z_n^*$ and $z_n^{(-1)} = z_n$, scattered for the gain term, and let unchanged for the loss term, so that the collision is always incoming, allowing to pursue the backwards method of characteristics with the next transport operator. Hence, for given collision parameters $(i, s_1, \omega_1, v_{n+1})$, this operator can be seen as a weighted adjunction of a particle to the characteristics—or pseudo-trajectory—which scatters (or not, according to s_1) with particle i , creating a new state $z_{n+1}^{[t_1]} \doteq (z_n^{[t_1]}, x_i + s_1 \varepsilon \omega_1, v_{n+1}^{(s_1)})$. The integration and sum over these collision parameters will yield an integral over pseudo-trajectories. Iterating this extended method of characteristics and tracking the pseudo-trajectories $(z_{n+j}^{[t_j]})$ thus constructed, we bring the analysis back to the value of the functional at time $\tau = 0$, in the state $z_{n+k}^{[0]}$.

We will have to record the numbering labels of the existing particles meeting the new one, the velocities of the particles that spring up, the angles at which the encounters happen, and whether they scatter or not. The pseudo-trajectories will also keep track of the tags of the encountered particles.

Here is precisely how we construct the pseudo-trajectories. The choice of the successive encountered tags is registered in $\underline{\ell}_k^* = (\tilde{\ell}_{n+1}, \dots, \tilde{\ell}_{n+k})$, and expanding all the sums in all the collision operators (2.2), we can sum them up to the history (m_1, \dots, m_k) of which particle encountered the $(n+i)$ -th new one. These particles naturally belong to the following set

$$\mathcal{M}_{n,k} \doteq \left\{ (m_1, \dots, m_k) \mid \forall i \leq k, m_i \leq n + i - 1 \right\}.$$

We consider the scattering labels $(s_1, \dots, s_k) \in \{\pm 1\}^k$. The fact that some encounters do not scatter, along with the fact that some particles are artificially added, is why the pseudo-trajectories are not physical trajectories. Once the total history

$$\chi_k \doteq (\underline{m}_k, \underline{\ell}_k^*, \underline{s}_k) \in \mathcal{H}_{n,k} \doteq \mathcal{M}_{n,k} \times \Lambda_k \times \{\pm 1\}^k \quad (2.4)$$

is fixed, for given collision parameters $(\omega_k, v_{n+1}, \dots, v_{n+k})$, we can construct the pseudo-trajectories for every endstate $\underline{z}_n = \underline{z}_n^{[t]}$, backwards in time to an initial configuration $\underline{z}_{n+m}^{[0]}$, following the inductive procedure below:

$$\begin{cases} \underline{z}_n^{[t]} \doteq \underline{z}_n \\ \forall i \in \llbracket 0, k \rrbracket, \forall \tau \in (t_{i+1}, t_i), \underline{z}_{n+i}^{[\tau]} \text{ follows (backwards) the physical hard sphere dynamics} \\ \forall i \in \llbracket 1, k \rrbracket, \underline{z}_{n+i}^{[t_i]} = \left(\underline{z}_{n+i-1}^{[t_i^+, \langle s_i \rangle]}, x_{m_i} + s_i \varepsilon \omega_i, v_{n+1}^{\langle s_i \rangle} \right). \end{cases} \quad (2.5)$$

One may observe that the change of velocities in the last step is automatic by the hard sphere dynamics' boundary condition, but it will not be for the limit version of pseudo-trajectories, since the limit particles are formally pointwise. In the end, one can write the *pseudo-trajectory formulation* of the Dyson expansion

$$F_n^\varepsilon(t) = \sum_{k \geq 0} \sum_{\chi_k} p_\mu^{|\underline{\ell}_k^*|} \int_{T_k(t)} dt_k \int d\underline{\omega}_k dv_{n+1} \dots dv_{n+k} \prod_{i=1}^k s_i \langle \omega_i, v_{n+i} - v_{m_i}^{[t_i^+]} \rangle_+ F_{n+k}^\varepsilon(0, \underline{z}_{n+k}^{[0]}, \tilde{\underline{\ell}}_{n+k}). \quad (2.6)$$

A small technical detail lies in the fact that the added particles must satisfy the exclusion condition. A way to deal with it may be to impose a condition on the domain of integration of the collision angles [6], yet here to simplify we merely change the definition of the pseudo-trajectories: if at any moment the exclusion condition is violated by the adjunction of a particle, then the trajectories are frozen in this state until time $\tau = 0$, so that the integral formally vanishes thanks to the initial distribution $F_{n+k}^\varepsilon(0)$ being 0 outside of $\mathcal{D}_{n+k}^\varepsilon$.

Indeed, our goal is now to prove the convergence of the correlation functions to the **limit densities**

$$G_n(t, \underline{z}_n, \underline{\ell}_n) \doteq M_\beta^{\otimes n}(\underline{v}_n) \varphi^{\otimes \underline{\ell}_n}(t, \underline{z}, \underline{\ell}_n),$$

where φ is the solution to the linear Rayleigh–Boltzmann equation (1.7). Because of the structure of this equation, this family satisfies the following hierarchy

$$(\partial_t + \underline{v} \cdot \nabla_{\underline{x}}) G_n = \sum_{i=1}^n \sum_{s_c = \pm 1} s_c \int d\underline{\omega} dv_{n+1} \langle \omega, v_{n+1} - v_i \rangle_+ G_{n+1}(\underline{z}_n^{\langle s_c \rangle}, \underline{\ell}_n, x_i, v_{n+1}^{\langle s_c \rangle}, 0), \quad (2.7)$$

noticing that the terms vanish when the scattering occur between two particles distributed according to the equilibrium $M_\beta^{\otimes n}$. This equation is the formal limit of the BBGKY hierarchy (1.13) in the mixed low density regime $(\mathcal{S}_{\varepsilon, \mu, \lambda})$: it makes only appear the collision operator linked to equilibrium particles, tagged 0, since the other one has a factor λ/μ that vanishes at the limit. It leads to the following limit version of the Dyson expansion (2.1)

$$G_n(t) = \sum_{k \geq 0} Q_{n, \underline{0}_k}^{\text{lim}}(t) G_{n+k}(0), \quad (2.8)$$

where the following successive-collision operators contain only collisions with particles at equilibrium, and limit free transport of pointwise particles, without scattering:

$$Q_{n, \underline{0}_k}^{\text{lim}}(t) \doteq \int_{T_k(t)} \Theta_n^{\text{lim}}(t - t_1) \mathcal{C}_n^{(0), \text{lim}} \Theta_{n+1}^{\text{lim}}(t_1 - t_2) \dots \mathcal{C}_{n+k-1}^{(0), \text{lim}} \Theta_{n+k}^{\text{lim}}(t_k) dt_k. \quad (2.9)$$

The limit collision operators $\mathcal{C}_n^{(0),\text{lim}}$ are the formal limit of the collision operators (1.12), for $\varepsilon = 0$. The same computation as below for this limit hierarchy leads to a similar writing in terms of pseudo-trajectories

$$G_n(t) = \sum_{k \geq 0} \sum_{\underline{X}_k} \mathbf{1}_{\ell_k^* = 0_k} \int_{T_k(t)} d\underline{t}_k \int d\underline{\omega}_k dv_{n+1} \dots dv_{n+k} \prod_{i=1}^k s_i \langle \omega_i, v_{n+i} - v_{m_i}^{[t_i^+]} \rangle_+ G_{n+k}(0, \underline{\zeta}_{n+k}^{[0]}, \tilde{\underline{\ell}}_{n+k}), \quad (2.10)$$

where the limit pseudo-trajectories $(\underline{\zeta}_{n+i}^{[t_i]})_{i \leq k}$ are defined as their hard sphere versions (2.5) for $\varepsilon = 0$, with the noticeable difference that in the dynamics followed on each time interval (t_{i+1}, t_i) , the particles are pointwise and hence follow the free flow without any scattering.

Strategy of proof We will couple both pseudo-trajectory formulations (2.6) and (2.10), bringing down the difference at a certain time $(F_n^\varepsilon - G_n)(t, z_n)$ to the difference at time 0 of higher correlation functions $[F_{n+k}^\varepsilon(0, z_{n+k}^{[0]}) - G_{n+k}(0, \underline{\zeta}_{n+k}^{[0]})]$. For these hierarchies to be well coupled, the two pseudo-trajectories must be close; the classical argument is to show that the transport operators Θ_k , appearing between the collision operators, do not imply additional recollisions, since the trajectories would diverge from the limit transport operator Θ_k^{lim} , defined on the whole space \mathcal{D}^k without recollisions.

Indeed with this method, it will be enough to use **continuity estimates** on the operators to bring the convergence back to time $\tau = 0$ where we can use explicit **initial proximity**. Nevertheless, these continuity estimates demand to work with trajectories that do not contain too many particles, so that we will first compute a **tree pruning**, and control the pruned-out term using some **a priori estimates** on the densities. The last step before proving the **convergence** will be to **discard the pseudo-trajectories** in which some perturbed particles are encountered, as this situation does not happen in the limit pseudo-trajectories. This strategy will be followed in the following sections.

2.2 Initial proximity

The total kinetic energy, preserved at fixed number of particles by the transport and by elastic collisions, will be denoted

$$\|\underline{v}_k\|^2 \doteq \sum_{i=1}^k |v_i|^2,$$

where $|v_i|$ is the Euclidean norm of the velocity $v_i \in \mathbb{R}^d$ of particle i . For an inverse temperature $\beta > 0$ and $k \in \mathbb{N}^*$, we consider the space $\mathcal{F}_{k,\beta}$ of measurable functions defined almost everywhere on the domain \mathcal{D}^k such that

$$\|f_k\|_{k,\beta} \doteq \sup_{\underline{z}_k \in \mathcal{D}^k} |f_k(\underline{z}_k) \exp(\beta \|\underline{v}_k\|^2)| < \infty, \quad (2.11)$$

hence decreasing at least as the Gaussian equilibrium $M_{2\beta}^{\otimes k}$ in velocities. We denote

$$C_0 \doteq \max \left[\|M_\beta\|_{1,\beta/2} ; \|M_\beta \varphi_0\|_{1,\beta/4} ; \left\| M_\beta M_{\beta/2}^{-\frac{1}{2}} \varphi_0 \right\|_{\mathbb{L}^\infty(\mathcal{D})} \right]. \quad (2.12)$$

The initial error between the microscopic densities and the limit ones is mainly due to the exclusion condition, of which we can compute an explicit control.

Proposition 2.1 (Initial proximity). *For all $n \in \mathbb{N}$ and tags $\underline{\ell}_n \in \Lambda_n$, for any $\lambda, \mu > 0$ in the scaling $(\mathcal{S}_{\varepsilon,\mu,\lambda})$, one has*

$$\left\| \mathbf{1}_{\mathcal{X}_n^\varepsilon} M_\beta^{\otimes n} \varphi_0^{\otimes \underline{\ell}_n} - F_n^\varepsilon(0, \underline{\ell}_n) \right\|_{n,\beta/4} \leq C_0^n \varepsilon. \quad (2.13)$$

Proof of the proposition. We denote $\mathbf{1}_{i \not\sim j} \doteq \mathbf{1}_{d(x_i, x_j) > \varepsilon}$ and $\mathbf{1}_{i \sim j} \doteq \mathbf{1}_{d(x_i, x_j) \leq \varepsilon}$ the indicator of exclusion between i and j and its complementary. Recalling that $\mathbf{1}_{\mathcal{X}_N^\varepsilon}$ denotes the exclusion condition (1.1) and by definition (1.9) of the partition function \mathcal{Z}_μ , we can write, once again denoting $\tilde{\ell}_{n+p} \doteq (\ell_n, \ell_p^*)$:

$$\begin{aligned} & \mathbf{1}_{\mathcal{X}_n^\varepsilon} M_\beta^{\otimes n} \varphi_0^{\otimes \ell_n} - F_n^\varepsilon(0) \\ &= \frac{1}{\mathcal{Z}_\mu} \sum_{p \geq 0} \frac{1}{p!} \left[\sum_{\ell_p^* \in \Lambda_p} \lambda^{|\ell_p^*|} \mu^{p-|\ell_p^*|} \mathbf{1}_{\mathcal{X}_n^\varepsilon} M_\beta^{\otimes n} \varphi_0^{\otimes \ell_n} \int \varphi_0^{\otimes \ell_p^*} M_\beta^{\otimes p} \mathbf{1}_{\mathcal{X}_p^\varepsilon} - \left(\varphi_0^{\otimes \tilde{\ell}_{n+p}} M_\beta^{\otimes n+p} \mathbf{1}_{\mathcal{X}_{n+p}^\varepsilon} \right)^{(n)} \right]. \end{aligned}$$

Expanding the marginals' formula, the following term appears, in which we decompose the exclusion indicator $\mathbf{1}_{\mathcal{X}_{n+p}^\varepsilon}$ as

$$\begin{aligned} & \mathbf{1}_{\mathcal{X}_n^\varepsilon} M_\beta^{\otimes n} \varphi_0^{\otimes \ell_n} \int \varphi_0^{\otimes \ell_p^*} M_\beta^{\otimes p} \mathbf{1}_{\mathcal{X}_p^\varepsilon} - \int \varphi_0^{\otimes \tilde{\ell}_{n+p}} M_\beta^{\otimes n+p} \mathbf{1}_{\mathcal{X}_{n+p}^\varepsilon} \mathrm{d}z_p^* \\ &= \mathbf{1}_{\mathcal{X}_n^\varepsilon} M_\beta^{\otimes n} \varphi_0^{\otimes \ell_n} \int (M_\beta \varphi_0)^{\otimes \ell_p^*} (z_p^*) \mathbf{1}_{\mathcal{X}_p^\varepsilon}(x_p^*) \left(1 - \prod_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \mathbf{1}_{x_i \not\sim x_j^*} \right) \mathrm{d}x_p^* \mathrm{d}v_{\ell_p^*}^*, \end{aligned}$$

using that M_β is of integral 1. Then, we will harness the following basic set property,

$$1 - \prod_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \mathbf{1}_{x_i \not\sim x_j^*} \leq \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \mathbf{1}_{x_i \sim x_j^*},$$

yielding

$$\left| \mathbf{1}_{\mathcal{X}_n^\varepsilon} M_\beta^{\otimes n} \varphi_0^{\otimes \ell_n} - F_n^\varepsilon(0) \right| \leq \frac{\mathbf{1}_{\mathcal{X}_n^\varepsilon} M_\beta^{\otimes n} \varphi_0^{\otimes \ell_n}}{\mathcal{Z}_\mu} \sum_{p \geq 0} \sum_{\ell_p^* \in \Lambda_p} \frac{\lambda^{|\ell_p^*|} \mu^{p-|\ell_p^*|}}{p!} \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \int (M_\beta \varphi_0)^{\otimes \ell_p^*} \mathbf{1}_{\mathcal{X}_p^\varepsilon} \mathbf{1}_{x_i \sim x_j^*} \mathrm{d}x_p^* \mathrm{d}v_{\ell_p^*}^*.$$

To be able to integrate over x_j^* , we denote $\check{\ell}_p^{(j)} \doteq (\ell_1^*, \dots, \ell_{j-1}^*, \ell_{j+1}^*, \dots, \ell_p^*)$ the vector of all tags apart from j . Using the definition (2.12) of C_0 and summing over ℓ_j^* , we get

$$\begin{aligned} & \left\| \mathbf{1}_{\mathcal{X}_n^\varepsilon} M_\beta^{\otimes n} \varphi_0^{\otimes \ell_n} - F_n^\varepsilon(0) \right\|_{n, \beta} \\ & \leq \frac{C_0^n}{\mathcal{Z}_\mu} \sum_{p \geq 0} \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} \sum_{\substack{\check{\ell}_p^{(j)} \in \Lambda_{p-1}}} \frac{\lambda^{|\check{\ell}_p^{(j)}|} \mu^{p-1-|\check{\ell}_p^{(j)}|}}{p!} (\lambda C_0 + \mu) \int (M_\beta \varphi_0)^{\otimes \check{\ell}_p^{(j)}} \mathbf{1}_{\mathcal{X}_p^\varepsilon} \mathbf{1}_{x_i \sim x_j^*} \\ & \leq \frac{C_0^n}{\mathcal{Z}_\mu} \sum_{p \geq 0} np \sum_{\ell_{p-1}^* \in \Lambda_{p-1}} \frac{\lambda^{|\ell_{p-1}^*|} \mu^{p-1-|\ell_{p-1}^*|}}{p!} 2\mu |\mathbb{S}^{d-1}| \varepsilon^d \int (M_\beta \varphi_0)^{\otimes \ell_{p-1}^*} \mathbf{1}_{\mathcal{X}_{p-1}^\varepsilon} \end{aligned}$$

using the exchangeability of identical particles, integrating over x_j^* and using the fact that $C_0 \lambda < \mu$. We get in the mixed low density scaling $(S_{\varepsilon, \mu, \lambda})$

$$\left\| \mathbf{1}_{\mathcal{X}_n^\varepsilon} M_\beta^{\otimes n} \varphi_0^{\otimes \ell_n} - F_n^\varepsilon(0) \right\|_{n, \beta} \leq 2 |\mathbb{S}^{d-1}| \varepsilon \frac{C_0^n}{\mathcal{Z}_\mu} n \sum_{p \geq 1} \sum_{\ell_{p-1}^* \in \Lambda_{p-1}} \frac{\lambda^{|\ell_{p-1}^*|} \mu^{p-1-|\ell_{p-1}^*|}}{(p-1)!} \int (M_\beta \varphi_0)^{\otimes \ell_{p-1}^*} \mathbf{1}_{\mathcal{X}_{p-1}^\varepsilon},$$

which concludes the proof recognizing the partition function \mathcal{Z}_μ (1.9) after an index shift. \square

2.3 A priori estimates

As in the previous works [5, 12] about the Rayleigh gas, the long-time derivation is allowed thanks to a priori estimates yielded by the rigid structure of the equilibrium. Here, some additional technical difficulties, dealt with in Appendix B, appear in the proof of these estimates because of the structure of the grand canonical mixture and its partition function. First, let us observe that the initial canonical densities defined in (1.8) satisfy for all $(\underline{z}_n, \underline{\ell}_n) \in \mathcal{D}^n \times \Lambda_n$ that

$$W_n^\varepsilon(0, \underline{z}_n, \underline{\ell}_n) \leq \frac{\lambda^{|\underline{\ell}_n|} \mu^{n-|\underline{\ell}_n|}}{\mathcal{Z}_\mu} \left\| M_\beta M_{\beta/2}^{-\frac{1}{2}} \varphi_0 \right\|_{\mathbb{L}^\infty}^{|\underline{\ell}_n|} M_{\beta/2}^{\otimes n}(\underline{v}_n) \mathbb{1}_{\mathcal{X}_n^\varepsilon}(\underline{x}_n). \quad (2.14)$$

Since their evolution is simply given by the global transport of n particles, by which the equilibrium $M_{\beta/2}^{\otimes n} \mathbb{1}_{\mathcal{X}_n^\varepsilon}$ is invariant, for all times $t \geq 0$ the bound (2.14) is propagated by the transport and remains true. Hence, taking the marginals we get

$$\begin{aligned} W_n^{\varepsilon, (k)}(t, \underline{z}_k, \underline{\ell}_k) &\leq \sum_{\underline{\ell}_{n-k}^* \in \Lambda_{n-k}} \frac{(\lambda C_0)^{|\underline{\ell}_k| + |\underline{\ell}_{n-k}^*|} \mu^{n-|\underline{\ell}_k| - |\underline{\ell}_{n-k}^*|}}{\mathcal{Z}_\mu} M_{\beta/2}^{\otimes k}(\underline{v}_k) (\mathbb{1}_{\mathcal{X}_n^\varepsilon})^{(k)}(\underline{x}_k) \\ &\leq (\lambda C_0 + \mu)^{n-k} \frac{(\lambda C_0)^{|\underline{\ell}_k|} \mu^{k-|\underline{\ell}_k|}}{\mathcal{Z}_\mu} M_{\beta/2}^{\otimes k}(\underline{v}_k) \mathbb{1}_{\mathcal{X}_n^\varepsilon}^{(k)}(\underline{x}_k), \end{aligned}$$

where the factor $(\lambda C_0 + \mu)^{n-k}$ stems from the sum over $\underline{\ell}_{n-k}^*$, using the binomial theorem. The main difference in the linear case, compared to the general non-linear one, is that the bound (2.14) on the canonical densities uses the invariant density $M_{\beta/2}^{\otimes n}$, that passes to the k -th marginals to become $M_{\beta/2}^{\otimes k}$, contrary to the constant C_0^n in the general case.

Eventually, these bounds over the marginals of the canonical densities leaves the following a priori estimate for the correlation functions

$$\begin{aligned} F_n^\varepsilon(t, \underline{z}_n, \underline{\ell}_n) &\leq \frac{M_{\beta/2}^{\otimes n}}{\mu^{n-|\underline{\ell}_n|} \lambda^{|\underline{\ell}_n|}} \sum_{p \geq 0} \frac{1}{p!} \frac{(\lambda C_0 + \mu)^p (\lambda C_0)^{|\underline{\ell}_n|} \mu^{n-|\underline{\ell}_n|}}{\mathcal{Z}_\mu} \mathbb{1}_{\mathcal{X}_{n+p}^\varepsilon}^{(n)} \\ &\leq \frac{C_0^{|\underline{\ell}_n|} M_{\beta/2}^{\otimes n}}{\mathcal{Z}_\mu} \sum_{p \geq 0} \frac{(\lambda C_0 + \mu)^p}{p!} \int \mathbb{1}_{\mathcal{X}_p^\varepsilon} \\ &\leq \frac{C_0^{|\underline{\ell}_n|} M_{\beta/2}^{\otimes n}}{\mathcal{Z}_\mu} \sum_{q, r \geq 0} \frac{(\lambda C_0)^q \mu^r}{q! r!} \int \mathbb{1}_{\mathcal{X}_{q+r}^\varepsilon} \end{aligned} \quad (2.15)$$

where at line (2.15) we computed a direct binomial theorem. The key point to end our a priori estimate is now to control the remaining quotient implying the partition function \mathcal{Z}_μ and the slightly modified version of it, which is the following proposition.

Proposition 2.2. *There exists a constant C_d depending only on the dimension such that for μ large enough, in our mixed Boltzmann-Grad scaling $(S_{\varepsilon, \mu, \lambda})$, we have*

$$\frac{1}{\mathcal{Z}_\mu} \sum_{q, r \geq 0} \frac{(\lambda C_0)^q \mu^r}{q! r!} \int \mathbb{1}_{\mathcal{X}_{q+r}^\varepsilon} \leq C_d^{C_0 \lambda}.$$

The proof of this technical result is given in Section B, using an explicit expansion of the partition function according to the cumulants of the exclusion. Eventually, this leads to the following proposition, which is the main argument of our long time analysis.

Proposition 2.3 (A priori estimates for the correlation functions). *For any $n \in \mathbb{N}$ and $\varepsilon > 0$ in the mixed scaling $(S_{\varepsilon, \mu, \lambda})$, one has*

$$F_n^\varepsilon(t, \underline{z}_n, \underline{\ell}_n) \leq C_0^{|\underline{\ell}_n|} M_{\beta/2}^{\otimes n} \times C^{C_0 \lambda}. \quad (2.16)$$

2.4 Continuity estimates

Thanks to the a priori estimates (2.16) given in the previous Section 2.3, all the correlation functions (F_n^ε) extended by 0 out of $\mathcal{D}_n^\varepsilon$ belong to the space $\mathcal{F}_{n,\beta/4}$ defined in (2.11).

In the following, up to change the initial temperature β to 4β , to simplify the notation in the computation below we will assume that they belong to the space $\mathcal{F}_{n,\beta}$, with

$$\sup_{t \geq 0} \|F_n^\varepsilon(t)\|_{n,\beta} \leq C_0^n C^{C_0 \lambda}, \quad (2.17)$$

thanks to the estimates above, and similarly for the initial proximity, from Proposition 2.1,

$$\left\| \mathbb{1}_{\mathcal{X}_n^\varepsilon} M_\beta^{\otimes n} \varphi_0^{\otimes \ell_n} - F_n^\varepsilon(0, \underline{\ell}_n) \right\|_{n,\beta} \leq C_0^n \varepsilon. \quad (2.18)$$

Our whole long time derivation is allowed thanks to the fact that the a priori bounds (2.17) are valid for every time with the same inverse temperature β . Indeed, in the non-linear case, this parameter is downgraded over time until vanishing in finite time [15, Section 5].

Technically, the downgrading of this norm parameter stems from the fact that to control the collision operators, the sum over the velocities is resorbed thanks to a fraction of the sub-Gaussian decreasing, eventually providing the following continuity estimates.

Proposition 2.4 (Continuity of the successive-collision operators).

There exists a constant C_d depending only on the dimension such that for all $n, s \in \mathbb{N}^$, and all times $t > 0$, fixing tags $\underline{\ell}_s \in \Lambda_k$ and choosing two inverse temperature $\beta' < \beta$, we have*

$$f_{n+s} \in \mathcal{F}_{n+s,\beta} \Rightarrow Q_{n,\underline{\ell}_s}(t) f_{n+s} \in \mathcal{F}_{n,\beta'}, \text{ with} \quad (2.19)$$

$$\|Q_{n,\underline{\ell}_s}(t) f_{n+s}\|_{n,\beta'} \leq e^n \left(\frac{C_d t}{\sqrt{\beta^d (\beta - \beta')}} \right)^s \|f_{n+s}\|_{n+s,\beta}.$$

For fixed tags, the proof of this proposition is exactly the same as in the paper [12] about non-ideal Rayleigh gas, adapted from the article by Bodineau, Gallagher and Saint-Raymond, which derived the hydrodynamic limit of this gas [5]. Compared to this proof, we merely chose to write $\beta' = \beta \left(1 - \frac{1}{b}\right)$ for better clarity in what follows. One can find it written in Fougères' thesis [14].

2.5 Adaptive tree pruning of the Dyson expansion for long times

The continuity estimates presented in the previous section justify the wellposedness of the Dyson series (2.1) for short times. To perform a derivation for long times, we will iterate the Dyson series, but each iteration will make appear a factor e^n stemming from the estimate of Proposition 2.4. These factors stack, so that, without further adjustment, at each iteration the successive times would decrease extremely fast and their sum would eventually be summable, leaving the derivation on a finite short time (see [14] for details). The method we use here, introduced in [5], consists in putting aside these stacking factors e^n and to control them by bounding the number of collisions appearing in the Dyson series.

More precisely, for a fixed time $t > 1$ we will split the time interval $[0, t]$ into K pieces and impose a piece-dependent amount of collisions on each small interval of this cutting. The number K will be tuned in the end of the proof, and henceforth we write

$$t = \sum_{i=1}^K h_i, \text{ with time steps } t_k^D \doteq t - \sum_{j=1}^k h_j, \quad (2.20)$$

so that $t_K^p = 0$. Like in [12], and contrary to [5], we will not choose a uniform cutting, but an adaptive one. At the k -th time quantum of length h_k , we want at most 2^k collisions to have happened; we prune the collision tree every time it becomes more than exponentially big. Explicitly, between t and $t_1^p = t - h_1$, we first truncate the Dyson series (2.1) to 2 collisions, then expand it again between $t - h_1$ and $t - h_2$ truncated to 2^2 collisions, and iterate this process K times.

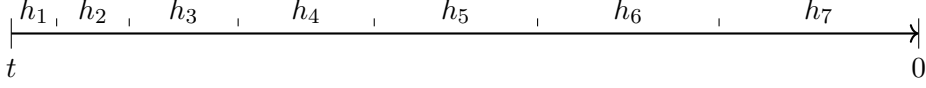


Figure 1: Backwards division of the time interval under study

We denote the step number of added tagged particles and the step total number of particles

$$L_k \doteq \sum_{i=1}^k |\ell_{j_i}^*| \quad \text{and} \quad N_k \doteq n + j_1 + \dots + j_k. \quad (2.21)$$

This yields, similarly as in [5, 12], the following pruned expansion

$$F_n^\varepsilon(t) = \sum_{\substack{j_i \leq 2^i \\ \ell_{j_i}^* \in \Lambda_{j_i}}} p_\mu^{L_K} Q_{n, \ell_{j_1}^*}(h_1) \dots Q_{N_{K-1}, \ell_{j_K}^*}(h_K) F_{N_K}^\varepsilon(0) + R_n^{[K]}(t) \quad (2.22)$$

where the remainder is defined as

$$R_n^{[K]}(t) \doteq \sum_{k=1}^K \sum_{\substack{j_i \leq 2^i \\ \ell_{j_i}^* \in \Lambda_{j_i}}} Q_{n, \ell_{j_1}^*}(h_1) \dots Q_{N_{k-2}, \ell_{j_{k-1}}^*}(h_{k-1}) \sum_{j_k > 2^k} \sum_{\ell_{j_k}^* \in \Lambda_{j_k}} p_\mu^{L_k} Q_{N_{k-1}, \ell_{j_k}^*}(h_k) F_{N_k}^\varepsilon(t_k^p). \quad (2.23)$$

Based on the limit expansion (2.8), one can write the same decomposition for the limit family $(G_n)_{n \geq 0}$:

$$G_n(t) = \sum_{(j_i \leq 2^i)_{1 \leq i \leq K}} Q_{n, \underline{0}_{j_1}}^{\text{lim}}(h_1) \dots Q_{N_{K-1}, \underline{0}_{j_K}}^{\text{lim}}(h_K) G_{N_K}(0) + R_n^{[K], \text{lim}}(t). \quad (2.24)$$

Let us denote $\widehat{G}_n(t) \doteq G_n(t) - R_n^{[K], \text{lim}}(t)$ the limit pruned expansion. Since the chosen condition of a sub-exponential number of collisions is very restrictive at first, and then gradually relaxed, the adaptive cutting times are chosen small at first and then progressively bigger (see Fig. 1). The bound on the pruned-out remainder is given in the following proposition.

Proposition 2.5 (Estimate of the pruned-out term). *With the previous notation, for any choice of power $\alpha \in (0, 3/4)$, and K large enough satisfying $t \lesssim K^{\frac{3}{4}-\alpha}$, a good choice of time cutting*

$$\underline{h} = (h_1, \dots, h_K)$$

provides the following estimate

$$\left\| R_n^{[K]}(t) \right\|_{\mathbb{L}^\infty(\mathcal{D}^d)} + \left\| R_n^{[K], \text{lim}}(t) \right\|_{\mathbb{L}^\infty(\mathcal{D}^d)} \leq C^{C_0 \lambda} n^{cn} e^{-2^{K-K\alpha}}. \quad (2.25)$$

Note that this estimate imposes a technical condition on the mixed scaling of λ and K for the error to be small; we will tune this scaling in Section 2.7. The factor n^{cn} stems from the fact that we generalize the result to all the correlation functions, and not only the first one. Indeed, one will see that this factor is due to the usual bound C^n stacking at each iteration of the cutting.

Proof. The proof is very similar to the one found in [12]. We give it for the hard sphere version, since the limit version is identical. Using the a priori estimates (2.16) on the densities and the continuity estimate on successive-collision operators given in Proposition 2.4, at given $(\ell_{j_i}^*)_{i \leq k-1}$ one has for every $k \in \llbracket 1, K \rrbracket$

$$\begin{aligned} \left\| \sum_{j_k > 2^k} \sum_{\ell_{j_k}^* \in \Lambda_{j_k}} p_\mu^{L_k} Q_{N_{k-1}, \ell_{j_k}^*} (h_k) F_{N_k}^\varepsilon (t_k^{\text{p}}) \right\|_{N_{k-1}, \beta/2} &\leq e^{N_{k-1}} \sum_{j_k > 2^k} \sum_{\ell_{j_k}^* \in \Lambda_{j_k}} p_\mu^{L_k} \left(\frac{\sqrt{2} C_d h_k}{\beta^{(d+1)/2}} \right)^{j_k} \left\| F_{N_k}^\varepsilon (t_k^{\text{p}}) \right\|_{N_k, \beta} \\ &\leq e^{N_{k-1}} \sum_{j_k > 2^k} \sum_{\ell_{j_k}^* \in \Lambda_{j_k}} p_\mu^{L_k} \left(\frac{\sqrt{2} C_d h_k}{\beta^{(d+1)/2}} \right)^{j_k} C^{N_k} C^{C_0 \lambda}. \end{aligned}$$

For μ large enough in the scaling $(S_{\varepsilon, \mu, \lambda})$, we have $p_\mu \leq 1$, so that the sum over $\ell_{j_k}^* \in \Lambda_{j_k}$ only gives a factor 2^{N_k} that can be resorbed in the term C^{N_k} . We then iterate Proposition 2.4, downgrading the parameter $\beta/2$ by $\beta/(4k)$ at each step, so that it remains greater than $\beta/4$. Hence we can write, grouping $C^{N_{k-1}}$ and all the appearing terms of the form e^{N_i} together as a power of a constant \hat{C} , that

$$\begin{aligned} &\left\| Q_{n, \ell_{j_1}^*} (h_1) \dots Q_{N_{k-2}, \ell_{j_{k-1}}^*} (h_{k-1}) \sum_{j_k > 2^k} \sum_{\ell_{j_k}^* \in \Lambda_{j_k}} p_\mu^{L_k} Q_{N_{k-1}, \ell_{j_k}^*} (h_k) F_{N_k}^\varepsilon (t_k^{\text{p}}) \right\|_{n, \beta/4} \quad (2.26) \\ &\leq C^{C_0 \lambda} \hat{C}^{\sum_{i=0}^{k-1} N_i} \left(\left(\frac{4}{\beta} \right)^{\frac{d+1}{2}} \sqrt{4k} C_d h_1 \right)^{j_1} \dots \left(\left(\frac{4}{\beta} \right)^{\frac{d+1}{2}} \sqrt{4k} C_d h_{k-1} \right)^{j_{k-1}} \sum_{j_k > 2^k} \left(\frac{\sqrt{2} C_d h_k}{\sqrt{\beta}} \right)^{j_k}. \end{aligned}$$

We now observe, on the one hand, that recalling notation (2.21) for N_i and since for $i \leq k-1$, $j_i \leq 2^i$, we have $\sum_{i=0}^{k-1} N_i \leq nk + 2^{k+1}$. On the other hand, we can also put aside from the sum the following factors

$$\left(\left(\frac{4}{\beta} \right)^{\frac{d+1}{2}} 2k^{\frac{1}{4}-\alpha} C_d \right)^{\sum_{i=0}^{k-1} j_i} \leq (\hat{C}_\beta k^{\frac{1}{4}-\alpha})^{2^k}.$$

This computation is what allows us to gain a power $\frac{1}{4}$ on the time scaling that we imposed in Theorem 1, compared to [12]. Indeed, in the following, the computation above harnesses the full power decay of the last time interval h_k , whereas some of it was lost in [12]. In the end, for a possibly larger constant C depending on d and β , we get

$$\left\| R_n^{[K]}(t) \right\|_{\mathbb{L}^\infty} \leq C^{C_0 \lambda} \sum_{k=1}^K C^{nk} (C k^{\frac{1}{4}-\alpha})^{2^k} \sum_{j_1=0}^2 \dots \sum_{j_{k-1}=0}^{2^{k-1}} (k^{\frac{1}{4}+\alpha} h_1)^{j_1} \dots (k^{\frac{1}{4}+\alpha} h_{k-1})^{j_{k-1}} \sum_{j_k > 2^k} (C h_k)^{j_k}.$$

Eventually, we consider similarly as in [12], for all $1 \leq i \leq K$,

$$\tilde{h}_i \doteq \frac{e^{-2(K-K^{1-\alpha-i})}}{2K^{\frac{1}{4}+\alpha}} \leq \frac{1}{2K^{\frac{1}{4}+\alpha}}, \quad (2.27)$$

and renormalize them such that

$$h_i \doteq \frac{t}{\sum_{j=1}^K \tilde{h}_j} \tilde{h}_i \leq \tilde{h}_i.$$

Indeed, as soon as $t \lesssim K^{\frac{3}{4}-2\alpha}$, we can write

$$\begin{aligned} \frac{1}{t} \sum_{i=1}^K \tilde{h}_i &\geq \frac{1}{t} \sum_{j=0}^{\lfloor K^{1-\alpha} \rfloor} \tilde{h}_{K-j} \\ &\geq \frac{1}{t} \sum_{j=0}^{\lfloor K^{1-\alpha} \rfloor} \frac{e^{-1}}{2K^{\frac{1}{4}+\alpha}} \geq 1. \end{aligned}$$

Now, the time interval lengths $(h_i)_{1 \leq i \leq K}$ cover t as imposed by (2.20), and their choice provides, summing the geometric series over $(j_i)_{1 \leq i \leq K}$,

$$\begin{aligned} \left\| R_n^{[K]}(t) \right\|_{\mathbb{L}^\infty} &\leq C^{C_0\lambda} \sum_{k=1}^K C^{nk} (CK^{\frac{1}{4}-\alpha})^{2k} \times 2^k \sum_{j_k > 2^k} \left(\frac{e^{-2K-K^{1-\alpha}-k}}{2K^{\frac{1}{4}+\alpha}} \right)^{j_k} \\ &\leq C^{C_0\lambda} \sum_{k=1}^K C^{mk} \left(\frac{CK^{\frac{1}{4}-\alpha}}{K^{\frac{1}{4}+\alpha}} \right)^{2k} \left(\frac{e^{-2K-K^{1-\alpha}-k}}{2} \right)^{2k}. \end{aligned}$$

Observe now that there exists a constant c such that $(k \geq c \log n) \Rightarrow (nk \leq 2^k)$, so that in this case the factor C^{nk} is absorbed by C^{2k} , and for $k \leq c \log n$, then $C^{nk} \leq C^{cn \log n} = n^{(c \log C)n}$. Now, the denominator $K^{\frac{1}{4}+\alpha}$ crushes the term $CK^{\frac{1}{4}-\alpha}$ for K large enough and we end up with

$$\left\| R_n^{[K]}(t) \right\|_{\mathbb{L}^\infty} \leq C^{C_0\lambda} \tilde{C}^{n \log n} e^{-2K-K^{1-\alpha}}.$$

This completes the proof of Proposition 2.5. \square

2.6 Discarding trajectories implying several labelled particles

Since the limit pseudo-trajectories defined from the limit hierarchy (2.10) only imply collisions with particles at equilibrium (which are in wide majority), we must get rid of the ones including collisions with perturbed tagged particles. We hence write our pruned expansion as

$$\begin{aligned} &\sum_{\substack{(j_i \leq 2^i) \\ (\ell_{j_i}^* \in \Lambda_{j_i})}}_{1 \leq i \leq K} p_\mu^{L_K} Q_{n, \ell_{j_1}^*} (h_1) \dots Q_{N_{K-1}, \ell_{j_K}^*} (h_K) F_{N_K}^\varepsilon (0) \\ &= \sum_{(j_i \leq 2^i)_{1 \leq i \leq K}} Q_{n, 0_{j_1}} (h_1) \dots Q_{N_{K-1}, 0_{j_K}} (h_K) F_{N_K}^\varepsilon (0) \end{aligned} \quad (2.28)$$

$$\begin{aligned} &+ \sum_{(j_i \leq 2^i)} \sum_{k=1}^K \sum_{\substack{(\ell_{j_i}^*)_{i \leq K} \\ \ell_{j_k}^* \neq 0_{j_k}}} p_\mu^{L_K} Q_{n, \ell_{j_1}^*} (h_1) \dots Q_{N_{K-1}, \ell_{j_K}^*} (h_K) F_{N_K}^\varepsilon (0), \quad (2.29) \\ &\doteq \widehat{F}_n^\varepsilon (t) + F_n^{\varepsilon, \text{u.e.}} (t), \end{aligned}$$

where we denote $\widehat{F}_n^\varepsilon (t)$ the main term (2.28) containing only collisions with equilibrium, and $F_n^{\varepsilon, \text{u.e.}} (t)$ the one implying unwanted encounters (2.29). We bound the latter in the following proposition.

Proposition 2.6 (Encountering tagged particles is rare). *As long as*

$$p_\mu \leq \frac{1}{2^K},$$

in the same time setting as in Proposition 2.5, the unwanted pseudotrajectories implying tagged particles are bounded by

$$\|F_n^{\varepsilon, \text{u.e.}} (t)\|_{\mathbb{L}^\infty(\mathcal{D}^d)} \leq C^{mK+A^K} p_\mu. \quad (2.30)$$

The reader may think of the factor C^{A^K} as a small negative power of ε in the final scaling, which we will compute in the following Section 2.7: the scaling proportion p_μ will have to compensate it.

Proof of the proposition. A computation directly adapted from the initial proximity bound (2.18) leads to the estimate

$$\left\| F_{N_K}^\varepsilon(0, \tilde{\ell}_{N_K}) \right\|_{N_K, \beta} \leq (C_0)^{N_K},$$

so that, using the binomial identity

$$\sum_{\ell_{j_i}^* \in \Lambda_{j_i}} p_\mu^{|\ell_{j_i}^*|} = (1 + p_\mu)^{j_i}$$

and the same continuity estimates on the successive-collision operators as in the proof of Proposition 2.5, one has

$$\|F_n^{\varepsilon, \text{u.e.}}(t)\|_{\mathbb{L}^\infty(\mathcal{D}^d)} \leq C^{nK} (CK^{\frac{1}{4}-\alpha})^{2^K} \sum_{(j_i \leq 2^i)_{i \leq K}} \sum_{k=1}^K \left((1 + p_\mu)^{j_k} - 1 \right) \prod_{i=1}^K \left(\frac{1}{2} \right)^{j_i}.$$

Notice on the one hand that

$$\begin{aligned} (CK^{\frac{1}{4}-\alpha})^{2^K} &= \exp \left[2^K \log(CK^{\frac{1}{4}-\alpha}) \right] \\ &\leq \exp \left(A^K \right) \end{aligned}$$

for any $A > 2$ as long as K is large enough (depending on A). On the other hand, using $j_k \leq 2^k$ and taking $p_\mu \leq 2^{-K} \leq 2^{-k}$, we have by convexity on $[0, 2^{-k}]$ that

$$\begin{aligned} (1 + p_\mu)^{j_k} - 1 &\leq (1 + p_\mu)^{2^k} - 1 \\ &\leq (e - 1) 2^k p_\mu, \end{aligned}$$

so that eventually

$$\begin{aligned} \|F_n^{\varepsilon, \text{u.e.}}(t)\|_{\mathbb{L}^\infty(\mathcal{D}^d)} &\leq (e - 1) p_\mu \times C^{nK+A^K} \sum_{k=1}^K 2^k \sum_{(j_i \leq 2^i)_{i \leq K}} \prod_{i=1}^K \left(\frac{1}{2} \right)^{j_i} \\ &\leq p_\mu \widehat{C}^{nK+A^K} \end{aligned}$$

for another constant \widehat{C} absorbing the factor $(e - 1) 2^{K+1} \times 2^K$, concluding the proof. \square

2.7 Proof of the convergence

Now that the pruned-out terms have been controlled, we want to compare the pruned terms $\widehat{F}_n^\varepsilon(t)$ and $\widehat{G}_n(t)$, which is the last step before our choice of scaling for K and the conclusion of Theorem 1. As explained in Section 2.1, our strategy relies on considering pseudo-trajectories without recollisions. This method is an adaptation of [5, Section 5] which is now classical; it follows from several approximations: an energy truncation and a time separation of the collisions are operated, so as to be able to construct a small set of bad collision parameters, outside of which there will be no recollision.

This time separation method does not yield the best quantitative estimates, but in our study in long time the worst error is made with the pruned-out terms, so that here we use this method anyway. This allows concision on the one hand, and on the other hand one can thus compare it with the optimized method presented in the companion paper [13].

Like in [12], we still refine the quantitative bounds using the same control on

$$Q_{n, \underline{\ell}_{j_1}^*}^*(h_1) \dots Q_{N_K-1, \underline{\ell}_{j_K}^*}^*(h_K)$$

as in Propositions 2.5 and 2.6, which induces our factor $C_\beta^{nK+A^K}$, instead of the original crude bound with $|Q|_{1, J_k}(t)$, which gave a factor $(Ct)^{2^K}$ (see [5]). From a physical perspective, we decompose the time interval into small pieces whose lengths are adapted to the maximum number of particles that may appear in them, so that the dynamics behaves similarly during each one of them. Hence, as long as time does not get too big with respect to the number of pieces, none of the estimates depends on the total time length.

Pseudo-trajectory formulation First of all, let us notice that the pruned expansion (2.22) has a pseudo-trajectory formulation similar to the original one (2.6), summing over the successive numbers of collisions $(j_i \leq 2^i)_{i \leq K}$, with the following additional condition on the collision times, located between the time steps (2.20):

$$\underline{t}_{J_K} \in T_{\underline{j}_K}(t) \doteq \left\{ (t_1, \dots, t_{J_K}) \in T_{J_K}(t), \{t_{J_k+1}, \dots, t_{J_{k+1}}\} \subset [t_{k+1}^p, t_k^p] \right\},$$

where $J_K \doteq N_K - n = j_1 + \dots + j_K$ denotes the step number of collisions. As announced above, we start by computing a few approximations, initiated by an energy truncation, so as to work with bounded velocities.

Energy truncation We hence consider the following functionals with truncated energy

$$\begin{aligned} \widehat{F}_n^{\varepsilon, [\mathbf{V}]}(t) \doteq & \sum_{\underline{j}_K} \sum_{\substack{\underline{x}_{J_K} \\ L_K=0}} \int_{T_{\underline{j}_K}(t)} dt_{J_K} \int d\underline{\omega}_{J_K} dv_{n+1} \dots dv_{N_K} \prod_{i=1}^{J_K} s_i \langle \omega_i, v_{n+i} - v_{m_i}^{[t_i^+]} \rangle_+ \\ & \times F_{N_K}^\varepsilon(0, \underline{z}_{N_K}^{[0]}) \mathbf{1}_{\left\| \underline{v}_{N_K}^{[0]} \right\|^2 \leq \mathbf{V}^2}, \end{aligned} \quad (2.31)$$

and similarly the truncated limit functions $\widehat{G}_n^{[\mathbf{V}]}$. The error made by truncating the velocities is

$$\widehat{F}_n^\varepsilon(t) - \widehat{F}_n^{\varepsilon, [\mathbf{V}]}(t) = \sum_{(j_i \leq 2^i)} Q_{n, 0_{j_1}} \dots Q_{N_K-1, 0_{j_K}}(h_K) \left[F_{N_K}^\varepsilon(0) \mathbf{1}_{\left\| \underline{v}_{N_K}^{[0]} \right\|^2 \geq \mathbf{V}^2} \right],$$

with the following estimate on the initial functional on the right:

$$\begin{aligned} \left\| F_{N_K}^\varepsilon(0, \underline{z}_{N_K}^{[0]}, \tilde{\underline{z}}_{N_K}) \mathbf{1}_{\left\| \underline{v}_{N_K}^{[0]} \right\|^2 \geq \mathbf{V}^2} \right\|_{N_K, \beta/2} & \leq \sup_{\underline{u}_{N_K} \in \mathbb{R}^{dN_K}} \left| F_{N_K}^\varepsilon(0) e^{\beta \|\underline{u}_{N_K}\|^2} e^{-\frac{\beta}{2} \|\underline{u}_{N_K}\|^2} \mathbf{1}_{\|\underline{u}_{N_K}\|^2 \geq \mathbf{V}^2} \right| \\ & \leq \left\| F_{N_K}^\varepsilon(0) \right\|_{N_K, \beta} e^{-\frac{\beta}{2} \mathbf{V}^2}. \end{aligned}$$

Hence, applying the same bounds as in Propositions 2.5 and 2.6 in the same setting, we end up with the following lemma.

Lemma 2.1 (Energy truncation error). *The error due to the energy truncation is bounded by*

$$\left\| \widehat{F}_n^\varepsilon(t) - \widehat{F}_n^{\varepsilon, [\mathbf{V}]}(t) \right\|_{\mathbb{L}^\infty} \leq C^{nK+A^K} \exp\left(-\frac{\beta}{2} \mathbf{V}^2\right). \quad (2.32)$$

The same holds for its limit version $\widehat{G}_n(t) - \widehat{G}_n^{[\mathbf{V}]}(t)$.

Time separation We now need the successive collisions to be separated enough in time, to avoid pathological geometric recollisions. Like in the previous section, let us define

$$\widehat{F}_n^{\varepsilon, [\mathbf{V}, \delta]}(t) \doteq \sum_{\underline{j}_K} \sum_{\substack{\underline{x}_{J_K} \\ L_K=0}} \int_{T_{\underline{j}_K}^{[\delta]}(t)} dt_{J_K} \int d\underline{\omega}_{J_K} dv_{n+1} \dots dv_{N_K} \prod_{i=1}^{J_K} s_i \langle \omega_i, v_{n+i} - v_{m_i}^{[t_i^+]} \rangle_+ \quad (2.33)$$

$$\times F_{N_K}^\varepsilon(0, \underline{z}_{N_K}^{[0]}) \mathbb{1}_{\left\| \underline{v}_{N_K}^{[0]} \right\|^2 \leq \mathbf{V}^2},$$

with the separation condition encoded in the following time set, over which we integrate,

$$T_{\underline{j}_K}^{[\delta]}(t) = \left\{ \underline{t} \in T_{\underline{j}_K}(t), t_{i-1} - t_i > \delta \right\}. \quad (2.34)$$

The limit version $\widehat{G}_n^{[\mathbf{V}, \delta]}$ is defined by the same time restriction. The error of time separation is

$$\widehat{F}_n^{\varepsilon, [\mathbf{V}]} - \widehat{F}_n^{\varepsilon, [\mathbf{V}, \delta]} \doteq \sum_{\underline{j}_K} \sum_{\substack{\underline{x}_{J_K} \\ L_K=0}} \int \left(T_{\underline{j}_K}^{[\delta]}(t) \right)^c dt_{J_K} \int d\underline{\omega}_{J_K} dv_{n+1} \dots dv_{N_K}$$

$$\times \prod_{i=1}^{J_K} s_i \langle \omega_i, v_{m_i}^{[t_i^+]} - v_{n+i} \rangle_+ \times F_{N_K}^\varepsilon(0, \underline{z}_{N_K}^{[0]}) \mathbb{1}_{\left\| \underline{v}_{N_K}^{[0]} \right\|^2 \leq \mathbf{V}^2},$$

where one can write

$$\left(T_{\underline{j}_K}^{[\delta]}(t) \right)^c = \bigcup_{i=1}^{J_K-1} \left\{ \underline{t} \in T_{\underline{j}_K}(t), t_i - t_{i+1} \leq \delta \right\}. \quad (2.35)$$

Now, the integral in time over one of these sets, using the same method as before, changes the estimate of the corresponding successive-collision operator from $\frac{h_k^{j_k}}{j_k!}$ to $\frac{\delta h_k^{j_k-1}}{(j_k-1)!}$. Since the loss of $\frac{h_k}{j_k}$ and the factor (j_k-1) coming from the union (2.35) easily resorb into the bigger factor C^{A^K} , we get in the end the following lemma.

Lemma 2.2 (Time separation error). *For this error, we have the following estimate*

$$\left\| \widehat{F}_n^{\varepsilon, [\mathbf{V}]} - \widehat{F}_n^{\varepsilon, [\mathbf{V}, \delta]} \right\|_{\mathbb{L}^\infty} \leq C^{mK+A^K} \delta. \quad (2.36)$$

Once again, the same holds for its limit version $\widehat{G}_n^{[\mathbf{V}]} - \widehat{G}_n^{[\mathbf{V}, \delta]}$ for similar reasons.

Restriction to non-pathological collision parameters Finally, at fixed ε , we have to restrict the collision parameters to non-pathological configurations, leading to no recollision during the transport flow. First, to compute the convergence of the n -th marginal, we have to consider final configurations $\underline{z}_n = \underline{z}_n^{[t]}$ that do not directly lead to recollisions, i.e. that belong to the following set of past-excluding configurations, with some extra room $\varepsilon_d \doteq \varepsilon^{\frac{d}{d+1}}$:

$$\mathcal{E}_n^t(\varepsilon_d) \doteq \left\{ \underline{z}_n \in \mathcal{D}_n^\varepsilon \mid \forall \tau \in [0, t], (\underline{x}_n - \tau \underline{v}_n) \in \mathcal{D}_n^{\varepsilon_d} \right\}. \quad (2.37)$$

This set is the whole domain for $n=1$, like in [5, 12]. Then, the restriction on the collision parameters is done based on the following geometric result, proved in [15, 5] and formalizing the arguments of Lanford [20], based on billiards theory.

Similarly to the notation $\underline{z}_n^{[\tau]} = (\underline{x}_n^{[\tau]}, \underline{v}_n^{[\tau]})$ for the hard sphere pseudo-trajectories, let us denote $\underline{z}_n^{[\tau]} = (\underline{y}_n^{[\tau]}, \underline{u}_n^{[\tau]})$ the limit pseudo-trajectories, and $N[\tau]$ the number of particles in the trajectory at time $\tau \in [0, t]$, from $N[t] = n$ to $N[0] = N_K$. Hence, the following lemma asserts that once the previous truncations computed, choosing collision parameters away from a set of small measure, the hard sphere and limit pseudo-trajectories are easy to compare.

Lemma 2.3. *Considering a history $(\underline{j}_K, \chi_{J_K})$ with collision times $t_{J_K} \in T_{\underline{j}_K}^{[\delta]}(t)$ (δ -separated) and a final configuration $\underline{z}_n^{[t]}$, given a maximum energy $\mathbf{V}^2 > 0$, there exists a set of pathological collision parameters*

$$\Pi(\underline{z}_n^{[t]}, \underline{j}_K, \chi_{J_K}) \subset (\mathbb{S}^{d-1} \times \mathbb{R}^d)^{J_K}$$

with small volume

$$|\Pi(\underline{z}_n^{[t]}, \underline{j}_K, \chi_{J_K})| \leq C J_K N_K \left(\varepsilon^{\frac{d}{d+2}} + \mathbf{V}^d \times \varepsilon^{\frac{d-1}{3(d+1)}} + \mathbf{V}^{\frac{d+1}{2}} \left(\frac{\varepsilon^{\frac{d}{d+1}}}{\delta} \right)^{\frac{d-1}{2}} \right), \quad (2.38)$$

and such that, assuming

- i. the collision parameters are non-pathological: $(\underline{\omega}_{J_K}, v_{n+1}, \dots, v_{N_K}) \notin \Pi(\underline{z}_n^{[t]}, \underline{j}_K, \chi_{J_K})$
- ii. the energy of the corresponding pseudo-trajectory remains bounded: $\|\underline{v}_{N_K}^{[0]}\|^2 \leq \mathbf{V}^2$
- iii. the final configuration is past-excluding for the free-flow: $\underline{z}_n^{[t]} \in \mathcal{E}_n^t(\varepsilon_d)$,

then

1. the hard sphere pseudo-positions remain sufficiently far away: $\forall \tau \in [0, t], \underline{x}_{N[\tau]}^{[\tau]} \in \mathcal{D}_{N[\tau]}^{\varepsilon_d/2}$
2. the velocities of the hard sphere and limit trajectories coincide: $\forall \tau \in [0, t], \underline{v}_{N[\tau]}^{[\tau]} = \underline{v}_{N[\tau]}^{[0]}$
3. the positions of both trajectories remain close: $\forall \tau \in [0, t], \forall i \leq N[\tau], d(x_i^{[\tau]}, y_i^{[\tau]}) \leq J_K \varepsilon$.

Let us observe that 3. is a consequence of 2. since, when the velocities coincide, the only difference between the positions is the shift of $\varepsilon \omega_i$ that happens at each particle adjunction. Furthermore, 2. is a consequence of 1. since when the collision parameters are the same, and if the particles do not collide between the particle adjunctions, the velocities of both pseudo-trajectories are identically determined. This result has been proved [15, 5] for a one-species gas, but the dynamics is strictly identical for our mixture gas.

Hence, we consider the following functional restricted to non-pathological collision parameters

$$\begin{aligned} \tilde{F}_n^{\varepsilon, [\mathbf{V}, \delta]}(t) &\doteq \sum_{\underline{j}_K} \sum_{\substack{\chi_{J_K} \\ L_K=0}} \int_{T_{\underline{j}_K}^{[\delta]}(t)} dt_{J_K} \int_{\Pi(\underline{z}_n^{[t]}, \underline{j}_K, \chi_{J_K})^c} d\omega_{J_K} dv_{n+1} \dots dv_{N_K} \\ &\quad \times \prod_{i=1}^{J_K} s_i \langle \omega_i, v_{n+i} - v_{m_i}^{[t_i^+]}\rangle_+ \times F_{N_K}^\varepsilon(0, \underline{z}_{N_K}^{[0]}) \mathbf{1}_{\|\underline{v}_{N_K}^{[0]}\|^2 \leq \mathbf{V}^2}, \end{aligned} \quad (2.39)$$

and its limit version $\tilde{G}_n^{[\mathbf{V}, \delta]}$. Now, the error $\tilde{F}_n^{\varepsilon, [\mathbf{V}, \delta]} - \tilde{G}_n^{[\mathbf{V}, \delta]}$ is supported on $\Pi(\underline{z}_n^{[t]}, \underline{j}_K, \chi_{J_K})$, so that we will use the control on its volume (2.38) to control the successive-collision operators, concluding the bounds with the usual computation. More precisely, in the proof [12] of Proposition 2.4, we bound the collision operators (1.12) in the following way:

$$\begin{aligned} e^{\beta \|\underline{v}_j\|^2} \left| \mathcal{C}_j^\ell f_{j+1} \right| &\leq \sum_{i=1}^j \int d\omega_j dv_{j+1} |v_{j+1} - v_i| \cdot \|f_{j+1}\|_{j+1, \beta'} \times e^{-(\beta' - \beta) \|\underline{v}_j\|^2 - \beta' |v_{j+1}|^2} \\ &\leq \|f_{j+1}\|_{j+1, \beta'} \int d\omega_j dv_{j+1} \left(j |v_{j+1}| + \sum_{i=1}^j |v_i| \right) e^{-(\beta' - \beta) \|\underline{v}_j\|^2 - \beta' |v_{j+1}|^2}. \end{aligned}$$

By the Cauchy-Schwarz inequality, we used to bound the quantity above as

$$\begin{aligned} e^{\beta\|v_j\|^2} \left| \mathcal{C}_j^\ell f_{j+1} \right| &\leq \|f_{j+1}\|_{j+1, \beta'} \int d\omega_j dv_{j+1} \left(j |v_{j+1}| + \sqrt{\frac{j}{2e(\beta' - \beta)}} \right) e^{-\beta'|v_{j+1}|^2} \\ &\leq \|f_{j+1}\|_{j+1, \beta'} \left(j \frac{c_d}{\sqrt{\beta'^{d+1}}} + \sqrt{\frac{j}{2e(\beta' - \beta)}} \frac{\tilde{c}_d}{\sqrt{\beta'^d}} \right), \end{aligned}$$

for some constants c_d, \tilde{c}_d depending only on the dimension. For the present bound, one can write instead

$$\int d\omega_j dv_{j+1} \left(j |v_{j+1}| + \sqrt{\frac{j}{2e(\beta' - \beta)}} \right) e^{-\beta'|v_{j+1}|^2} \leq \int d\omega_j dv_j \left(\frac{1}{\sqrt{2e\beta}} + \sqrt{\frac{e^{-1}}{\beta - \beta'}} \right),$$

making appear the volume of collision parameters bounded in (2.38), leading to the following estimate (the factors J_K and N_K resorb in the bigger factor C^{nK+A^K} , for a slightly different constant C depending on β).

Lemma 2.4 (Error of restriction to non-pathological collision parameters). *The restriction of considering only collision parameters chosen so as to avoid recollisions, leads to an error of order*

$$\|\tilde{F}_n^{\varepsilon, [\mathbf{V}, \delta]} - \widehat{F}_n^{\varepsilon, [\mathbf{V}, \delta]}\|_{\mathbb{L}^\infty} \leq C^{nK+A^K} \left(\varepsilon^{\frac{d}{d+2}} + \mathbf{V}^d \times \varepsilon^{\frac{d-1}{3(d+1)}} + \mathbf{V}^{\frac{d+1}{2}} \left(\frac{\varepsilon^{\frac{d}{d+1}}}{\delta} \right)^{\frac{d-1}{2}} \right), \quad (2.40)$$

and similarly for the limit $\widehat{G}_n^{[\mathbf{V}]} - \widehat{G}_n^{\varepsilon, [\mathbf{V}, \delta]}$ from the same computation.

Harnessing initial proximity Now that we have constructed approximations of our distributions that avoid recollisions, we can at last compare both the BBGKY and limiting distributions thanks to the coupled pseudo-trajectories. Indeed, we can write the coupling

$$\begin{aligned} \tilde{F}_n^{\varepsilon, [\mathbf{V}, \delta]} - \tilde{G}_n^{\varepsilon, [\mathbf{V}, \delta]} &= \sum_{\underline{j}_K} \sum_{\underline{x}_{J_K}} p_\mu^{|\underline{\ell}_{J_K}^*|} \int_{T_{\underline{j}_K}^{[\delta]}(t)} dt J_K \int_{\Pi(\underline{z}_n^{[t]}, \underline{j}_K, \underline{x}_{J_K})^c} d\omega_{J_K} dv_{n+1} \dots dv_{N_K} \\ &\quad \times \prod_{i=1}^{J_K} s_i \langle \omega_i, v_{n+i} - v_{m_i}^{[t_i^+]} \rangle_+ \left[F_{N_K}^\varepsilon(0, \underline{z}_{N_K}^{[0]}) - G_{N_K}(0, \underline{\zeta}_{N_K}^{[0]}) \right] \mathbf{1}_{\|\underline{v}_{N_K}^{[0]}\|^2 \leq \mathbf{V}^2}, \end{aligned}$$

with the same collision parameters for both pseudo-trajectories, the hard-sphere one and the limit one. Since by construction $\underline{x}_n^{[t]} = \underline{y}_n^{[t]}$, and by Lemma 2.3 for $\underline{z}_n^{[t]} \in \mathcal{E}_n^t(\varepsilon_d)$, the n first particles also have identical velocities on $[0, t]$, then for all times $\tau \in [0, t]$, we have $\underline{z}_n^{[\tau]} = \underline{\zeta}_n^{[\tau]}$ also in positions. Henceforth, since by the work done in Section 2.6 all the added particles are particles at equilibrium tagged 0, one has (as $\underline{\ell}_n \subset \llbracket 1, n \rrbracket$),

$$\begin{aligned} G_{N_K}(0, \underline{\zeta}_{N_K}^{[0]}) &= M_\beta^{\otimes N_K}(\underline{v}_{N_K}^{[0]}) \varphi_0^{\otimes \ell_n}(\underline{\zeta}_{\ell_n}^{[0]}) \\ &= M_\beta^{\otimes N_K}(\underline{v}_{N_K}^{[0]}) \varphi_0^{\otimes \ell_n}(\underline{z}_{\ell_n}^{[0]}), \end{aligned} \quad (2.41)$$

and so

$$\begin{aligned} \left| F_{N_K}^\varepsilon(0, \underline{z}_{N_K}^{[0]}) - G_{N_K}(0, \underline{\zeta}_{N_K}^{[0]}) \right| &= \left| F_{N_K}^\varepsilon(0, \underline{z}_{N_K}^{[0]}) - G_{N_K}(0, \underline{z}_{N_K}^{[0]}) \right| \\ &\leq (C_0)^{N_K} e^{-\beta\|\underline{v}_{N_K}^{[0]}\|^2} \cdot \varepsilon \end{aligned}$$

by the initial proximity result (2.18). Hence, the previous estimates on the successive-collision operators eventually yield the following result.

Lemma 2.5 (Initial value error). *Conditionally to $z_n \in \mathcal{E}_n^t(\varepsilon_d)$, the initial error is bounded by*

$$\left\| \tilde{F}_n^{\varepsilon, [\mathbf{V}, \delta]} - \tilde{G}_n^{[\mathbf{V}, \delta]} \right\|_{\mathbb{L}^\infty(\mathcal{E}_n^t(\varepsilon_d))} \leq C^{nK+A^K} \varepsilon. \quad (2.42)$$

Coherent choice of truncation parameters (K, \mathbf{V}, δ) Eventually, the last step is to tune the truncation parameters according to ε , so as to obtain the convergence we want. Actually, there is room to choose the scaling, since one can set all the errors as powers of ε , except for the pruning error (2.25) which is significantly bigger. Explicitly, stacking all the errors (pruning (2.25), removing additional tagged particle (2.30), energy truncation (2.32), time separation (2.36) and removing pathological trajectories (2.40)) then using the coupling result (2.42), one has

$$\begin{aligned} \|F_n^\varepsilon(t) - G_n(t)\|_{\mathbb{L}^\infty(\mathcal{E}_n^t(\varepsilon_d))} &\leq C^{C_0 \lambda} n^{cn} e^{-2K-K\alpha} \\ &+ C^{nK+A^K} \left(p_\mu + e^{-\frac{\beta}{2} \mathbf{V}^2} + \delta + \left[\varepsilon^{\frac{d}{d+2}} + \mathbf{V}^d \times \varepsilon^{\frac{d-1}{3(d+1)}} + \frac{\mathbf{V}^{\frac{d+1}{2}} \varepsilon^{\frac{d(d-1)}{2(d+1)}}}{\delta^{\frac{d-1}{2}}} \right] + \varepsilon \right). \end{aligned}$$

Hence, choosing the scaling

$$\delta = \varepsilon^{\frac{d-1}{d+1}}, \quad \text{and } \mathbf{V}^2 = \frac{2}{\beta} |\log \varepsilon|,$$

one gets

$$\begin{aligned} e^{-\frac{\beta}{2} \mathbf{V}^2} + \delta + \mathbf{V}^d \times \varepsilon^{\frac{d-1}{4(d+1)}} + \frac{\mathbf{V}^{\frac{d+1}{2}} \varepsilon^{\frac{d(d-1)}{2(d+1)}}}{\delta^{\frac{d-1}{2}}} &\leq \varepsilon + \varepsilon^{\frac{d-1}{d+1}} + \left| \frac{2}{\beta} \log \varepsilon \right|^{\frac{d}{2}} \varepsilon^{\frac{d-1}{3(d+1)}} + \left| \frac{2}{\beta} \log \varepsilon \right|^{\frac{d+1}{4}} \varepsilon^{\frac{d-1}{2(d+1)}} \\ &\leq \varepsilon^{\frac{d-1}{4(d+1)}} \end{aligned} \quad (2.43)$$

for ε small enough. This way, if we pick

$$K = \left\lfloor \frac{1}{\log A} \log \left(\frac{(d-1) |\log \varepsilon|}{8(d+1) \log C} \right) \right\rfloor,$$

then $C^{AK} \leq \varepsilon^{-\frac{d-1}{8(d+1)}}$, and we can deal with the term C^{nK} like in the proof of Proposition 2.5, yielding the same factor n^{cn} . Hence, for K large enough, denoting $c_\beta \doteq (d-1)/8(d+1) \log C$, we have

$$e^{-2K-K\alpha} \leq e^{-2(1-\alpha)K} \leq \exp \left(-c_\beta |\log \varepsilon|^{(1-\alpha) \frac{\log 2}{\log A}} \right).$$

As $A > 2$ is arbitrary, let us choose A such that $(1-\alpha) \frac{\log 2}{\log A} \geq 1-2\alpha$. This gives us the final condition on

$$\lambda \leq \frac{c_\beta}{2C_0 \log C} |\log \varepsilon|^{1-2\alpha}.$$

Note that we take this scaling so as to have the biggest λ possible in the hypothesis, pushing the associated error at the same level as the truncation error above, which is the limiting one. It yields

$$C^{C_0 \lambda} e^{-2K-K\alpha} \leq \exp \left(-\frac{c_\beta}{2} |\log \varepsilon|^{1-2\alpha} \right).$$

The final verification is the condition $p_\mu \leq 2^{-K}$ of Proposition 2.6, which is satisfied with room to spare considering the choices above. Piling all of this, the pruning error is bigger than all the other ones, and we end up with the very last inequality for ε small enough

$$\|F_n^\varepsilon(t) - G_n(t)\|_{\mathbb{L}^\infty(\mathcal{E}_n^t(\varepsilon_d))} \leq n^{cn} \exp \left(-\frac{c_\beta}{2} |\log \varepsilon|^{1-2\alpha} \right).$$

Eventually, the indicator $\mathbb{1}_{\mathcal{E}_n^t(\varepsilon_d)}$ of the set of past-excluding configurations (2.37) pointwise converges, as ε goes to 0, to the indicator of the following set of full measure

$$\left\{ z_n \in \mathcal{D}^n, \forall 1 \leq i < j \leq n, x_i \neq x_j \text{ and } v_i - v_j \notin \text{Vect}(x_i - x_j) \right\},$$

which concludes the proof of Theorem 1. \square

Appendix

A Cumulants of the exclusion

This section is devoted to the study of the cumulants of the exclusion indicator associated with the hard sphere domain (1.1) (see the companion paper [13] for a generalized definition of the cumulants).

Denoting \mathcal{G}_S the set of graphs on a set S and $\mathcal{G}_n \doteq \mathcal{G}_{[1,n]}$ the set of graphs on $\{1, \dots, n\}$, one may write

$$\begin{aligned} \mathbb{1}_{\mathcal{X}_n^\varepsilon}(x_1, \dots, x_n) &= \prod_{1 \leq i < j \leq n} (1 - \mathbb{1}_{x_i \sim x_j}) \\ &= \sum_{G \in \mathcal{G}_n} \prod_{\{i,j\} \in E_G} (-\mathbb{1}_{x_i \sim x_j}). \end{aligned}$$

One can even further decompose the graphs into their connected components. We denote $\mathcal{C}_S \subset \mathcal{G}_S$ the set of *connected* graphs on S and \mathcal{P}_n the set of partitions of $\{1, \dots, n\}$, so that we can write

$$\mathbb{1}_{\mathcal{X}_n^\varepsilon}(x_1, \dots, x_n) = \sum_{\sigma \in \mathcal{P}_n} \prod_{k=1}^{|\sigma|} \left(\sum_{G_k \in \mathcal{C}_{\sigma_k}} \prod_{\{i,j\} \in E_{G_k}} (-\mathbb{1}_{x_i \sim x_j}) \right).$$

We thus define the cumulants of the exclusion as

$$\phi_k(x_1, \dots, x_k) \doteq \sum_{G \in \mathcal{C}_k} \prod_{\{i,j\} \in E_G} (-\mathbb{1}_{x_i \sim x_j}). \quad (\text{A.1})$$

The very specific structure of these cumulants yields a strong bound on them, called the tree inequality, exposed in the following Proposition. This bound on the integral of the cumulants allows a strong control of the particle correlations, used in Appendix B to gain estimates on the partition function.

Proposition A.1 (Tree inequality).

(i) *The modulus of the cumulants may be controlled restricting the sum defining them to the trees $\mathcal{T}_k \subset \mathcal{C}_k$ (i.e. to the minimally connected graphs) as such:*

$$|\phi_k(x_1, \dots, x_k)| \leq \sum_{T \in \mathcal{T}_k} \prod_{\{i,j\} \in E_T} \mathbb{1}_{x_i \sim x_j}.$$

(ii) *As a consequence, we have the following control over their integral*

$$\int |\phi_k(\underline{x}_k)| d\underline{x}_k \leq k^{k-2} (|\mathcal{B}_d| \varepsilon^d)^{k-1}.$$

Proof. The proof of this proposition relies on a partition scheme due to Penrose [25], and may also be found in [6].

1. The key argument is to find a map $\pi : \mathcal{C}_k \rightarrow \mathcal{T}_k$ such that for any tree $T \in \mathcal{T}_k$, there is a connected graph $R(T) \in \mathcal{C}_k$ satisfying

$$\pi^{-1}(T) = \{G \in \mathcal{C}_k, E_T \subset E_G \subset E_{R(T)}\}.$$

This means that we can partition \mathcal{C}_k into subsets corresponding each to a single tree, and containing all the graphs that are both compatible with this tree and smaller than an upper graph $R(T)$. One can find the construction of this partition scheme for example in [14].

Now, we decompose the sum defining the k -th cumulant according to this mapping, and using its structure we get

$$\begin{aligned} \sum_{G \in \mathcal{C}_k} \prod_{\{i,j\} \in E_G} (-\mathbb{1}_{i \sim j}) &= \sum_{T \in \mathcal{T}_k} \sum_{G \in \pi^{-1}(T)} \prod_{\{i,j\} \in E_G} (-\mathbb{1}_{i \sim j}) \\ &= \sum_{T \in \mathcal{T}_k} \left(\prod_{\{i,j\} \in E_T} (-\mathbb{1}_{i \sim j}) \right) \left(\sum_{E' \subset E_{R(T)} \setminus E_T} \prod_{\{i,j\} \in E'} (-\mathbb{1}_{i \sim j}) \right) \\ &= \sum_{T \in \mathcal{T}_k} \left(\prod_{\{i,j\} \in E_T} (-\mathbb{1}_{i \sim j}) \right) \left(\prod_{\{i,j\} \in E_{R(T)} \setminus E_T} (1 - \mathbb{1}_{i \sim j}) \right) \end{aligned}$$

reversing the usual computation so as to harness the cancellations, which yields the result since $1 - \mathbb{1}_{i \sim j} \in \{0, 1\}$.

2. By the first part of the proposition, the problem is reduced to the integration over the vertices of a tree, for which we know a very simple algorithm. We start integrating over a leaf i_ℓ of T , which appears in only one condition $\mathbb{1}_{i_\ell \sim j}$, and we then iterate for a new leaf until the tree is reduced to a single root, so that

$$\begin{aligned} \int |\phi_k| &\leq \int \sum_{T \in \mathcal{T}_k} \prod_{\{i,j\} \in E_T} \mathbb{1}_{i \sim j} d\mathbf{x}_k \\ &\leq \sum_{T \in \mathcal{T}_k} |\mathcal{B}_d| \varepsilon^d \int \prod_{\{i,j\} \in E_T \setminus \{i_\ell\}} \mathbb{1}_{i \sim j} dx_1 \dots d\check{x}_{i_\ell} \dots dx_k \\ &\leq \sum_{T \in \mathcal{T}_k} (|\mathcal{B}_d| \varepsilon^d)^{k-1} \\ &= k^{k-2} (|\mathcal{B}_d| \varepsilon^d)^{k-1}, \end{aligned}$$

concluding with Cayley's formula, that gives the number of trees on $\llbracket 1, k \rrbracket$. □

B Partition function estimates

This section is dedicated to the proof of Proposition 2.2, which yields the following estimate on the partition function, for μ large enough in our mixed low density scaling $(\mathcal{S}_{\varepsilon, \mu, \lambda})$:

$$\frac{1}{\mathcal{Z}_\mu} \sum_{q,r \geq 0} \frac{(\lambda C_0)^q \mu^r}{q! r!} \int \mathbb{1}_{\mathcal{X}_{q+r}^\varepsilon} \leq C_d^{C_0 \lambda}.$$

First of all, let us use the symmetry of the particles (among each labelled group) to observe that the partition function can be rewritten as

$$\begin{aligned} \mathcal{Z}_\mu &= \sum_{p \geq 0} \sum_{\ell_p \in \Lambda_p} \frac{\lambda^{|\ell_p|} \mu^{p-|\ell_p|}}{p!} \int (M_\beta \varphi_0)^{\otimes \ell_p}(\underline{z}_{\ell_p}) \mathbf{1}_{\mathcal{X}_p^\varepsilon}(\underline{x}_p) d\underline{x}_p d\underline{v}_{\ell_p} \\ &= \sum_{p \geq 0} \sum_{k=0}^p \frac{\lambda^k \mu^{p-k}}{k!(p-k)!} \int (M_\beta \varphi_0)^{\otimes k}(\underline{z}_k) \mathbf{1}_{\mathcal{X}_p^\varepsilon}(\underline{x}_p) d\underline{x}_p d\underline{v}_k \end{aligned} \quad (\text{B.1})$$

$$= \sum_{p, q \geq 0} \frac{\lambda^q \mu^p}{q! p!} \int (M_\beta \varphi_0)^{\otimes q}(\underline{x}_q) \mathbf{1}_{\mathcal{X}_{p+q}^\varepsilon}(\underline{x}_{p+q}) d\underline{x}_{p+q} d\underline{v}_q. \quad (\text{B.2})$$

We now start by proving the computational lemma below, giving an explicit and exact formulation of the partition function using the cumulants of the exclusion, defined in the Appendix A above.

Lemma B.1. *For all $\lambda, \mu > 0$, with (ϕ_k) the cumulants associated to the exclusion $(\mathbf{1}_{\mathcal{X}_k^\varepsilon})$, one has*

$$\mathcal{Z}_\mu = \exp \left(\sum_{(k,l) \neq (0,0)} \frac{\mu^k \lambda^l}{k! l!} \int (M_\beta \varphi_0)^{\otimes l} \phi_{k+l} \right). \quad (\text{B.3})$$

Proof. We start from the formulation (B.2) of the partition function, and we will write the following cumulant expansion, denoting \mathcal{P}_{p+q}^s the partitions of $[[1, p+q]]$ into s subsets,

$$\begin{aligned} \mathcal{Z}_\mu &= 1 + \sum_{p, q \neq 0, 0} \frac{\mu^p \lambda^q}{p! q!} \sum_{s=1}^{p+q} \sum_{\sigma \in \mathcal{P}_{p+q}^s} \int \prod_{k=1}^s \phi_{|\sigma_k|}(\underline{x}_{\sigma_k}) (M_\beta \varphi_0)^{\otimes q}(\underline{x}_q) d\underline{x}_{p+q} \\ &= 1 + \sum_{p, q \neq 0, 0} \frac{\mu^p \lambda^q}{p! q!} \sum_{s=1}^{p+q} \sum_{\substack{(k_i, l_i) \neq (0,0) \\ \sum k_j = p, \sum l_j = q}} \frac{1}{s!} \frac{p!}{k_1! \cdots k_s!} \frac{q!}{l_1! \cdots l_s!} \int \prod_{i=1}^s \phi_{k_i+l_i} (M_\beta \varphi_0)^{\otimes l_i}. \end{aligned}$$

To perform this last equality, denoting $P = [[1, p]]$ and $Q = [[p+1, p+q]]$, we consider the following surjection that counts the number of elements from each partition subset in both P and Q , defined up to an arbitrary choice of a partition order:

$$\begin{aligned} \Phi : \mathcal{P}_{p+q}^s &\longrightarrow \left\{ (\underline{k}_s, \underline{l}_s) \mid (k_i, l_i) \neq (0, 0), \sum k_j = p, \sum l_j = q \right\} \\ \sigma &\longmapsto (|\sigma_1 \cap P|, \dots, |\sigma_s \cap P|, |\sigma_1 \cap Q|, \dots, |\sigma_s \cap Q|). \end{aligned}$$

Its defect of injectivity is indeed given by

$$\# \left\{ \sigma \in \mathcal{P}_{p+q}^s, \Phi(\sigma) = (\underline{k}_s, \underline{l}_s) \right\} = \frac{1}{s!} \cdot \frac{p!}{k_1! \cdots k_s!} \cdot \frac{q!}{l_1! \cdots l_s!},$$

providing the combinatorial factors. Now, permuting the sums so as to get rid of the condition on the value of the index sums, we get

$$\begin{aligned} \mathcal{Z}_\mu &= 1 + \sum_{s \geq 1} \frac{1}{s!} \prod_{i=1}^s \left(\sum_{(k_i, l_i) \neq (0,0)} \frac{\mu^{k_i} \lambda^{l_i}}{k_i! l_i!} \int \phi_{k_i+l_i}(\underline{x}_{k_i+l_i}) (M_\beta \varphi_0)^{\otimes l_i}(\underline{x}_{l_i}) d\underline{x}_{k_i+l_i} \right) \\ &= \exp \left(\sum_{(k_i, l_i) \neq (0,0)} \frac{\mu^{k_i} \lambda^{l_i}}{k_i! l_i!} \int \phi_{k_i+l_i} \cdot (M_\beta \varphi_0)^{\otimes l_i} \right), \end{aligned} \quad (\text{B.4})$$

concluding the proof. \square

Proof of the proposition. Now, Lemma B.1 applied to both the numerator and the denominator—that have the same structure—gives

$$\frac{1}{\mathcal{Z}_\mu} \sum_{p,q \geq 0} \frac{\mu^p (\lambda C_0)^q}{p! q!} \int \mathbb{1}_{\mathcal{X}_{p+q}^\varepsilon} = \exp \left(\sum_{(k,l) \neq 0,0} \frac{\mu^k \lambda^l}{k! l!} \int \phi_{k+l} \cdot [C_0^l - (M_\beta \varphi_0)^{\otimes l}] \right).$$

The terms in the sum are vanishing for $l = 0$, so that

$$\begin{aligned} \left| \sum_{(k,l) \neq 0,0} \frac{\mu^k \lambda^l}{k! l!} \int \phi_{k+l} \cdot [C_0^l - (M_\beta \varphi_0)^{\otimes l}] \right| &\leq \sum_{\substack{k \geq 0 \\ l \geq 1}} \frac{\mu^k \lambda^l}{k! l!} \int |\phi_{k+l}| \cdot 2C_0^l \\ &\leq 2 \sum_{r \geq 1} \sum_{l=1}^r \frac{\mu^{r-l} (\lambda C_0)^l}{(r-l)! l!} \int |\phi_r| \\ &\leq 2 \sum_{r \geq 1} \sum_{l=1}^r \frac{\mu^{r-l} (\lambda C_0)^l}{(r-l)! l!} r^{r-2} (|\mathcal{B}_d| \varepsilon^d)^{r-1}, \end{aligned}$$

by the point (ii) of Proposition A.1, stemming from the tree inequality. Using Stirling's approximation, we have for any $l \leq r$

$$\frac{r^{r-2}}{(r-l)! l!} \leq \frac{e^r}{r^2} \cdot \frac{r!}{(r-l)! l!} \leq \frac{(2e)^r}{r^2},$$

so that

$$\left| \sum_{(k,l) \neq 0,0} \frac{\mu^k \lambda^l}{k! l!} \int \phi_{k+l} \cdot [C_0^l - (M_\beta \varphi_0)^{\otimes l}] \right| \leq 4eC_0 \lambda \sum_{r \geq 1} \sum_{l=1}^r \frac{\mu^{r-l} (\lambda C_0)^{l-1}}{r^2} (2e|\mathcal{B}_d| \varepsilon^d)^{r-1}.$$

Eventually, thanks to the scaling $\mu \varepsilon^{d-1} = 1$ and $1 \ll \lambda \ll \mu$, we have

$$\begin{aligned} \sum_{r \geq 1} \sum_{l=1}^r \frac{\mu^{r-l} (\lambda C_0)^{l-1}}{r^2} (2e|\mathcal{B}_d| \varepsilon^d)^{r-1} &\leq \sum_{r \geq 1} \sum_{l=1}^r \frac{\mu^{r-1}}{r^2} (2e|\mathcal{B}_d| \varepsilon^d)^{r-1} \\ &\leq \sum_{r \geq 1} (2e|\mathcal{B}_d| \varepsilon)^{r-1} \\ &\leq 2, \end{aligned}$$

concluding the proof. □

References

- [1] R. K. Alexander. *The infinite hard-sphere system*. ProQuest LLC, Ann Arbor, MI, 1975. Thesis (Ph.D.)–University of California, Berkeley.
- [2] I. Ampatzoglou, J. K. Miller, and N. Pavlović. A rigorous derivation of a Boltzmann system for a mixture of hard-sphere gases. *SIAM J. Math. Anal.*, 54(2):2320–2372, 2022.
- [3] N. Ayi. From Newton’s law to the linear Boltzmann equation without cut-off. *Comm. Math. Phys.*, 350(3):1219–1274, 2017.
- [4] C. Bardos, F. Golse, and D. Levermore. Fluid dynamic limits of kinetic equations. I. Formal derivations. *J. Statist. Phys.*, 63(1-2):323–344, 1991.
- [5] T. Bodineau, I. Gallagher, and L. Saint-Raymond. The Brownian motion as the limit of a deterministic system of hard-spheres. *Invent. Math.*, 203(2):493–553, 2016.
- [6] T. Bodineau, I. Gallagher, L. Saint-Raymond, and S. Simonella. Statistical dynamics of a hard sphere gas: fluctuating Boltzmann equation and large deviations. *Ann. of Math. (2)*, 198(3):1047–1201, 2023.
- [7] L. Boltzmann. *Lectures on gas theory*. University of California Press, Berkeley-Los Angeles, Calif., 1964. Translated by Stephen G. Brush.
- [8] C. Cercignani, R. Illner, and M. Pulvirenti. *The mathematical theory of dilute gases*, volume 106 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1994.
- [9] Y. Deng, Z. Hani, and X. Ma. Hilbert’s sixth problem: derivation of fluid equations via Boltzmann’s kinetic theory, 2025. Preprint.
- [10] L. Desvillettes and M. Pulvirenti. The linear Boltzmann equation for long-range forces: a derivation from particle systems. *Math. Models Methods Appl. Sci.*, 9(8):1123–1145, 1999.
- [11] L. Erdős and D. Q. Tuyen. Central limit theorems for the one-dimensional Rayleigh gas with semipermeable barriers. *Comm. Math. Phys.*, 143(3):451–466, 1992.
- [12] F. Fougères. On the derivation of the linear Boltzmann equation from the nonideal Rayleigh gas. *Journal of Statistical Physics*, 191, 10 2024.
- [13] F. Fougères. Cumulants of the Rayleigh gas mixture model: statistical results, 2026. Preprint.
- [14] F. Fougères. Refining the kinetic limit of the nonideal Rayleigh gas, 2026. PhD thesis.
- [15] I. Gallagher, L. Saint-Raymond, and B. Texier. *From Newton to Boltzmann: hard spheres and short-range potentials*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2013.
- [16] T. Gianoli, J.-F. Boussuge, P. Sagaut, and J. de Laborderie. Development and validation of Navier-Stokes characteristic boundary conditions applied to turbomachinery simulations using the lattice Boltzmann method. *Internat. J. Numer. Methods Fluids*, 95(4):528–556, 2023.
- [17] F. Golse. On the periodic Lorentz gas and the Lorentz kinetic equation. *Ann. Fac. Sci. Toulouse Math. (6)*, 17(4):735–749, 2008.
- [18] F. Golse and L. Saint-Raymond. The Navier-Stokes limit for the Boltzmann equation. *C. R. Acad. Sci. Paris Sér. I Math.*, 333(9):897–902, 2001.

- [19] F. G. King. *BBGKY hierarchy for positive potentials*. ProQuest LLC, Ann Arbor, MI, 1975. Thesis (Ph.D.)—University of California, Berkeley.
- [20] O. E. Lanford, III. Time evolution of large classical systems. In *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*, volume Vol. 38 of *Lecture Notes in Phys.*, pages 1–111. Springer, Berlin-New York, 1975.
- [21] O. E. Lanford, III. On a derivation of the Boltzmann equation. In *International Conference on Dynamical Systems in Mathematical Physics (Rennes, 1975)*, volume No. 40 of *Astérisque*, pages 117–137. Soc. Math. France, Paris, 1976.
- [22] J. L. Lebowitz and H. Spohn. Steady state self-diffusion at low density. *J. Statist. Phys.*, 29(1):39–55, 1982.
- [23] K. Matthies, G. Stone, and F. Theil. The derivation of the linear Boltzmann equation from a Rayleigh gas particle model. *Kinet. Relat. Models*, 11(1):137–177, 2018.
- [24] A. Nota, R. Winter, and B. Lods. Kinetic description of a Rayleigh gas with annihilation. *J. Stat. Phys.*, 176(6):1434–1462, 2019.
- [25] O. Penrose. Convergence of fugacity expansions for classical systems. *Statistical mechanics: foundations and applications*, page 101, 1967.
- [26] L. Saint-Raymond. From Boltzmann’s kinetic theory to Euler’s equations. *Phys. D*, 237(14–17):2028–2036, 2008.
- [27] H. Spohn. *Kinetic Equations from Hamiltonian Dynamics: The Markovian Approximations*, pages 183–211. Springer Vienna, Vienna, 1988.
- [28] H. Spohn. *Large Scale Dynamics of Interacting Particles*, volume vol. 174 of *Theoretical and Mathematical Physics*. Springer, 1991.
- [29] H. van Beijeren, O. E. Lanford, III, J. L. Lebowitz, and H. Spohn. Equilibrium time correlation functions in the low-density limit. *J. Statist. Phys.*, 22(2):237–257, 1980.
- [30] C. Villani. A review of mathematical topics in collisional kinetic theory. In *Handbook of mathematical fluid dynamics, Vol. I*, pages 71–305. North-Holland, Amsterdam, 2002.
- [31] G. Wissocq and P. Sagaut. Hydrodynamic limits and numerical errors of isothermal lattice Boltzmann schemes. *J. Comput. Phys.*, 450:Paper No. 110858, 61, 2022.