



# Fluctuation Theory in the Boltzmann–Grad Limit

Thierry Bodineau<sup>1</sup> · Isabelle Gallagher<sup>2</sup> · Laure Saint-Raymond<sup>3</sup> · Sergio Simonella<sup>4</sup>

Received: 30 March 2020 / Accepted: 7 April 2020 / Published online: 3 May 2020  
© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

We develop a rigorous theory of hard-sphere dynamics in the kinetic regime, away from thermal equilibrium. In the low density limit, the empirical density obeys a law of large numbers and the dynamics is governed by the Boltzmann equation. Deviations from this behaviour are described by dynamical correlations, which can be fully characterized for short times. This provides both a fluctuating Boltzmann equation and large deviation asymptotics.

**Keywords** Kinetic theory · Hard sphere dynamics · Low density limit · Boltzmann equation · Fluctuations · Large deviations · cumulants

## 1 Introduction

In this note we report on some recent progress on the origins of the fluctuation theory from the fundamental laws of motion. For states far from equilibrium, the macroscopic fluctuation theory has been investigated intensively, but microscopic derivations are mainly focused on stochastic lattice gases (see e.g. [3,12,32]). We study here classical deterministic particles in a rarefied gas. In its rigorous version, the issue is then connected with the problem of the

---

This paper is dedicated to Joel Lebowitz for his 90th birthday.

---

Communicated by Eric A. Carlen.

---

✉ Laure Saint-Raymond  
laure.saint-raymond@ens-lyon.fr

Thierry Bodineau  
thierry.bodineau@polytechnique.edu

Isabelle Gallagher  
isabelle.gallagher@ens.fr

Sergio Simonella  
sergio.simonella@ens-lyon.fr

<sup>1</sup> CMAP, CNRS, Ecole Polytechnique, I.P. Paris, Palaiseau, France

<sup>2</sup> DMA, Ecole Normale Supérieure, Paris, France

<sup>3</sup> UMPA, Ecole Normale Supérieure de Lyon, Lyon, France

<sup>4</sup> UMPA, CNRS, Ecole Normale Supérieure de Lyon, Lyon, France

mathematical validity of the Boltzmann equation, in the limit introduced by Grad [16]. This limit procedure states that in a Hamiltonian system of  $N$  particles, strongly interacting at distance  $\varepsilon$ , the particle density approximates the solution to the Boltzmann equation when  $N \rightarrow \infty$ ,  $\varepsilon \rightarrow 0$ , in such a way that the collision frequency (proportional to  $N\varepsilon^{d-1}$  in dimension  $d = 2, 3$ ) remains bounded; the volume density scales like  $\varepsilon$ , and both collisions and transport have a finite effect in the limit.

The Boltzmann gas is a simple case featuring a nonlinear dynamics, and a rich structure for the fluctuations. At the macroscopic scale, typical particles behave as i.i.d. variables. Small fluctuations admit already an interesting theory. In particular, they exhibit spatial correlations, and noise originating from (deterministic) collisions [13,31]. Moreover, rare fluctuations satisfy a large deviation principle. We refer to the companion work by Bouchet [7], where the large deviation theory for the Boltzmann equation has been discussed first.

On the mathematical side, the only result we are aware of in this direction, is the convergence of the second moment of the small fluctuations proved by Spohn in [30]. Similar results are available for linear regimes close to equilibrium, both for short [2] and for large times [4]. As suggested in [31], the fluctuation theory should not be merely a phenomenological theory, but a rigorous consequence of the laws of mechanics. Our aim is to support this assertion, providing a robust mathematical framework.

We shall state several theorems (Theorems 1, 3, 4) describing the behaviour of the empirical density

$$\pi_t^\varepsilon := \frac{1}{\mu_\varepsilon} \sum_{i=1}^N \delta_{z_i^\varepsilon(t)}, \quad \mu_\varepsilon = \varepsilon^{-(d-1)} \tag{1.1}$$

for a Newtonian evolution of  $N$  particles with positions and velocities

$$z_i^\varepsilon(t) = (x_i^\varepsilon(t), v_i^\varepsilon(t)), \quad i = 1, \dots, N.$$

We assume that the particles are approximately Poisson-distributed at time  $t = 0$ , with (random) total number of particles  $\mathcal{N}$  and regular phase-space density  $f^0 = f^0(x, v)$ . Probability and expectation with respect to this initial measure are denoted by  $\mathbb{P}_\varepsilon$  and  $\mathbb{E}_\varepsilon$ . Then, in the Boltzmann–Grad limit  $\varepsilon \rightarrow 0$ ,  $\mathbb{E}_\varepsilon(\mathcal{N})/\mu_\varepsilon \rightarrow 1$ , the following properties hold.

1. *Law of large numbers:*

$$\pi_t^\varepsilon \rightarrow f_t, \quad t \in [0, T^*] \tag{1.2}$$

weakly (in probability) for some  $T^* > 0$ , where  $f_t$  is the solution of

$$\partial_t f + v \cdot \nabla_x f = C(f, f) \tag{1.3}$$

with initial datum  $f^0$  and  $C$  is Boltzmann’s collision operator [21]; in particular, the *chaos* property of the initial measure propagates in time (rescaled correlation functions converge to a tensor product).

2. *Central limit theorem:* the fluctuation field

$$\zeta_t^\varepsilon := \sqrt{\mu_\varepsilon} (\pi_t^\varepsilon - \mathbb{E}_\varepsilon(\pi_t^\varepsilon)) \tag{1.4}$$

describing the small deviations of the empirical density from its average, converges in law on  $[0, T^*]$  to the Gaussian process  $\zeta_t$  governed by the fluctuating Boltzmann equation

$$d\zeta_t = \mathcal{L}_t \zeta_t dt + d\eta_t, \tag{1.5}$$

where  $\mathcal{L}_t$  is Boltzmann’s operator linearized around  $f_t$ , and  $d\eta_t$  is Gaussian noise (with covariance defined in (5.2)), as predicted in [30].

3. *Large deviations* are exponentially small in  $\mu_\varepsilon$  and characterized, at least in a regime of strong regularity, by the same large deviation functional as heuristically derived in [7] (and previously obtained, rigorously, in [28] from a one-dimensional stochastic process). That is, the probability of observing a path  $\varphi_t = \varphi(t, x, v)$  satisfies

$$\mathbb{P}_\varepsilon (\pi_t^\varepsilon \approx \varphi_t, t \in [0, T^*]) \asymp \exp(-\mu_\varepsilon \mathcal{F}(T^*, \varphi)), \tag{1.6}$$

where  $\mathcal{F}$  is defined as the Legendre transform of a functional  $\mathcal{J} = \mathcal{J}(T^*, \varphi)$ , solution of a Hamilton–Jacobi equation.

The Boltzmann equation is naturally suited to a probabilistic interpretation, and its mathematical validity can be based on the construction of a stochastic particle system mimicking the microscopic collisions. The basic example is the Kac model [17], from which the spatially homogeneous Boltzmann equation can indeed be recovered. Fluctuations in this type of process can also be analysed (see e.g. [18,19,23]), including large deviations ([22]) and spatially inhomogeneous variants ([27,28]). Our results show that the analogy between the deterministic hard-sphere dynamics and the stochastic model goes far beyond the typical behavior as it remains valid for extremely rare events. The statistical behavior of the hard-sphere gas described above is in fact the same as the one derived in [27,28]. For physical observables as the empirical measure, the deterministic dynamics and the stochastic approximation (as often used in simulations) cannot be distinguished, even at the level of fluctuations.

Our main restriction is the smallness of the time  $T^*$ . This time (depending only on  $f^0$ ) is actually a fraction of the time of validity of the Boltzmann equation obtained by Lanford in [21]. We will further restrict to a gas of hard spheres, though we believe that the results could be proved for smooth and compactly supported interactions, adopting known techniques [14, 20,25].

The Hamilton–Jacobi equation determining  $\mathcal{J}$  (Theorem 2) is our ultimate point of arrival in the derivation from a microscopic mechanical model. A stationary solution of this equation is given by the (dual) Boltzmann’s  $H$  functional, which describes large deviations of the equilibrium state. Moreover,  $\mathcal{F}$  has an invariance encoding the microscopic reversibility, a symmetry inherited from the equality between the probability of a path and the probability of the time-reversed path. This is an indication on the amount of recovered information which was “lost” in Lanford’s Theorem, proving the transition from a reversible to a dissipative model.

Our method is far from standard approaches to a large deviation problem. For stochastic dynamics, large deviations can be evaluated by modifying the underlying stochastic process in a time dependent way in order to produce an atypical trajectory. The optimal cost for inducing such a bias on the stochastic dynamics is precisely the large deviation rate. For hard-sphere systems, there is no underlying stochastic dynamics as all the randomness lies in the initial data. It seems exceedingly hard to figure out a way to bias the initial probability measure in order to produce a given admissible path  $\varphi_t$ . Indeed, the deterministic dynamics is responsible for an intricate relation between the path and the initial distribution of spheres.

We therefore turn back to the more modest problem of analysing the error in (1.2).<sup>1</sup> It is already evident in Lanford’s proof, that the dynamical information lives on precise little regions of the  $j$ –particle phase space, converging to measure-zero sets as  $\varepsilon \rightarrow 0$ , for any finite  $j$ . In little regions of the same size, correlations are generated by the collision events, which break the propagation of chaos (see e.g. [5]). These correlation sets do not encode the most probable future dynamics and they can be neglected when proving (1.2). However we

<sup>1</sup> For previous quantitative investigations of the correlation error, we refer to [14,26].

can extract much more information, by looking for mathematically tractable quantities which are concentrated exactly on these sets, and retaining the information which is lost in (1.2).

A natural candidate is provided by *cumulants*, which can be obtained by the series expansion of the generating function

$$\Lambda_t^\varepsilon(h) := \frac{1}{\mu_\varepsilon} \log \mathbb{E}_\varepsilon \left( \exp \left( \sum_{i=1}^{\mathcal{N}} h(\mathbf{z}_i^\varepsilon(t)) \right) \right) \tag{1.7}$$

where  $h$  is any test function. The order  $n$  in this expansion is given by a function  $f_n^\varepsilon = f_n^\varepsilon(t)$  on the  $n$ -particle phase space (formula (4.6)) describing a cluster of particles mutually correlated by a chain of interactions. It has been noted in [13] that the hierarchy of cumulants determines *all* the properties of the fluctuations in a gas, and that their exact computation furnishes a theory of fluctuations at the same time. In order to prove rigorous results and reach large deviations, we will construct the limit of the exponential moment (1.7), and link it to the function  $\mathcal{J}$ .

The expansion of (1.7) leads to a combinatorial problem, which can be dealt with by the cluster expansion method [29]. Indeed this method fits very well with the dynamics at low density, when combined with geometrical estimates on hard-sphere trajectories.

We organize the paper as follows. Section 2 is a brief introduction to our strategy. Section 3 presents the model and the fundamental result leading to (1.2), and explains the basic dynamical formula expressing the main quantities of interest in terms of the initial data. In Sect. 4 we state our main results on the dynamical correlations and their limiting structure, and derive the Hamilton–Jacobi equation. Finally, the last two sections are devoted to the fluctuating Boltzmann equation and the large deviations respectively. In this paper we shall only sketch the proof of our results, the complete version of which will be provided in a longer publication [6].

## 2 Strategy

Lanford’s method [21] is based on the BBGKY hierarchy governing the evolution of the family of (properly rescaled) correlation functions  $(F_n^\varepsilon)_{n \geq 1}$ . This hierarchy is completely equivalent to the Liouville equation describing interacting transport of  $N$  hard spheres. In the Boltzmann–Grad limit, the probability densities concentrate on an infinite-dimensional space, and the BBGKY hierarchy is convenient to capture the relevant information. One thus introduces the (rescaled) correlation functions  $F_n^\varepsilon(t, Z_n)$  such that

$$\mathbb{E}_\varepsilon \left( \sum_{\substack{i_1, \dots, i_n \\ i_j \neq i_k, j \neq k}} h_n(\mathbf{z}_{i_1}^\varepsilon(t), \dots, \mathbf{z}_{i_n}^\varepsilon(t)) \right) = \mu_\varepsilon^n \int_{\mathbb{D}^n} dZ_n F_n^\varepsilon(t, Z_n) h_n(Z_n),$$

for any test function  $h_n$ , where  $\mathbb{D}$  is the 1-particle phase space. The family  $(F_n^\varepsilon)_{n \geq 1}$  is suited to the description of typical events: in the limit  $\varepsilon \rightarrow 0$ ,  $F_n^\varepsilon \rightarrow f^{\otimes n}$  so that everything is coded in  $f$  (solution of (1.3)), no matter how large  $n$ .

We need to go beyond the BBGKY hierarchy and turn to a more powerful representation of the dynamics. We shall replace the family  $(F_n^\varepsilon)_{n \geq 1}$  with an equivalent family of (rescaled) truncated correlation functions  $(f_n^\varepsilon)_{n \geq 1}$ , called *cumulants*. Their role is to grasp information on the dynamics on finer and finer scales. Loosely speaking,  $f_n^\varepsilon$  will collect events where  $n$  particles are “completely connected” by a chain of interactions. We shall say that the  $n$

particles form a connected *cluster*. Since a collision between two given particles is typically of order  $\mu_\varepsilon^{-1}$  (the size of the “collision tube” spanned by one particle in time 1), a complete connection would account for events of probability of order  $\mu_\varepsilon^{-(n-1)}$ . We therefore end up with a hierarchy of rare events, which we would like to control at arbitrary order. At variance with  $(F_n^\varepsilon)_{n \geq 1}$ , even *after* the limit  $\mu_\varepsilon \rightarrow \infty$  is taken, the cumulant  $f_n^\varepsilon$  cannot be trivially obtained from the cumulant  $f_{n-1}^\varepsilon$ . Each step entails extra information, and events of increasing complexity, and decreasing probability.

Unfortunately, the equations for  $(f_n^\varepsilon)_{n \geq 1}$  are difficult to handle. But the moment-to-cumulant relation  $(F_n^\varepsilon)_{n \geq 1} \rightarrow (f_n^\varepsilon)_{n \geq 1}$  is a bijection and, in order to construct  $f_n^\varepsilon(t)$ , we can still resort to the same solution representation of [21] for the correlation functions  $(F_n^\varepsilon(t))_{n \geq 1}$ . This formula is an expansion over *collision trees*, meaning that it has a geometrical representation as a sum over binary tree graphs, with vertices accounting for collisions (see Sect. 3). Two particles are correlated if their generated trees are *connected* by a “recollision”, which is an event of weight  $\mu_\varepsilon^{-1}$  (see Sect. 3.3.3 for a precise notion of recollision).

In Proposition 2 we will state the main technical advance of this paper: the cumulant (rescaled by the factor  $\mu_\varepsilon^{n-1}$ ) grows as  $n^{n-2}$  in  $L^1$ -norm. This estimate is intuitively simple. We have at disposal a geometric notion of correlation as a *link* between two collision trees. Based on this notion, we can draw a random graph on  $n$  vertices telling us which particles are correlated and which particles are not (each collision tree being one vertex of the graph). Since the cumulant  $f_n^\varepsilon$  corresponds to  $n$  completely correlated particles, there will be at least  $n - 1$  edges, each one of small ‘volume’  $\mu_\varepsilon^{-1}$ . Of course there could be more than  $n - 1$  connections (the random graph has cycles), but these are hopefully unlikely as they produce extra smallness in  $\varepsilon$ . If we ignore all of them, we are left with minimally connected graphs, whose total number is  $n^{n-2}$  by Cayley’s formula.

The limiting equations for the family  $(f_n^\varepsilon)_{n \geq 1}$  form a *Boltzmann cumulant hierarchy*, displaying a remarkable structure [10,13]. The first equation ( $n = 1$ ) is just the Boltzmann equation. The second equation ( $n = 2$ ) is driven by a linearized Boltzmann operator  $\mathcal{L}_1$ , plus a singular “recollision operator”, acting on  $f_1$  only, generating the “connection” (correlation) between two particles and suited to be interpreted as noise source [30]. The higher order equations ( $n > 2$ ) have an increasingly complex structure, combining the action of the two operators (of standard linearized type, and of connecting type) on  $n$  different particles, in all possible ways. But the good  $n$ -dependence of the uniform bounds allows to sum up the cumulants into an analytic series. This finally translates the cumulant hierarchy into the Hamilton–Jacobi equation, which stands as a compact, nonlinear representation of the correlation dynamics.

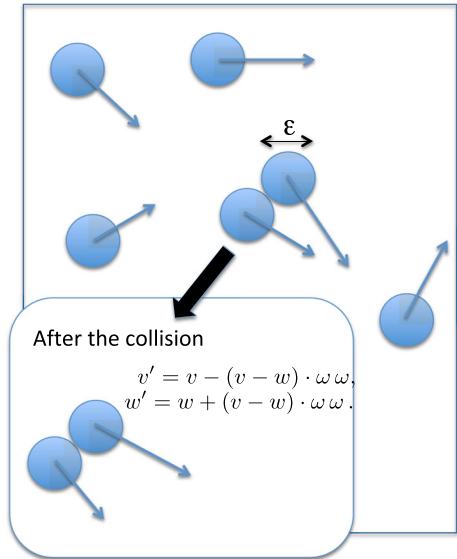
### 3 Collision Trees

In this section we introduce the geometrical representation of the hard-sphere dynamics with random initial data, which will be our basic tool.

#### 3.1 Hard-Sphere Model

The microscopic model consists of  $N$  identical hard spheres of unit mass and of diameter  $\varepsilon$ . Their motion is governed by a system of ordinary differential equations, which are set in  $\mathbb{D}^N := (\mathbb{T}^d \times \mathbb{R}^d)^N$  where  $\mathbb{T}^d$  is the unit  $d$ -dimensional periodic box:

**Fig. 1** Transport and collisions in a hard-sphere gas



$$\frac{d\mathbf{x}_i^\varepsilon}{dt} = \mathbf{v}_i^\varepsilon, \quad \frac{d\mathbf{v}_i^\varepsilon}{dt} = 0 \quad \text{as long as } |\mathbf{x}_i^\varepsilon(t) - \mathbf{x}_j^\varepsilon(t)| > \varepsilon \quad \text{for } 1 \leq i \neq j \leq N, \quad (3.1)$$

with specular reflection at collisions:

$$\left. \begin{aligned} (\mathbf{v}_i^\varepsilon)' &:= \mathbf{v}_i^\varepsilon - \frac{1}{\varepsilon^2} (\mathbf{v}_i^\varepsilon - \mathbf{v}_j^\varepsilon) \cdot (\mathbf{x}_i^\varepsilon - \mathbf{x}_j^\varepsilon) (\mathbf{x}_i^\varepsilon - \mathbf{x}_j^\varepsilon) \\ (\mathbf{v}_j^\varepsilon)' &:= \mathbf{v}_j^\varepsilon + \frac{1}{\varepsilon^2} (\mathbf{v}_i^\varepsilon - \mathbf{v}_j^\varepsilon) \cdot (\mathbf{x}_i^\varepsilon - \mathbf{x}_j^\varepsilon) (\mathbf{x}_i^\varepsilon - \mathbf{x}_j^\varepsilon) \end{aligned} \right\} \quad \text{if } |\mathbf{x}_i^\varepsilon(t) - \mathbf{x}_j^\varepsilon(t)| = \varepsilon. \quad (3.2)$$

The sign of the scalar product  $(\mathbf{v}_i^\varepsilon - \mathbf{v}_j^\varepsilon) \cdot (\mathbf{x}_i^\varepsilon - \mathbf{x}_j^\varepsilon)$  identifies post-collisional (+) and pre-collisional (−) configurations. This flow does not cover all possible situations, as multiple collisions are excluded. But one can show (see [1]) that, for almost every initial configuration  $(\mathbf{x}_i^{\varepsilon 0}, \mathbf{v}_i^{\varepsilon 0})_{1 \leq i \leq N}$ , there are neither multiple collisions, nor accumulations of collision times, so that the dynamics is globally well defined (Fig. 1).

Below, we shall denote collections of positions and velocities respectively by  $X_N := (x_1, \dots, x_N) \in \mathbb{T}^{dN}$  and  $V_N := (v_1, \dots, v_N) \in \mathbb{R}^{dN}$ , and we set  $Z_N := (X_N, V_N) \in (\mathbb{T}^d \times \mathbb{R}^d)^N$ ,  $Z_N = (z_1, \dots, z_N)$ .

Let  $f^0$  be a probability density on  $\mathbb{D}$  with Gaussian decay in velocity

$$|f^0(x, v)| + |\nabla_x f^0(x, v)| \leq C_0 \exp\left(-\frac{\beta_0}{2} |v|^2\right), \quad (3.3)$$

where  $C_0, \beta_0 > 0$ . Because of the condition of hard-sphere exclusion, the positions of the particles cannot be independent of each other. To better focus on the dynamical issue, we shall choose, as initial measure, the  $N$ -particle distribution with minimal correlations. In particular, to avoid spurious correlations due to a given total number of particles, we shall consider a grand canonical state. The initial probability density of finding  $N$  particles in  $Z_N$  is given by

$$\frac{1}{N!} W_N^{\varepsilon 0}(Z_N) := \frac{1}{Z^\varepsilon} \frac{\mu_\varepsilon^N}{N!} \mathbf{1}_{\mathcal{D}_N^\varepsilon} \prod_{i=1}^N f^0(z_i) \quad (3.4)$$

where the domain encodes the exclusion:

$$\mathcal{D}_N^\varepsilon := \{Z_N \in \mathbb{D}^N \mid \forall i \neq j, \quad |x_i - x_j| > \varepsilon\},$$

and the normalization constant  $\mathcal{Z}^\varepsilon$  is given by

$$\mathcal{Z}^\varepsilon := 1 + \sum_{N \geq 1} \frac{\mu_\varepsilon^N}{N!} \int_{\mathbb{D}^N} dZ_N \mathbf{1}_{\mathcal{D}_N^\varepsilon} \prod_{i=1}^N f^0(z_i).$$

With this definition, if

$$\mu_\varepsilon \varepsilon^{d-1} = 1,$$

then the average number of particles satisfies

$$\lim_{\varepsilon \rightarrow 0} \mathbb{E}_\varepsilon(\mathcal{N}) \varepsilon^{d-1} = 1$$

(Boltzmann–Grad scaling).

The rescaled  $n$ -particle correlation function is defined by

$$F_n^{\varepsilon 0}(Z_n) := \mu_\varepsilon^{-n} \sum_{p=0}^\infty \frac{1}{p!} \int dz_{n+1} \dots dz_{n+p} W_{n+p}^{\varepsilon 0}(Z_{n+p}). \tag{3.5}$$

For any symmetric test function  $h_n : \mathbb{D}^n \rightarrow \mathbb{R}$ , one can check that

$$\mathbb{E}_\varepsilon \left( \sum_{\substack{i_1, \dots, i_n \\ i_j \neq i_k, j \neq k}} h_n(\mathbf{z}_{i_1}^{\varepsilon 0}, \dots, \mathbf{z}_{i_n}^{\varepsilon 0}) \right) = \mu_\varepsilon^n \int_{\mathbb{D}^n} dZ_n F_n^{\varepsilon 0}(Z_n) h_n(Z_n). \tag{3.6}$$

Moreover one can prove that, in the Boltzmann–Grad limit,

$$\forall n \geq 1, \quad F_n^{\varepsilon 0}(Z_n) \longrightarrow \prod_{i=1}^n f^0(z_i) \text{ as } \varepsilon \rightarrow 0$$

on the set  $\{x_i \neq x_j, \forall i \neq j\}$ . That is, at leading order, the initial distribution is chaotic.

Starting from the dynamical equations (3.1) we get that, for each fixed  $N$ , the probability density at time  $t > 0$  is determined by the Liouville equation

$$\partial_t W_N^\varepsilon + V_N \cdot \nabla_{X_N} W_N^\varepsilon = 0 \quad \text{on } \mathcal{D}_N^\varepsilon, \tag{3.7}$$

with specular reflection (3.2) on the boundary  $|x_i - x_j| = \varepsilon$ .

By integration of the Liouville equation for fixed  $\varepsilon$ , we get that the one-particle correlation function  $F_1^\varepsilon$  satisfies an equation

$$\partial_t F_1^\varepsilon + v \cdot \nabla_x F_1^\varepsilon = C_{1,2}^\varepsilon F_2^\varepsilon \tag{3.8}$$

where the collision operator comes from the boundary terms in Green’s formula (using the reflection condition to rewrite the gain part in terms of pre-collisional velocities):

$$\begin{aligned} (C_{1,2}^\varepsilon F_2^\varepsilon)(x, v) &:= \int F_2^\varepsilon(x, v', x + \varepsilon \omega, w') ((w - v) \cdot \omega)_+ d\omega dw \\ &\quad - \int F_2^\varepsilon(x, v, x + \varepsilon \omega, w) ((w - v) \cdot \omega)_- d\omega dw, \end{aligned}$$

with

$$v' = v - (v - w) \cdot \omega \omega, \quad w' = w + (v - w) \cdot \omega \omega.$$

As in (3.6),  $F_1^\varepsilon(t)$  describes the average behavior of (identical) particles at time  $t$ :

$$\mathbb{E}_\varepsilon \left( \frac{1}{\mu_\varepsilon} \sum_{i=1}^{\mathcal{N}} h(\mathbf{z}_i^\varepsilon(t)) \right) = \int F_1^\varepsilon(t, z) h(z) dz,$$

for any test function  $h : \mathbb{D} \rightarrow \mathbb{R}$ . Similarly for any test function  $h_2 : \mathbb{D}^2 \rightarrow \mathbb{R}$ , the two-particle correlation function satisfies

$$\mathbb{E}_\varepsilon \left( \frac{1}{\mu_\varepsilon^2} \sum_{i \neq j} h_2(\mathbf{z}_i(t), \mathbf{z}_j(t)) \right) = \int F_2^\varepsilon(t, Z_2) h_2(Z_2) dZ_2.$$

### 3.2 Law of Large Numbers

The issue with equation (3.8) is that it is not closed: it involves  $F_2^\varepsilon$ . At the level of (3.8), Boltzmann’s main assumption would correspond to the replacement

$$F_2^\varepsilon(t, x_1, v_1, x_2, v_2) \sim F_1^\varepsilon(t, x_1, v_1) F_1^\varepsilon(t, x_2, v_2), \quad \text{as } \varepsilon \rightarrow 0,$$

when  $|x_1 - x_2| = \varepsilon, \quad (x_1 - x_2) \cdot (v_1 - v_2) < 0.$

In other words, particles are assumed to be statistically independent, at least in pre-collisional configurations. This very *strong chaos property* (which we assumed at time 0) is supposed to be valid for all times.

In the Boltzmann–Grad limit  $\varepsilon \rightarrow 0$ , we then expect  $F_1^\varepsilon$  to be well approximated by the solution to the *Boltzmann equation*

$$\partial_t f + v \cdot \nabla_x f = C(f, f) \tag{3.9}$$

with

$$C(f, f)(t, x, v) := \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} \left( f(t, x, w') f(t, x, v') - f(t, x, w) f(t, x, v) \right) ((v - w) \cdot \omega)_+ d\omega dw.$$

The claim by Boltzmann that the particle density is well approximated by Eq. (3.9) has been made rigorous by Lanford for short times.

**Theorem 1** (Lanford, [21]) *Consider a gas of hard spheres initially distributed according to (3.4). Then, in the Boltzmann–Grad limit  $\mu_\varepsilon \rightarrow \infty$  with  $\mu_\varepsilon \varepsilon^{d-1} = 1$ , the 1-particle distribution  $F_1^\varepsilon$  converges, uniformly on compact sets, towards the solution  $f$  of the Boltzmann equation (3.9) on a short time interval  $[0, T^*]$  (where  $T^*$  depends on the initial distribution  $f^0$  through  $C_0, \beta_0$  in (3.3)).*

*Furthermore for each  $n$ , the  $n$ -particle correlation function  $F_n^\varepsilon(t)$  converges almost everywhere to  $f^{\otimes n}(t)$  on the same time interval.*

The propagation of chaos obtained in Lanford’s theorem implies in particular that the empirical measure  $\pi_t^\varepsilon$ , defined by (1.1), concentrates on the solution to the Boltzmann equa-



tion. Indeed computing the variance we get, for any test function  $h$ , that

$$\begin{aligned} & \mathbb{E}_\varepsilon \left( \left( \pi_t^\varepsilon(h) - \int F_1^\varepsilon(t, z) h(z) dz \right)^2 \right) \\ &= \mathbb{E}_\varepsilon \left( \frac{1}{\mu_\varepsilon^2} \sum_{i=1}^{\mathcal{N}} h^2(\mathbf{z}_i^\varepsilon(t)) + \frac{1}{\mu_\varepsilon^2} \sum_{i \neq j} h(\mathbf{z}_i^\varepsilon(t)) h(\mathbf{z}_j^\varepsilon(t)) \right) - \left( \int F_1^\varepsilon(t, z) h(z) dz \right)^2 \\ &= \frac{1}{\mu_\varepsilon} \int F_1^\varepsilon h^2 dz_1 + \int F_2^\varepsilon h^{\otimes 2} dZ_2 - \left( \int F_1^\varepsilon h dz \right)^2 \end{aligned} \tag{3.10}$$

which converges to 0 as  $\varepsilon \rightarrow 0$  since  $F_2^\varepsilon$  converges to  $f^{\otimes 2}$  and  $F_1^\varepsilon$  to  $f$ . This computation can be interpreted as a law of large numbers.

Theorem 1 entails a drastic loss of information, which (as becomes clear from the proof) is retained in particular in “recollision sets” of measure zero. Some of the microscopic time-reversible structure can be recovered by looking at correlations on finer scales. This is the role played by the rescaled dynamical *cumulants*, defined by

$$f_n^\varepsilon(t, Z_n) := \mu_\varepsilon^{n-1} \sum_{s=1}^n \sum_{\sigma \in \mathcal{P}_n^s} (-1)^{s-1} (s-1)! \prod_{i=1}^s F_{|\sigma_i|}^\varepsilon(t, Z_{\sigma_i}). \tag{3.11}$$

Here we denoted by  $\mathcal{P}_n^s$  the set of partitions of  $\{1, \dots, n\}$  in  $s$  parts,  $\mathcal{P}_n^s \ni \sigma = (\sigma_1, \dots, \sigma_s)$ , by  $|\sigma_i|$  the cardinality of the set  $\sigma_i$  and by  $Z_{\sigma_i} = (z_j)_{j \in \sigma_i}$ . This formula is cooked up to extract the effect of recollisions and therefore to obtain the detailed correlation structure at arbitrarily small scales. Note that, for fixed  $\varepsilon > 0$ ,  $(F_n^\varepsilon)_{n \geq 1}$  and  $(f_n^\varepsilon)_{n \geq 1}$  provide the same amount of information, as shown by the inversion formula :

$$F_n^\varepsilon(t, Z_n) = \sum_{s=1}^n \sum_{\sigma \in \mathcal{P}_n^s} \mu_\varepsilon^{-(n-s)} \prod_{i=1}^s f_{|\sigma_i|}^\varepsilon(t, Z_{\sigma_i}). \tag{3.12}$$

Before passing to the investigation of (3.11), we need to recall the main features of the proof of Theorem 1 (see [9,14,32] for more details).

### 3.3 Hierarchy and Pseudotrajectories

The starting point is the equation (3.8) for the 1-particle correlation function  $F_1^\varepsilon$ . In order to get a closed system, we write similar equations for all correlation functions  $F_n^\varepsilon$

$$\partial_t F_n^\varepsilon + V_n \cdot \nabla_{X_n} F_n^\varepsilon = C_{n,n+1}^\varepsilon F_{n+1}^\varepsilon \quad \text{on } \mathcal{D}_n^\varepsilon, \tag{3.13}$$

with specular boundary reflection as in (3.7) [8]. As  $C_{1,2}^\varepsilon$  above,  $C_{n,n+1}^\varepsilon$  describes collisions between one “fresh” particle (labelled  $n + 1$ ) and one given particle  $i \in \{1, \dots, n\}$ .

We denote by  $S_n^\varepsilon$  the group associated with free transport in  $\mathcal{D}_n^\varepsilon$  (with specular reflection at collisions). Iterating Duhamel’s formula, we can express the solution as a sum of operators acting on the initial data :

$$F_n^\varepsilon(t) = \sum_{m \geq 0} Q_{n,n+m}^\varepsilon(t) F_{n+m}^{\varepsilon 0}, \tag{3.14}$$

where we have defined for  $t > 0$

$$\begin{aligned} Q_{n,n+m}^\varepsilon(t) F_{n+m}^{\varepsilon 0} := & \int_0^t \int_0^{t_1} \dots \int_0^{t_{m-1}} S_n^\varepsilon(t - t_1) C_{n,n+1}^\varepsilon S_{n+1}^\varepsilon(t_1 - t_2) C_{n+1,n+2}^\varepsilon \\ & \dots S_{n+m}^\varepsilon(t_m) F_{n+m}^{\varepsilon 0} dt_m \dots dt_1 \end{aligned}$$

and  $Q_{n,n}^\varepsilon(t)F_n^{\varepsilon 0} := S_n^\varepsilon(t)F_n^0$ ,  $Q_{n,n+m}^\varepsilon(0)F_{n+m}^{\varepsilon 0} := \delta_{m,0}F_{n+m}^{\varepsilon 0}$ .

In the following, we shall label  $1^*, \dots, n^*$  the  $n$  particles with configuration  $Z_n$  at time  $t$ , and  $1, \dots, m$  the  $m$  “fresh” particles which are added by the collision operators. The configuration of the particle labeled  $i^*$  will be denoted indifferently  $z_i^* = (x_i^*, v_i^*)$  or  $z_{i^*} = (x_{i^*}, v_{i^*})$ .

### 3.3.1 The Tree Structure

Each term of the series expansion (3.14) (after inserting the explicit definition of the collision operators) can be represented by a collision tree  $a = (a_i)_{i=1, \dots, m}$ , which records the combinatorics of collisions : the colliding particles at time  $t_i$  are  $i$  and  $a_i \in \{1^*, \dots, n^*\} \cup \{1, \dots, i-1\}$ . We define the set  $\mathcal{A}_{n,m}$  of all possible such trees. Note that  $|\mathcal{A}_{n,m}| = n(n+1) \dots (n+m-1)$ . Note also that, graphically,  $a \in \mathcal{A}_{n,m}$  is represented by  $n$  binary tree graphs (below, we will call collision tree both  $a \in \mathcal{A}_{n,m}$  and each of its  $n$  components).

For all collision trees  $a \in \mathcal{A}_{n,m}$  and all parameters  $(t_i, \omega_i, v_i)_{i=1, \dots, m}$  with  $t_1 > t_2 > \dots > t_m$ , one constructs *pseudo-trajectories* on  $[0, t]$

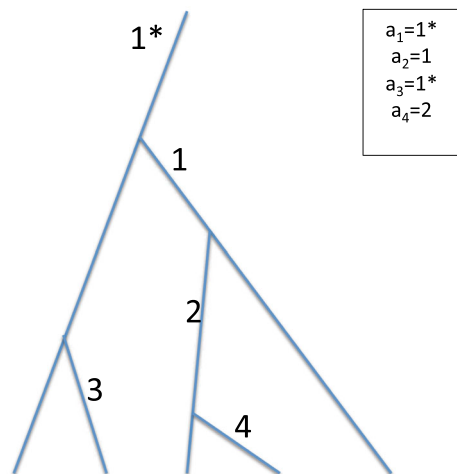
$$\Psi_{n,m}^\varepsilon = \Psi_{n,m}^\varepsilon \left( Z_n^*, (a_i, t_i, \omega_i, v_i)_{i=1, \dots, m} \right)$$

iteratively on  $i = 1, 2, \dots, m$  (denoting by  $Z_{n,m}(\tau) = (Z_n^*(\tau), Z_m(\tau))$  the coordinates of particles at time  $\tau \leq t_m$ ):

- starting from  $(Z_n^*)$  at time  $t =: t_0$ ,
- transporting all existing particles backward on  $[t_i, t_{i-1}]$  (on  $\mathcal{D}_{n+i-1}^\varepsilon$  with specular reflection at collisions),
- adding a new particle  $i$  at time  $t_i$ , with position  $x_{a_i}(t_i) + \varepsilon \omega_i$  and velocity  $v_i$ ,
- and applying the scattering rule (3.2) if  $(x_{a_i}(t_i), v_{a_i}(t_i^+), x_{a_i}(t_i) + \varepsilon \omega_i, v_i)$  is a post-collisional configuration.

We discard non admissible parameters for which this procedure is ill-defined; in particular we exclude values of  $\omega_i$  corresponding to an overlap of particles (two spheres at distance strictly smaller than  $\varepsilon$ ). In the following we denote by  $\mathcal{G}_m^\varepsilon(a, Z_n^*)$  the set of admissible parameters. Figure 2 is an example of such flow (for  $n = 1, m = 4$ ).

**Fig. 2** The tree structure of collisions



With these notations, one gets the following geometric representation of the correlation function  $F_n^\varepsilon$  :

$$F_n^\varepsilon(t, Z_n^*) = \sum_{m \geq 0} \sum_{a \in \mathcal{A}_{n,m}} \int_{\mathcal{G}_m^\varepsilon(a, Z_n^*)} dT_m d\Omega_m dV_m \left( \prod_{i=1}^m (v_i - v_{a_i}(t_i^+)) \cdot \omega_i \right) F_{n+m}^{\varepsilon 0}(\Psi_{n,m}^{\varepsilon 0}),$$

where  $(T_m, \Omega_m, V_m) := (t_i, \omega_i, v_i)_{1 \leq i \leq m}$ , and  $\Psi_{n,m}^{\varepsilon 0}$  is the  $(n + m)$ -particle configuration of the pseudo-trajectory at time zero. Or, in short,

$$F_n^\varepsilon(t, Z_n^*) = \int \mu(d\Psi_n^\varepsilon) \mathcal{C}(\Psi_n^\varepsilon) \mathbf{1}_{\mathcal{G}^\varepsilon}(\Psi_n^\varepsilon) F^{\varepsilon 0}(\Psi_n^{\varepsilon 0})$$

with  $\mu(d\Psi_n^\varepsilon) := \sum_m \sum_{a \in \mathcal{A}_{n,m}} dT_m d\Omega_m dV_m$ ,  $\mathcal{C}(\Psi_n^\varepsilon) := \prod_{i=1}^m (v_i - v_{a_i}(t_i^+)) \cdot \omega_i$  (3.15)

$\mathbf{1}_{\mathcal{G}^\varepsilon}(\Psi_n^\varepsilon) := \mathbf{1}_{\mathcal{G}_m^\varepsilon(a, Z_n^*)}$ , and  $F^{\varepsilon 0}(\Psi_n^{\varepsilon 0})$  the initial correlation function evaluated on the configuration at time 0 of the pseudo-trajectory (including  $n + m$  particles). From now on, we will indicate by  $\Psi_n^\varepsilon$  a generic pseudo-trajectory with  $n$  particles at time  $t = t_0$ .

### 3.3.2 A Short Time Estimate

Each elementary integral corresponding to a collision tree with  $m$  branching points involves a simplex in time  $(t_1 > t_2 > \dots > t_m)$ . Thus, if we replace, for simplicity, the cross-section factors  $\mathcal{C}(\Psi_1^\varepsilon)$  by a bounded function (cutting off high energies), we immediately get that the integrals for  $n = 1$  are bounded, for each fixed tree  $a \in \mathcal{A}_{1,m}$ , by

$$\left| \int dT_m d\Omega_m dV_m \mathcal{C}(\Psi_1^\varepsilon) \mathbf{1}_{\mathcal{G}^\varepsilon}(\Psi_1^\varepsilon) F^{\varepsilon 0}(\Psi_1^{\varepsilon 0}) \right| \leq \frac{(C'_0 t)^m}{m!},$$

where  $C'_0 > 0$  depends only on  $C_0, \beta_0$  of (3.3). Since  $|\mathcal{A}_{1,m}| = m!$ , the series expansion is therefore absolutely convergent for short times, uniformly in  $\varepsilon$ . A similar estimate holds for  $n > 1$ . Moreover in presence of the true factors  $\mathcal{C}(\Psi_n^\varepsilon)$ , the result remains valid (with a slightly different value of the convergence radius), though the proof requires some extra care [20].

Hence it is enough to study the convergence of each elementary term in the Boltzmann–Grad limit  $\varepsilon \rightarrow 0$ .

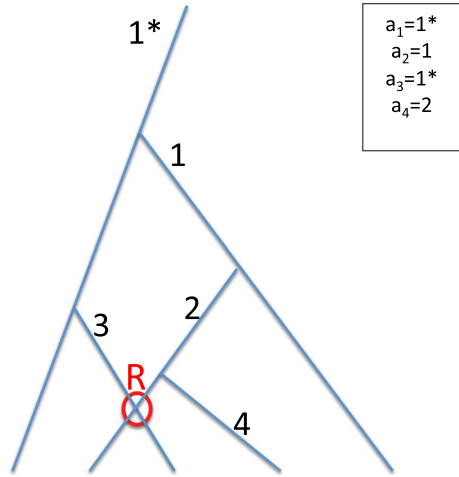
### 3.3.3 Removing Recollisions

When the size  $\varepsilon$  of the particles goes to 0, we expect the pseudo-trajectory  $\Psi_1^\varepsilon$  to converge to a limiting  $\Psi_1$ , defined iteratively on  $i = 1, 2, \dots, m$ :

- starting from  $z_1^*$  at time  $t = t_0$ ,
- transporting all existing particles backward on  $[t_i, t_{i-1}]$  (by free transport),
- adding a new particle  $i$  at time  $t_i$ , exactly at position  $x_{a_i}(t_i)$  and with velocity  $v_i$ ,
- and applying the scattering rule (3.2) if  $(v_i - v_{a_i}(t_i^+)) \cdot \omega_i > 0$  (post-collisional configuration).

The main obstacle to the convergence  $\Psi_1^\varepsilon \rightarrow \Psi_1$  are the so-called recollisions. In the language of pseudo-trajectories, a *recollision* is a collision between pre-existing particles (see Fig. 3), namely a collision which does not correspond to the addition of a fresh particle.

Fig. 3 An example of recollision



It is easy to realize that, in the absence of recollisions,  $\Psi_1^\varepsilon$  and  $\Psi_1$  differ only by small shifts in the positions.

A careful geometric analysis of recollisions shows that they can happen only for a small set of parameters, which is negligible in the limit  $\varepsilon \rightarrow 0$ . Roughly, if particles  $p$  and  $q$  are at positions  $x_p, x_q$  with  $x_p \neq x_q$  at time  $\tau > 0$ , then a recollision between these particles implies that there is a time  $t_{\text{rec}} < \tau$  such that  $x_p - x_q - (v_p - v_q)(\tau - t_{\text{rec}}) = O(\varepsilon)$ . As a consequence,  $v_p - v_q$  is constrained to be in a small cone of opening  $\varepsilon$ , and the integration parameters in (3.15) lie in a small set. Thanks to the uniform bounds, one concludes that pseudo-trajectories involving recollisions give an overall vanishing contribution to  $F_1^\varepsilon$ .

A similar analysis can be performed to study higher order correlation functions. However, in this case the convergence is slightly more subtle. Notice that, for the  $n$ -particle correlation function, the convergence will fail on some sets of parameters of volume  $\mu_\varepsilon^{-1}$ , which correspond to particles of different trees colliding in the backward dynamics (see e.g. Fig. 4). These “external recollisions” are apparently innocent, as they correspond again to small volume sets which do not contribute to the limit. On the other hand, it is the little failure of convergence of the  $F_n^\varepsilon$  which prevents Lanford’s theorem from being “reversible”, i.e. from being applicable to the state at time  $t > 0$  with reversed velocities [2,5]. This suggests that the relevant information to go backwards is hidden in singular directions where different trees merge. In Sect. 4, we will show that the dynamical cumulants  $f_n^\varepsilon$  defined by (3.11) “live” in these singular directions, thus allowing to investigate the  $n$ -particle correlations more precisely.

### 3.3.4 Averaging Over Trajectories

We conclude this section with a generalization of the previous discussion, which will be important in the following.

So far we discussed correlations in phase space, at a given time  $t$ . But clearly, spatio-temporal correlations are also of interest. We therefore need to study trajectories of particles, and not only their distribution at a given time. Pseudo-trajectories provide a geometric representation of the iterated Duhamel series, but they are not physical trajectories of the particle

system. Nevertheless, the probability of trajectories of  $n$  particles can be represented as above, by *conditioning* the Duhamel series.

**Proposition 1** [6] *Let  $H_n$  be a bounded measurable function on the Skorokhod space of trajectories over  $\mathbb{D}^n$  in  $[0, t]$ . Define*

$$F_{n,[0,t]}^\varepsilon(H_n) := \int dZ_n^* \int \mu(d\Psi_n^\varepsilon) \mathcal{C}(\Psi_n^\varepsilon) \mathbf{1}_{\mathcal{G}^\varepsilon}(\Psi_n^\varepsilon) H_n(Z_n^*([0, t])) F^{\varepsilon 0}(\Psi_n^{\varepsilon 0}), \tag{3.16}$$

where  $Z_n^*([0, t])$  are the trajectories of the  $n$   $*$ -tagged particles in the pseudo-trajectory  $\Psi_n^\varepsilon$ . Then,

$$\mathbb{E}_\varepsilon \left( \sum_{\substack{i_1, \dots, i_n \\ i_j \neq i_k, j \neq k}} H_n(\mathbf{z}_{i_1}^\varepsilon([0, t]), \dots, \mathbf{z}_{i_n}^\varepsilon([0, t])) \right) = \mu_\varepsilon^n F_{n,[0,t]}^\varepsilon(H_n),$$

where  $\mathbf{z}_{i_1}^\varepsilon([0, t]), \dots, \mathbf{z}_{i_n}^\varepsilon([0, t])$  is the sample path of  $n$  hard spheres labeled  $i_1, \dots, i_n$ , among the  $\mathcal{N}$  hard spheres randomly distributed at time zero.

This generalizes (3.6) and the representation (3.15), in the sense that, for  $H_n(Z_n^*([0, t])) = h_n(Z_n^*(t))$ , we obtain

$$F_{n,[0,t]}^\varepsilon(H_n) = \int F_n^\varepsilon(t, Z_n^*) h_n(Z_n^*) dZ_n^*.$$

## 4 Dynamical Correlations

### 4.1 Recollisions and Overlaps

We start from the representation (3.16) of  $F_{n,[0,t]}^\varepsilon(H_n)$  in terms of collision trees and pseudo-trajectories. We assume that

$$H_n = H^{\otimes n}$$

with  $H$  a measurable function on the Skorokhod space of trajectories  $D([0, t])$  in  $\mathbb{D}$ , and we abbreviate

$$\mathcal{H}(\Psi_n^\varepsilon) := H^{\otimes n}(Z_n^*([0, t])).$$

Recall that there are two types of interactions between particles:

- a collision correspond to the addition of a new particle;
- recollisions occurring when two pre-existing particles collide.

The elementary integrals in the series expansion of  $F_{n,[0,t]}^\varepsilon(H_n)$  can be decomposed depending on whether collision trees are correlated or not by recollisions (see Fig. 4). We then have a partition of  $\{1, \dots, n\}$  into a certain number (say  $\ell$ ) of *forests*  $(\lambda_i)_{i=1, \dots, \ell}$ , and we shall denote by  $\mathbb{A}_{\lambda_i}$  the characteristic function of the forest  $\lambda_i$ . Namely,  $\mathbb{A}_{\lambda_i} = 1$  if and only if any two elements of  $\lambda_i$  are connected (through their collision trees) by a chain of recollisions. We say that  $\mathbb{A}_{\lambda_i} = 1$  is supported on *clusters* of size  $|\lambda_i|$ , formed by  $|\lambda_i|$  “recolliding” collision trees. We will further indicate the decomposition in forests by  $\lambda = (\lambda_i)_{i=1, \dots, \ell}$ .

Formula (3.16) can then be rewritten as a partially factorized expression:

$$F_{n,[0,t]}^\varepsilon(H^{\otimes n}) = \int dZ_n^* \sum_{\ell=1}^n \sum_{\lambda \in \mathcal{P}_n^\ell} \int \mathcal{K}_\lambda(\Psi_\lambda^\varepsilon) \Phi_\ell F^{\varepsilon 0}(\Psi_n^{\varepsilon 0}) \tag{4.1}$$

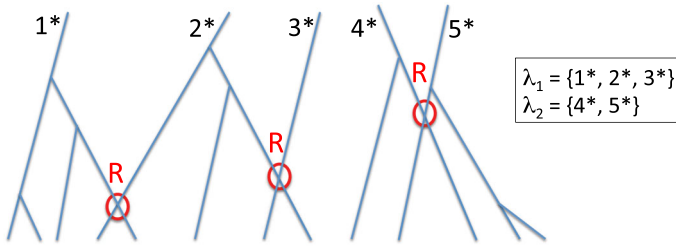


Fig. 4 Recollisions connect trees into forests

where

$$\mathcal{K}_\lambda(\Psi_\lambda^\varepsilon) = \prod_{i=1}^\ell [\mu(d\Psi_{\lambda_i}^\varepsilon) \Delta_{\lambda_i} \mathcal{C}(\Psi_{\lambda_i}^\varepsilon) \mathbf{1}_{\mathcal{G}^\varepsilon}(\Psi_{\lambda_i}^\varepsilon) \mathcal{H}(\Psi_{\lambda_i}^\varepsilon)] ,$$

and  $\Phi_\ell = \Phi_\ell(\lambda_1, \dots, \lambda_\ell)$  is the indicator function that particles belonging to different forests keep mutual distance larger than  $\varepsilon$ . Here and below, we indicate by  $\Psi_\alpha^\varepsilon$  the pseudo-trajectory constructed starting from  $Z_\alpha^*$ , for any  $\alpha$  subset of  $\{1, \dots, n\}$ .

Although there cannot be any recollision between particles of different forests  $\lambda_i$ , such particles are not yet independent, as the parameters of the pseudo-trajectories are constrained precisely by the fact that no recollision should occur. The characteristic function  $\Phi_\ell = \Phi_\ell(\lambda_1, \dots, \lambda_\ell)$  expresses this no-recollision condition. Next, we write its cumulant expansion (the analogue of (3.12)):

$$\begin{aligned} \Phi_\ell &= \sum_{r=1}^\ell \sum_{\rho \in \mathcal{P}_\ell^r} \varphi_\rho . \\ \varphi_\rho &= \prod_{i=1}^r \varphi_{\rho_i} . \end{aligned} \tag{4.2}$$

This formula reorganizes the  $\ell$  forests into a group of  $r$  jungles  $\rho = (\rho_i)_{i=1, \dots, r}$ .

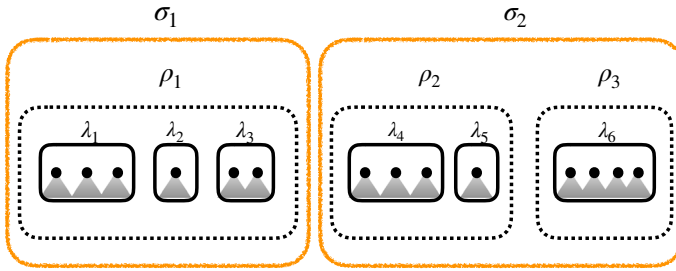
By construction, particles belonging to different forests will never collide among themselves. However they are allowed to “overlap”. We say that two different forests  $\lambda_i$  and  $\lambda_j$  overlap if two particles, belonging to the pseudo-trajectories  $\Psi_{\lambda_i}^\varepsilon$  and  $\Psi_{\lambda_j}^\varepsilon$  respectively, touch each other (without colliding) and cross each other freely. Standard combinatorial arguments show then that the cumulant  $\varphi_s$  of order  $s$  is supported on clusters of size  $s$ , formed by  $s$  overlapping forests (namely any two forests are connected by a chain of overlaps).

The last source of correlation in (3.16) comes from the initial data. For each given  $\rho$ , we introduce a cumulant expansion of the initial data associated with  $\rho$ :

$$F^{\varepsilon 0}(\Psi_n^{\varepsilon 0}) = \sum_{s=1}^r \sum_{\sigma \in \mathcal{P}_r^s} f_\sigma^{\varepsilon 0} , \quad f_\sigma^{\varepsilon 0} = \prod_{i=1}^s f_{\sigma_i}^{\varepsilon 0} , \quad f_{\sigma_i}^{\varepsilon 0} = f^{\varepsilon 0}(\Psi_{\sigma_i}^{\varepsilon 0}) . \tag{4.3}$$

Here and below, by abuse of notation, the partitions  $\sigma, \rho$  are also interpreted as a partition of  $\{1, \dots, n\}$ , coarser than the partition  $\lambda$ ; the relative coarseness will be denoted by  $\lambda \hookrightarrow \rho \hookrightarrow \sigma$ . Therefore  $f^{\varepsilon 0}(\Psi_{\sigma_i}^{\varepsilon 0})$  is the time-zero block cumulant evaluated on the configuration (at time 0) of the pseudo-trajectory starting from  $Z_{\sigma_i}^*$ .

We end up with a cluster structure on collision trees, of the form depicted in Fig. 5.



**Fig. 5** Clustering structure due to recollisions, overlaps and initial correlations

Replacing (4.2) and (4.3) into (4.1), we arrive to the following decomposition of correlation functions :

$$F_{n,[0,t]}^\varepsilon(H^{\otimes n}) = \int dZ_n^* \sum_{\substack{\lambda, \rho, \sigma \\ \lambda \hookrightarrow \rho \hookrightarrow \sigma}} \int \mathcal{K}_\lambda \varphi_\rho f_\sigma^{\varepsilon 0}, \tag{4.4}$$

where  $\lambda$  is the partition of  $\{1, \dots, n\}$  into  $\ell$  forests of recolliding trees,  $\rho$  is the partition of  $\{1, \dots, \ell\}$  into  $r$  jungles of overlapping forests, and  $\sigma$  is the partition of  $\{1, \dots, r\}$  into initially correlated clusters.

### 4.2 Cumulants and Clusters

Comparing formula (4.4) with (3.12), we finally identify the rescaled dynamical cumulants (averaged over trajectories):

$$f_{n,[0,t]}^\varepsilon(H^{\otimes n}) = \mu_\varepsilon^{n-1} \int dZ_n^* \sum_{\ell=1}^n \sum_{\lambda \in \mathcal{P}_n^\ell} \sum_{r=1}^\ell \sum_{\rho \in \mathcal{P}_\ell^r} \int \mathcal{K}_\lambda \varphi_\rho f_{\{1, \dots, r\}}^{\varepsilon 0}(\Psi_{\rho_1}^{\varepsilon 0}, \dots, \Psi_{\rho_r}^{\varepsilon 0}). \tag{4.5}$$

This result shows that the cumulant of order  $n$  is geometrically represented by connected clusters of size  $n$  :  $f_{n,[0,t]}^\varepsilon$  corresponds to pseudo-trajectories where the  $n$  collision trees are connected by recollisions, overlaps, or initial correlations. This graphical representation of cumulants leads to the following result.

**Proposition 2** (Convergence of dynamical cumulants, [6]) *Consider a gas of hard spheres initially distributed according to (3.4). Let  $H$  be a bounded continuous functional on  $D([0, T^*])$ . Define the rescaled cumulant  $f_{n,[0,t]}^\varepsilon(H^{\otimes n})$  by (4.5). Then,*

- there exists a positive constant  $C$  such that the following uniform a priori bound holds

$$|f_{n,[0,t]}^\varepsilon(H^{\otimes n})| \leq C^n \|H\|_\infty^n (t + \varepsilon)^{n-1} n!$$

uniformly in  $\varepsilon$  and  $n$ , for any  $t \leq T^*$ ;

- when  $\varepsilon \rightarrow 0$ , in the same time interval,  $f_{n,[0,t]}^\varepsilon(H^{\otimes n})$  converges to a limiting  $f_{n,[0,t]}(H^{\otimes n})$ , which is represented by a sum over minimally connected graphs, and by pseudo-trajectories with exactly  $n - 1$  pointwise recollisions or overlaps.

The key point to obtain the right scaling of cumulants is to identify “independent” clustering constraints : for fixed  $\lambda, \rho$ , collision parameters  $a$  and  $(T_m, \Omega_m, V_m)$  and initial velocities  $V_n^*$  :

- we extract a sequence of  $|\lambda_i| - 1$  clustering recollisions in each forest  $\lambda_i$ , a sequence of  $|\rho_j| - 1$  clustering overlaps in each jungle  $\rho_j$ , and a sequence of  $r - 1$  clustering initial correlations, and prove that the factor  $n!$  accounts for the combinatorics of these clustering constraints;
- we then show that the clustering constraints can be expressed as  $n - 1$  conditions on the positions at time  $t$  of the particles of the pseudo-trajectory  $(x_j^*)_{j=1, \dots, n}$ , which are satisfied on a set of volume  $O(\mu_\varepsilon^{-(n-1)})$ .

These estimates being essentially uniform with respect to the collision parameters  $(a, T_k, \Omega_k, V_k)$ , we can sum/integrate to get the  $L^1$ -bound.

There is a subtle point here: a brute expansion of the overlap constraint  $\varphi_s$  defined by (4.2), leads to  $2^{s^2}$  terms, and cancellations need to be exploited to show that the effective number is bounded by  $s!$ . How to do this is known by cluster expansion techniques (see e.g. [15,24,29]). In fact,  $\varphi_s$  can be regarded as an Ursell function ([29]) by writing formally “ $\Phi_\ell(\lambda_1, \dots, \lambda_\ell) = \exp(-U_\ell(\lambda_1, \dots, \lambda_\ell))$ ” and interpreting  $U$  as a hard core interaction on dynamical collision trees.

The proof of the second statement of Proposition 2 is very similar to Lanford’s proof. We first discard the contribution of initial correlations (which are of order  $O(\varepsilon^d)$  instead of  $O(\varepsilon^{d-1})$ ). We then prove (as discussed in Sect. 3.3.3) that any recollision which is *not* of clustering type will create some extra smallness, giving a vanishing contribution to the limit.

### 4.3 Cumulant Generating Function

The cumulants allow to characterize exponential moments of the empirical measure, as shown by the following identity :

$$\Lambda_{[0,t]}^\varepsilon(H) := \frac{1}{\mu_\varepsilon} \log \mathbb{E}_\varepsilon \left( \exp \left( \sum_{i=1}^{\mathcal{N}} H(\mathbf{z}_i^\varepsilon([0, t])) \right) \right) = \sum_{n=1}^\infty \frac{1}{n!} f_{n,[0,t]}^\varepsilon((e^H - 1)^{\otimes n}), \quad (4.6)$$

valid for functionals  $H : D([0, t]) \rightarrow \mathbb{R}$  such that the series is absolutely convergent. In order to describe the asymptotic behavior of these exponential moments, we need to obtain dynamical equations for the limiting cumulant generating function

$$\sum_{n=1}^\infty \frac{1}{n!} f_{n,[0,t]}((e^H - 1)^{\otimes n}), \quad (4.7)$$

which is well defined (as a corollary of Proposition 2) for  $t \in [0, T^*]$  provided that  $H$  is a continuous functional satisfying a suitable bound:

$$\left| \left( e^{H(z([0,t]))} - 1 \right)^{\otimes n} \right| \leq \exp \left( \alpha_0 n + \frac{\beta_0}{4} \sup_{s \in [0,t]} |V_n(s)|^2 \right),$$

for some  $\alpha_0$  (related to the constant  $C_0$  in (3.3) and to  $T^*$ ).

We shall not write here the hierarchical equations for the family of cumulants at equal times  $(f_n(t))_{n \geq 1}$ , obtained by choosing  $H(z([0, t])) = h(z(t))$  in (4.6). This hierarchy (mentioned in Sect. 2 as “Boltzmann cumulant hierarchy”) is derived and analysed in [13]. Our purpose is to focus directly on the full series (4.7), which we study for a class of regular test functionals.



For  $t \in [0, T^*]$ , denote by  $\mathcal{J}(t, \varphi, \gamma)$  the limiting cumulant generating function (4.7) associated with

$$e^{H(z(0,t))} = \gamma(z(t)) \exp\left(-\int_0^t \varphi(s, z(s)) ds\right),$$

where  $(\varphi, \gamma)$  belong to

$$\mathcal{B} := \left\{ (\varphi, \gamma) \in C^1([0, t] \times \mathbb{D}; \mathbb{C}) \times C^1(\mathbb{D}; \mathbb{C}) \mid |\gamma(z)| \leq e^{\frac{1}{2}(\alpha_0 + \frac{\beta_0}{4}|v|^2)}, \sup_{s \in [0, T^*]} |\varphi(s, z)| \leq \frac{1}{2T^*} \left(\alpha_0 + \frac{\beta_0}{4}|v|^2\right) \right\}.$$

We shall be interested in functions of the form

$$\varphi = D_s h \equiv (\partial_s + v \cdot \nabla_x)h \quad \text{and} \quad \gamma = \exp(h(t)),$$

therefore we simplify notation by setting

$$\mathcal{J}(t, h) := \mathcal{J}(t, Dh, \gamma)|_{\gamma=\exp(h(t))}.$$

We set  $\mathbb{B} := \{h \mid (D_t h, \exp(h(T^*))) \in \mathcal{B}\}$ . With these notations, the following result holds.

**Theorem 2** (Hamilton–Jacobi equations, [6]) *The functional  $\mathcal{J}$  is analytic with respect to  $\gamma$  on  $\mathbb{B}$ , and it satisfies on  $[0, T^*]$  the Hamilton–Jacobi equation*

$$\begin{aligned} \partial_t \mathcal{J}(t, h) &= \frac{1}{2} \int \frac{\partial \mathcal{J}(t, h)}{\partial \gamma}(z_1) \frac{\partial \mathcal{J}(t, h)}{\partial \gamma}(z_2) \\ &\quad \times \left( e^{h(t, z'_1) + h(t, z'_2)} - e^{h(t, z_1) + h(t, z_2)} \right) d\mu(z_1, z_2, \omega), \end{aligned} \tag{4.8}$$

where

$$d\mu(z_1, z_2, \omega) := \delta(x_1 - x_2) ((v_1 - v_2) \cdot \omega)_+ d\omega dv_1 dv_2 dx_1 dx_2. \tag{4.9}$$

The local existence and uniqueness of a solution for this Hamilton–Jacobi equation relies on a Cauchy–Kowalewski argument in a functional space, encoding the loss continuity estimates due to the divergence of the collision cross section (4.9) at large velocities.

### 5 The Fluctuating Boltzmann Equation

Describing the fluctuations around the Boltzmann equation is a first way to capture part of the information which has been lost in the limit (1.2). As in the standard central limit theorem, we expect these fluctuations to be of order  $1/\sqrt{\mu_\varepsilon}$ . We therefore define the fluctuation field  $\zeta^\varepsilon$  by (see (1.4))

$$\zeta_t^\varepsilon(h) := \sqrt{\mu_\varepsilon} \left( \pi_t^\varepsilon(h) - \int F_1^\varepsilon(t, z) h(z) dz \right),$$

for any test function  $h : \mathbb{D} \rightarrow \mathbb{R}$ .

It is easy to check that, in our assumptions, the empirical measure starts close to the density profile  $f^0$  and that  $\zeta_0^\varepsilon$  converges to a Gaussian white noise  $\zeta_0$  with covariance

$$\mathbb{E}(\zeta_0(h_1) \zeta_0(h_2)) = \int h_1(z) h_2(z) f^0(z) dz.$$

Moreover, it follows from Proposition 2 that  $\zeta^\varepsilon$  converges to a solution of the fluctuating Boltzmann equation

$$d\zeta_t = \mathcal{L}_t \zeta_t dt + d\eta_t, \tag{5.1}$$

where  $\mathcal{L}_t$  is the linearized Boltzmann operator around the solution  $f$  of the Boltzmann equation (3.9)

$$\begin{aligned} \mathcal{L}_t h(x, v) := & -v \cdot \nabla_x h(x, v) + \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} d\omega dw ((v-w) \cdot \omega)_+ (f(t, x, w')h(x, v') \\ & + f(t, x, v')h(x, w') - f(t, x, v)h(x, w) - f(t, x, w)h(x, v)), \end{aligned}$$

and  $d\eta_t(x, v)$  is a Gaussian noise with zero mean and covariance

$$\begin{aligned} \mathbb{E} \left( \int dt_1 dz_1 h_1(z_1) \eta_{t_1}(z_1) \int dt_2 dz_2 h_2(z_2) \eta_{t_2}(z_2) \right) \\ = \frac{1}{2} \int dt \int f(t, z_1) f(t, z_2) \Delta h_1 \Delta h_2 d\mu(z_1, z_2, \omega) \end{aligned} \tag{5.2}$$

with notation (4.9) and

$$\Delta h(z_1, z_2, \omega) := h(x_1, v'_1) + h(x_2, v'_2) - h(x_1, v_1) - h(x_2, v_2). \tag{5.3}$$

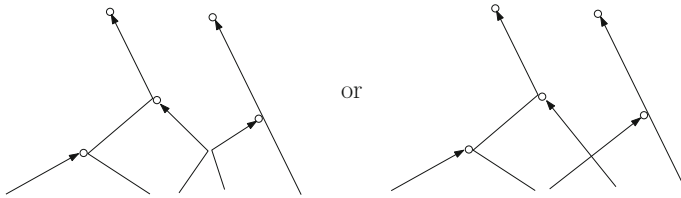
Our main result is then the following.

**Theorem 3** (Fluctuating Boltzmann equation, [6]) *Consider a system of hard spheres initially distributed according to (3.4). Then, in the Boltzmann–Grad limit  $\mu_\varepsilon \rightarrow \infty$ , the fluctuation field  $(\zeta_t^\varepsilon)_{t \geq 0}$  converges in law on  $[0, T^*]$  to the solution  $(\zeta_t)_{t \geq 0}$  of the fluctuating Boltzmann equation (5.1).*

The convergence towards the limiting process (5.1) was conjectured by Spohn in [31] and the non-equilibrium covariance of the process at two different times was obtained in [30]. The noise emerges after averaging the deterministic microscopic dynamics. It is white in time and space, but correlated in velocities so that momentum and energy are conserved.

We further recall a few properties (referring to [13,30,31] for details).

- In the equilibrium case ( $f^0(x, v) = M_\beta(v)$  where  $M_\beta$  is a Maxwellian with inverse temperature  $\beta$ ) the noise term compensates the dissipation induced by the (stationary) linearized collision operator  $\mathcal{L}$ , and the covariance of the noise can be predicted heuristically by using the invariant measure.
- Out of equilibrium, on the one hand, the noise covariance (5.2) can be simply understood as a generalization of the covariance at equilibrium, based on the assumption (which can be proved for short times [32]) that the system is locally Poisson distributed; i.e. on a small cube around  $x$  at time  $t$  we see a uniform ideal gas with density  $\int f(t, x, v)dv$  and velocity distribution  $f(t, x, v)/\int f(t, x, v_*)dv_*$ . The noise being delta-correlated in space and time, its structure is obtained from the equilibrium case after the replacement  $M_\beta(v) \rightarrow f(t, x, v)$ .
- On the other hand, the covariance of the fluctuation field out of equilibrium has a subtle microscopic structure originating from recollisions in the Newtonian dynamics. To see this, it is enough to compute the covariance of the fluctuation field at time  $t$  by using (3.10) :



**Fig. 6** Trajectories contributing to the equal time covariance at two different points in space (case of  $m = 3$  collisions with fresh particles): clustering recollisions or clustering overlaps

$$\begin{aligned} \mathbb{E}_\varepsilon \left( \zeta_t^\varepsilon(h)^2 \right) &= \int F_1^\varepsilon(t, z_1) h^2(z_1) dz_1 \\ &\quad + \int \mu_\varepsilon \left( F_2^\varepsilon(t, Z_2) - F_1^\varepsilon(t, z_1) F_1^\varepsilon(t, z_2) \right) h(z_1) h(z_2) dZ_2 \\ &= \int f_1^\varepsilon(t, z_1) h^2(z_1) dz_1 + \int f_2^\varepsilon(t, Z_2) h(z_1) h(z_2) dZ_2 \end{aligned}$$

where, in the second equality, we used the first two cumulants as defined by (3.12). The last term is zero at equilibrium, while out of equilibrium describes correlations visible at macroscopic distance in space. But, as made apparent from the geometrical representation (4.5),  $f_2^\varepsilon$  records the effect of one recollision/overlap; meaning precisely that the pseudo-trajectories contributing to  $f_2^\varepsilon$  have the form in Fig. 6. Contrary to the typical behavior of the hard sphere gas for which recollisions can be neglected, the covariance of the limiting Gaussian process encodes exactly the effect of a single recollision.

The uniform bounds on the cumulants discussed in the previous section are considerably better than what is required to obtain Theorem 3. The proof amounts indeed to looking at a characteristic function living on larger scales. A more technical part concerns the tightness of the process. This can be achieved adapting a Garsia–Rodemich–Rumsey’s inequality on the modulus of continuity, to the case of a discontinuous process. We omit the details, and focus on the characteristic function only.

Consider the function  $H$  defined by

$$H(z([0, t])) = \sum_{p=1}^P h_p(z(\theta_p)) \tag{5.4}$$

for a finite sequence of times  $(\theta_p)_{1 \leq p \leq P}$  and weights  $(h_p)_{1 \leq p \leq P}$ . The characteristic function can be rewritten in terms of the empirical measure

$$\begin{aligned} \log \mathbb{E}_\varepsilon \left( \exp \left( \sum_{p=1}^P \zeta_{\theta_p}^\varepsilon(h_p) \right) \right) &= \mu_\varepsilon \sum_{n=1}^\infty \frac{1}{n!} f_{n, [0, t]}^\varepsilon \left( \left( e^{\frac{H}{\sqrt{\mu_\varepsilon}}} - 1 \right)^{\otimes n} \right) \\ &\quad - \sqrt{\mu_\varepsilon} \sum_{p=1}^P \int F_1^\varepsilon(\theta_p, z) h_p(z) dz. \end{aligned}$$

At leading order, only the terms  $n = 1$  and  $n = 2$  will be relevant in the limit since

$$\left| f_{n, [0, t]}^\varepsilon \left( \left( e^{\frac{H}{\sqrt{\mu_\varepsilon}}} - 1 \right)^{\otimes n} \right) \right| \leq C^n \left\| e^{\frac{H}{\sqrt{\mu_\varepsilon}}} - 1 \right\|_\infty^n n!$$

Expanding the exponential with respect to  $\mu_\varepsilon$ , we also notice that the term of order  $\sqrt{\mu_\varepsilon}$  cancels and we find

$$\log \mathbb{E}_\varepsilon \left( \exp \left( \sum_{p=1}^P \zeta_{\theta_p}^\varepsilon(h_p) \right) \right) = \frac{1}{2} f_{1,[0,t]}^\varepsilon(H^2) + \frac{1}{2} f_{2,[0,t]}^\varepsilon(H^{\otimes 2}) + O \left( \frac{1}{\sqrt{\mu_\varepsilon}} \right).$$

Then the characteristic function  $\mathbb{E}_\varepsilon \left( \exp \left( \sum_{p=1}^P \zeta_{\theta_p}^\varepsilon(h_p) \right) \right)$  converges to the characteristic function of a Gaussian process.

From the equations on  $f_1$  and  $f_2$ , we deduce that the limiting covariance  $\mathcal{C} = \mathcal{C}(s, t, \varphi, \psi)$  satisfies the following dynamical equations, for test functions  $\varphi, \psi$  on  $\mathbb{D}$  :

$$\begin{cases} \partial_t \mathcal{C}(s, t, \varphi, \psi) = \mathcal{C}(s, t, \varphi, \mathcal{L}_t^* \psi), \\ \partial_t \mathcal{C}(t, t, \varphi, \psi) = \mathcal{C}(t, t, \varphi, \mathcal{L}_t^* \psi) + \mathcal{C}(t, t, \mathcal{L}_t^* \varphi, \psi) + \mathbf{Cov}_t(\psi, \varphi), \\ \mathcal{C}(0, 0, \varphi, \psi) = \int \varphi(z) \psi(z) f^0(z) dz, \end{cases}$$

where

$$\mathbf{Cov}_t(\varphi, \psi) := \frac{1}{2} \int d\mu(z_1, z_2, \omega) f(t, z_1) f(t, z_2) \Delta \psi \Delta \varphi$$

with notation (4.9) and (5.3), and  $\mathcal{L}_t^* := v \cdot \nabla_x + \mathbf{L}_t^*$  with

$$\mathbf{L}_t^* \varphi(v) := \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} d\omega dw ((v - w) \cdot v)_+ f(t, w) \Delta \varphi.$$

The covariance of the fluctuating Boltzmann equation (5.1) satisfies the same equations, and we conclude by a uniqueness argument that both processes coincide.

### 6 Large Deviations

While typical fluctuations are of order  $O(\mu_\varepsilon^{-1/2})$ , larger fluctuations may sometimes happen, leading to an evolution which is different from the typical one given by the Boltzmann equation. A classical problem is to evaluate the probability of such atypical events, namely that the empirical measure  $\pi_t^\varepsilon$ , defined in (1.1), remains close to a probability density  $\varphi_t$  during the time interval  $[0, T^*]$ .

In the Gärtner-Ellis theory of large deviations [11], the large deviation functional is given as the Legendre transform of the limiting cumulant generating function. The outcome of the cumulant analysis was the existence of the limiting exponential moment  $\mathcal{J}(t, h)$  and its characterization via the Hamilton–Jacobi equation in Theorem 2. For any  $t \leq T^*$ , we then define the large deviation functional on the time interval  $[0, t]$  as

$$\begin{aligned} \mathcal{F}(t, \varphi) := \sup_{h \in \mathbb{B}} \left\{ - \int_0^t \int_{\mathbb{D}} \varphi(s, z) D_s h(s, z) dz ds \right. \\ \left. + \int_{\mathbb{D}} \varphi(t, z) h(t, z) dz - \mathcal{J}(t, h) \right\}. \end{aligned} \tag{6.1}$$

Since the supremum is restricted, for technical reasons, to the test functions in  $\mathbb{B}$ , we do not expect  $\mathcal{F}$  to be the correct large deviation functional. However the following theorem shows that the functional  $\mathcal{F}$  fully describes the large deviation behavior for densities  $\varphi$  such that

the supremum in (6.1) is reached for some  $h \in \mathbb{B}$ . This restricted set of densities  $\varphi$  will be called  $\mathcal{R}$ .

A different, explicit formula for the large deviation functional was obtained by Rezakhanlou [28] in the case of a one-dimensional stochastic dynamics mimicking the hard-sphere dynamics, and then conjectured for the three-dimensional, deterministic hard-sphere dynamics by Bouchet [7] :

$$\widehat{\mathcal{F}}(t, \varphi) := \widehat{\mathcal{F}}(0, \varphi_0) + \sup_p \left\{ \int_0^t ds \left[ \int_{\mathbb{T}^d} dx \int_{\mathbb{R}^d} dv p(s, x, v) D_s \varphi(s, x, v) - \mathcal{H}(\varphi(s), p(s)) \right] \right\}, \tag{6.2}$$

where the supremum is taken over bounded measurable functions  $p$  growing at most quadratically in  $v$ , the Hamiltonian is given by

$$\mathcal{H}(\varphi, p) := \frac{1}{2} \int d\mu(z_1, z_2, \omega) \varphi(z_1) \varphi(z_2) (\exp(\Delta p) - 1)$$

and  $\widehat{\mathcal{F}}(0, \cdot)$  stands for the large deviation functional on the initial data

$$\widehat{\mathcal{F}}(0, \varphi_0) = \int dz \left( \varphi_0 \log \left( \frac{\varphi_0}{f^0} \right) - \varphi_0 + f^0 \right). \tag{6.3}$$

Let  $\widehat{\mathcal{R}}$  denote the set of densities  $\varphi$  such that the supremum in (6.2) is reached for some  $p \in \mathbb{B}$ .

Let  $\mathcal{M}(\mathbb{D})$  be the set of probability measures on  $\mathbb{D}$ .

Our main result is then the following.

**Theorem 4** (Large deviations, [6]) *Consider a system of hard spheres initially distributed according to (3.4). In the Boltzmann–Grad limit  $\mu_\varepsilon \rightarrow \infty$ , the empirical measure  $\pi^\varepsilon$  satisfies the following large deviation estimates for any  $t \in [0, T^*]$ .*

– For any compact set  $\mathbf{F}$  of the Skorokhod space  $D([0, T^*], \mathcal{M})$ ,

$$\limsup_{\mu_\varepsilon \rightarrow \infty} \frac{1}{\mu_\varepsilon} \log \mathbb{P}_\varepsilon (\pi^\varepsilon \in \mathbf{F}) \leq - \inf_{\varphi \in \mathbf{F}} \mathcal{F}(T^*, \varphi). \tag{6.4}$$

– For any open set  $\mathbf{O}$  of the Skorokhod space  $D([0, T^*], \mathcal{M})$ ,

$$\liminf_{\mu_\varepsilon \rightarrow \infty} \frac{1}{\mu_\varepsilon} \log \mathbb{P}_\varepsilon (\pi^\varepsilon \in \mathbf{O}) \geq - \inf_{\varphi \in \mathbf{O} \cap \mathcal{R}} \mathcal{F}(T^*, \varphi). \tag{6.5}$$

Moreover, for any  $\varphi \in \mathcal{R} \cap \widehat{\mathcal{R}}$  and  $t$  sufficiently small, one has that  $\mathcal{F}(t, \varphi) = \widehat{\mathcal{F}}(t, \varphi)$ .

Given our precise control of the exponential moments, the large deviation proof is standard. Note that, in absence of global convexity, we cannot succeed in proving a full large deviation principle. However, restricting to a class of regular profiles, the variational problem defining the dual of  $\widehat{\mathcal{F}}$  can be uniquely solved and identified with the solution of the Hamilton–Jacobi equation (4.8). The result then follows from a uniqueness property of (4.8).

**Acknowledgements** We are very grateful to H. Spohn, F. Bouchet, F. Rezakhanlou, G. Basile, D. Benedetto, and L. Bertini for many enlightening discussions on the subjects treated in this text. T.B. acknowledges the support of ANR-15-CE40-0020-01 Grant LSD.

## References

- Alexander, R.K.: The infinite hard sphere system. Ph.D. Thesis, Dep. of Math., University of California at Berkeley (1975)
- van Beijeren, H., Lanford III, O.E., Lebowitz, J.L., Spohn, H.: Equilibrium time correlation functions in the low-density limit. *J. Stat. Phys.* **22**, 2 (1980)
- Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., Landim, C.: Macroscopic fluctuation theory. *Rev. Mod. Phys.* **87**, 593–636 (2015)
- Bodineau, T., Gallagher, I., Saint-Raymond, L.: From hard sphere dynamics to the Stokes–Fourier equations: an  $L^2$  analysis of the Boltzmann–Grad limit. *Ann. PDE* **3**(1), 2 (2017)
- Bodineau, T., Gallagher, I., Saint-Raymond, L., Simonella, S.: One-sided convergence in the Boltzmann–Grad limit. *Ann. Fac. Sci.* **27**(5), 985–1022 (2018)
- Bodineau, T., Gallagher, I., Saint-Raymond, L., Simonella, S.: Fluctuating Boltzmann equation and large deviations for a hard sphere gas. In preparation
- Bouchet, F.: Is the Boltzmann equation reversible? A large deviation perspective on the irreversibility paradox. [arXiv:2002.10398](https://arxiv.org/abs/2002.10398)
- Cercignani, C.: On the Boltzmann equation for rigid spheres. *Transp. Theory Stat. Phys.* **2**, 211–225 (1972)
- Cercignani, C., Illner, R., Pulvirenti, M.: *The Mathematical Theory of Dilute Gases*. Springer, New York (1994)
- Cohen, E.G.D.: Cluster expansions and the hierarchy. I. Non-equilibrium distribution functions. *Physica* **28**, 1045–1059 (1962)
- Dembo, A., Zeitouni, O.: *Large Deviations Techniques and Applications*. Springer, New York (2010)
- Derrida, B.: Microscopic versus macroscopic approaches to non-equilibrium systems. *J. Stat. Mech.* **2011**, P01030 (2011)
- Ernst, M.H., Cohen, E.G.D.: Nonequilibrium fluctuations in  $\mu$  space. *J. Stat. Phys.* **25**, 1 (1981)
- Gallagher, I., Saint Raymond, L., Texier, B.: *From Newton to Boltzmann: Hard Spheres and Short-Range Potentials*. EMS, Zurich (2014)
- Gallavotti, G., Bonetto, F., Gentile, G.: *Aspects of Ergodic, Qualitative and Statistical Theory of Motion*. Springer-Verlag, New York (2004)
- Grad, H.: On the kinetic theory of rarefied gases. *Commun. Pure Appl. Math.* **2**(4), 331–407 (1949)
- Kac, M.: Foundations of kinetic theory. In: *Proceedings of The third Berkeley symposium on mathematical statistics and probability*. Berkeley and Los Angeles, California: University of California Press, pp. 171–197 (1956)
- Kac, M., Logan, J.: Fluctuations and the Boltzmann equation. *Phys. Rev. A* **13**, 458–470 (1976)
- van Kampen, N.G.: Fluctuations in Boltzmann’s equation. *Phys. Rev.* **50A**(4), 237 (1974)
- King, F.: BBGKY Hierarchy for Positive Potentials. Ph.D. Thesis, Dep. of Math., Univ. California, Berkeley (1975)
- Lanford, O.E.: Time evolution of large classical systems. In: Moser, J. (ed.) *Dynamical Systems, Theory and Applications*. Lecture Notes in Physics. Springer-Verlag, Berlin (1975)
- Léonard, C.: On large deviations for particle systems associated with spatially homogeneous Boltzmann type equations. *Probab. Theory Relat. Fields* **101**, 1–44 (1995)
- Méléard, S.: Convergence of the fluctuations for interacting diffusions with jumps associated with Boltzmann equations. *Stoch. Stoch. Rep.* **63**(3–4), 195–225 (1998)
- Poghosyan, O., Ueltschi, D.: Abstract cluster expansion with applications to statistical mechanical systems. *J. Math. Phys.* **50**, 050309 (2009)
- Pulvirenti, M., Saffirio, C., Simonella, S.: On the validity of the Boltzmann equation for short range potentials. *Rev. Math. Phys.* **26**, 2 (2014)
- Pulvirenti, M., Simonella, S.: The Boltzmann–Grad limit of a hard sphere system: analysis of the correlation error. *Inventiones* **207**(3), 1135–1237 (2017)
- Rezakhanlou, F.: Equilibrium fluctuations for the discrete Boltzmann equation. *Duke Math. J.* **93**(2), 257–288 (1998)
- Rezakhanlou, F.: Large deviations from a kinetic limit. *Ann. Probab.* **26**(3), 1259–1340 (1998)
- Ruelle, D.: *Statistical Mechanics. Rigorous Results*. W.A. Benjamin Inc., New York (1969)

30. Spohn, H.: Fluctuations around the Boltzmann equation. *J. Stat. Phys.* **26**, 2 (1981)
31. Spohn, H.: Fluctuation theory for the Boltzmann equation. In: Lebowitz, J.L., Montroll, E.W. (eds.) *Nonequilibrium Phenomena I: The Boltzmann Equation*. North-Holland, Amsterdam (1983)
32. Spohn, H.: *Large Scale Dynamics of Interacting Particles*. Texts and Monographs in Physics. Springer, Heidelberg (1991)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.