Journal of Statistical Physics, Vol. 66, Nos. 1/2, 1992

Kinetic Limits of the HPP Cellular Automaton

Anna De Masi,[1] Raffaele Esposito,[2] and Errico Presutti[2]

Received February 18, 1991

We study the Boltzmann–Grad limit in various versions of the two-dimensional HPP cellular automaton. In the completely deterministic case we prove convergence to an evolution that is not of kinetic type, a well-known phenomenon after Uchyiama's paper on the Broadwell gas, whereas the limiting equation becomes of kinetic type in the model with random collisions. The main part of the paper concerns the case where the collisions are deterministic and the randomness comes from inserting, between any two successive HPP updatings, $\varepsilon^{-\nu}$ stirring updatings, $\nu < 1$ being any fixed positive number and ε a parameter which tends to 0. The initial measure is a product measure with average occupation numbers of the order of ε (low-density limit) and varying on distances of the order of ε^{-1}. The limit as $\varepsilon \to 0$ of the system evolved for times of the order of $\varepsilon^{-1-\nu}$ corresponds to the Boltzmann–Grad limit. We prove propagation of chaos and that the renormalized average occupation numbers (i.e., divided by ε) converge to the solution of the Broadwell equation. Convergence is proven at all times for which the solution of the Broadwell equation is bounded.

KEY WORDS: Cellular automata; kinetic theory; stochastic processes.

1. INTRODUCTION

After the proof of the existence of global solutions to the Boltzmann equation obtained by Di Perna and Lions, one of the most interesting and challenging problems in the field is the derivation of the Boltzmann equation beyond the times for which Lanford's proof[8] applies. No real progres has been made except for the result by Illner and Pulvirenti[7] on a dilute gas which expands in the vacuum. The same problems have been more recently considered in the framework of stochastic interacting particle systems and discrete velocity Boltzmann equations, not only for the lack of

[1] Dipartimento di Matematica, Università di L'Aquila, L'Aquila 67100, Italy.
[2] Dipartimento di Matematica, Università di Roma Tor Vergata, 00133 Rome, Italy.

results on purely Hamiltonian systems, but also for the intrinsic interest that stochastic particle systems have both theoretically and in computer simulations. See Spohn[11] for a survey of these problems.

Discrete velocity Boltzmann equations and random evolutions are intimately related: Uchyiama[13] has in fact shown that even if all the correlation functions for the Broadwell gas, a system of hard squares with four velocities (the Broadwell analogue of the Boltzmann hard-sphere model) converge in the Boltzmann–Grad limit, yet the limiting density does not satisfy the desired Broadwell equation. The same arguments, presented, for the sake of completeness, in Section 3 of this paper for the HPP model, show that in general deterministic models are not suited for describing discrete velocity Boltzmann equations. In numerical experiments, too, the presence of the so-called spurious invariants seems to affect the validity of some computer simulations, while the addition of a few random updatings improves, at least in some cases, the outcome of the experiments, as noticed by S. Chen.

Randomness is not only necessary, but also sufficient, as shown by Caprino *et al.*,[2] where a stochastic version of the Broadwell gas is proven to converge to the Broadwell equation, without the short-times limitation present in the Lanford approach. It is therefore of interest to understand the true origin of such a result and in particular the role played by the stochasticity present in the evolution. Randomness may arise from a small noise added to the free motion of the particles and/or by randomizing their collisions, e.g., with probability p, any two particles entering a collision do in fact collide, while, with complementary probability, they do not. The Boltzmann–Grad limit can then be implemented by letting $p \to 0$ in such a way that the mean free path remains finite. In ref. 2 both sources of randomness are present and used in an essential way. The main object of this paper is the analysis of a stochastic version of the HPP cellular automaton,[6] where the collisions are deterministic, but the motion between collisions is random; see the next section for precise definitions. To have the mean free path finite we consider low densities, i.e., a vanishingly small initial density profile. We then prove convergence to the solution to the Broadwell equation up to the first explosion time, if any. Our results therefore cover several interesting cases, as in the analysis of the hydrodynamic limit of the Broadwell equation.[3,4]

While convergence at short times comes from the Lanford approach, convergence at longer times requires, at least here, a drastically different analysis. With respect to ref. 2 we miss an important step. By using Lanford's perturbative scheme, in ref. 2 it was possible to study the so-called BBGKY equations also for initial individual configurations and, when $\varepsilon \to 0$ (i.e., in the Boltzmann–Grad limit), for times which become infinitely long in microscopic time units. Such a result uses heavily that collisions have vanishing probability. The local occupation numbers may be large, even unbounded, due to local fluctuations, yet the number of times a particle has to be in the condition of colliding before this really happens diverges as $\varepsilon \to 0$. Consequently, the mean free time goes to infinity in microscopic units even though it might be macroscopically infinitesimal, due to local density fluctuations. In our case the occupation numbers in an initial lattice configuration are bounded, due to the exclusion rule in the HPP model, but since the collisions are deterministic, the mean free time is the time it takes for the first collision. We should therefore exploit that in typical configurations particles are far away from each other, but in doing this we miss the possibility of using sup norms as by Lanford. We could not avoid this problem, which is in fact the crucial problem in this paper. We have solved it by studying the dynamics iteratively in a sequence of different space-time scales. Only after a good control on a certain scale are we able to go to the next one, closer to the macroscopic scale. This multiscale analysis of the space-time process is carried through by means of techniques somewhat reminiscent of the renormalization group theory, as discussed in Section 4.

Another way to make the model random is to have stochastic collisions while keeping the free motion deterministic. It is not very difficult to prove that in this latter case there is, in the Boltzmann–Grad limit, global convergence to a kinetic equation which has bounded solutions at all times, due to the boundedness of the occupation numbers in the HPP model.

Deterministic collisions but random free motions were considered also by Lang and Nguyen[9] in a system of independent Brownian spheres which are removed when colliding, i.e., whenever any two of them are at a distance d, they both disappear. Letting $d \to 0$ as the number of particles suitably diverges so that the conditions of the Boltzmann–Grad limit are attained, propagation of chaos and global convergence to a kinetic-like equation are proven in ref. 9. The difficulties in this case are complementary to those met in the present paper, since in ref. 9 it is possible to control the growth of the correlation functions because the configurations at any time are random subsets of those without collisions. Both our approach and that in ref. 9 are related to the correlation function techniques, but ours, like that in ref. 2, focuses essentially on the evolution of the v-functions, sort of truncated correlation functions. We refer to ref. 12 for a survey on the correlation function techniques in the framework of stochastic systems and to ref. 5 for examples of applications of the v-functions.

In Section 2 we define the HPP model and its various stochastic versions, then we state the main results of this paper. In Section 3 we report the proofs relative to the deterministic case and to the case where the collisions are random. In Section 4 we outline the strategy used to study the model with deterministic collisions and random free motion. The proofs for this case are reported in Sections 5 and 6 and in a short Appendix.

2. THE MODEL AND THE RESULTS

Definition: The HPP Model. The one-particle phase space is $\mathbb{Z}^2 \times \mathscr{V}$, where $\mathscr{V}$ is the "velocity space," i.e.,

$$\mathscr{V} = \{c_1, c_2, c_3, c_4\}, \qquad c_1 = (1, 0); \quad c_2 = (0, 1), \quad c_3 = -c_1, \quad c_4 = -c_2 \tag{2.1}$$

An element in this space is denoted by (q, e). The many-particle phase space is $\{0, 1\}^{\mathbb{Z}^2 \times \mathscr{V}}$; its elements are the particle configurations, denoted by $\eta = \{\eta(q, e), (q, e) \in \mathbb{Z}^2 \times \mathscr{V}\}$, $\eta(q, e) = 0, 1$ being the occupation number at (q, e). The variable $\eta(q, e, t)$ is the occupation number in (q, e) at time $t \in \mathbb{N}$, where the evolution is given by the following deterministic updating rule, consisting of two successive subupdatings: call η the initial configuration, η' that after the first subupdating, and let η'' be the final one. Then for all (q, e)

$$\begin{aligned} \eta'(q, e) = {} & \eta(q, e)([1 - \eta(q, -e)] + \eta(q, -e)\{\eta(q, e^\perp) \\ & + [1 - \eta(q, e^\perp)]\, \eta(q, -e^\perp)\}) \\ & + [1 - \eta(q, e)][1 - \eta(q, -e)]\, \eta(q, e^\perp)\, \eta(q, -e^\perp) \end{aligned} \tag{2.2a}$$

Namely, there is a change at a site if and only if there are two particles which collide. This happens if and only if there are just two particles at that site with opposite velocities. After the collision each velocity changes by a clockwise rotation of $\pi/2$, $e \to e^\perp$. Equation (2.2a) can also be written as

$$\begin{aligned} \eta'(q, e) = {} & \eta(q, e) + \eta(q, e^\perp)\, \eta(q, -e^\perp)[1 - \eta(q, e)][1 - \eta(q', -e)] \\ & - \eta(q, e)\, \eta(q', -e)[1 - \eta(q, e^\perp)][1 - \eta(q, -e^\perp)] \end{aligned} \tag{2.2b}$$

Finally,

$$\eta''(q, e) = \eta'(q - e, e) \tag{2.2c}$$

namely η'' is obtained from η' by streaming. Time increases by one unit after the updating from η to η''.

The low-density limit is determined by an initial probability measure μ^ε which is a product measure on $\{0, 1\}^{\mathbb{Z}^2 \times \mathscr{V}}$ with averages

$$\mathbb{E}_{\mu^\varepsilon}(\eta(q, e)) = \varepsilon\rho(\varepsilon q, e) \tag{2.3}$$

where $\rho(r, e)$ is a bounded function on $\mathbb{R}^2 \times \mathscr{V}$.[3] We shall also assume that its derivatives exist and are bounded.

It is easily seen from (2.2) that for any $r \in \mathbb{R}$ (and denoting by $[r]$ the integer part of r), the limit of the following discrete-time derivative exists:

$$\lim_{\varepsilon \to 0} \varepsilon^{-2}[E_{\mu^\varepsilon}(\eta([\varepsilon^{-1}r], e, 1)) - \eta([\varepsilon^{-1}r], e, 0)]$$

(one ε comes from renormalizing times, the other one from renormalizing the densities). The above limit is equal to the time derivative at $(r, 0)$ of the solution of the Broadwell equation

$$\begin{aligned} \partial_t f_t(r, e) + e \cdot \nabla f_t(r, e) &= \mathscr{C} f_t(r, e) \\ f_0(r, e) &= \rho(r, e) \end{aligned} \tag{2.4a}$$

where

$$\begin{aligned} \nabla f(r, e) &= \left(\frac{\partial f(r, e)}{\partial r_1}, \frac{\partial f(r, e)}{\partial r_2} \right) \\ \mathscr{C} f(r, e) &= f(r, e^\perp) f(r, -e^\perp) - f(r, e) f(r, -e) \end{aligned} \tag{2.4b}$$

From this one might conjecture that the limiting behavior of the model is ruled by (2.4), but this is not what happens:

Theorem 2.1 (The deterministic HPP model). There exist τ^* and an L_∞ function $\rho(r, e, \tau)$, $\tau < \tau^*$, for which the following holds. Let $\phi(r, e)$ be any smooth function with compact support on $\mathbb{R}^2 \times \mathscr{V}$. Then for any positive δ and any $\tau < \tau^*$

$$\lim_{\varepsilon \to 0} \mathbb{P}_{\mu^\varepsilon} \left(\left| X_\tau^\varepsilon(\phi) - \sum_{e \in \mathscr{V}} \int_{\mathbb{R}^2} dr\, \phi(r, e)\, \rho(r, e, \tau) \right| > \delta \right) = 0 \tag{2.5a}$$

where the renormalized density field X_τ^ε is defined as

$$X_\tau^\varepsilon(\phi) = \varepsilon \sum_{e \in \mathscr{V}} \sum_{q \in \mathbb{Z}^2} \phi(\varepsilon q, e)\, \eta(q, e, \varepsilon^{-1}\tau) \tag{2.5b}$$

Furthermore, the function $\rho(r, e, \tau)$ does not satisfy the Broadwell equation.

This is just the Uchiyama theorem in the context of the HPP model; also the proof is essentially that of ref. 13, except for the fact that in HPP there are three-body collisions. A careful reading of Lanford's paper shows, however, that in this specific case no real problem arises. As this might not seem so obvious, we shall give some details in Section 3.

[3] More general initial measures may also be considered, but we shall not discuss this point.

From Theorem 2.1 it follows that if we want limiting equations of kinetic type, we have to add some randomness; one possibility is to make the collisions random.

Definition: The HPP Model with Random Collisions. The updating from η to η' and η'' [see (2.2)] is modified as follows. Given $\varepsilon > 0$, for each (q, e) we choose independently a number $\lambda_{(q,e)}$ equal to 1 with probability ε and to 0 with complementary probability. Then $\eta'(q, e)$ is equal to the expression in (2.2a) if $\lambda_{(q,e)} = 1$, otherwise $\eta'(q, e) = \eta(q, e)$. The η'' is obtained from η' as in (2.2c). The random variables $\lambda_{(q,e)}$ for different updatings are mutually independent.

We shall assume that μ^ε is still a product measure, but with averages

$$\mathbb{E}_{\mu^\varepsilon}(\eta(q, e)) = \rho(\varepsilon q, e) \tag{2.6}$$

where $0 \leqslant \rho \leqslant 1$ is a given smooth function, as before.

Definition: The Correlation Functions. These are the functions

$$u_n^\varepsilon(\underline{x}, t) = \mathbb{E}_{\mu^\varepsilon}(\eta(\underline{x}, t)) \tag{2.7a}$$

where $\underline{x} = (x_1, \ldots, x_n)$, $x_i = (q_i, e_i)$, are n different single-particle states, and

$$\eta(\underline{x}, t) = \prod_{i=1}^{n} \eta(x_i, t) \tag{2.7b}$$

We have the following:

Theorem 2.2 (The HPP model with random collisions). For any $\tau \geqslant 0$ and any $n \geqslant 1$

$$\lim_{\varepsilon \to 0} \sup_{\underline{x}} \left| u_n^\varepsilon(\underline{x}, \varepsilon^{-1}\tau) - \prod_{i=1}^{n} \rho(\varepsilon q_i, e_i, \tau) \right| = 0 \tag{2.8}$$

where the sup is over all n-tuples of different phase points $\underline{x} = (x_1, \ldots, x_n)$ and $x_i = (q_i, e_i)$. The function $\rho(r, e, t)$ is the solution of

$$\frac{\partial \rho(r, e, t)}{\partial t} = -e \cdot \nabla \rho + \mathscr{C}_1 \rho \tag{2.9a}$$

with initial datum $\rho(r, e)$, where

$$\begin{aligned}(\mathscr{C}_1 \rho)(q, e) = {} & \rho(q, e^\perp)\, \rho(q, -e^\perp)(1 - \rho(q, e))(1 - \rho(q, -e)) \\ & - \rho(q, e)\, \rho(q, -e)(1 - \rho(q, e^\perp))(1 - \rho(q, -e^\perp)) \end{aligned} \tag{2.9b}$$

Given $e \in \mathscr{V}$, we have denoted by $e^{\perp}$ the velocity obtained from e by a clockwise rotation of $\pi/2$.

We sketch the proof of this theorem in the next section, essentially for the sake of completeness, since the proof, based on the Lanford analysis, is quite standard. The trouble coming from recollisions, which were at the origin of the Uchiyama paradox, is avoided because the probability that two given particles have a collision is vanishingly small when $\varepsilon \to 0$. The global validity of the result is a consequence of the trivial uniform bound on the correlation functions; since the occupation numbers are 0 or 1, the density is bounded by 1.

We now turn to the model in which we are really interested: *the HPP model in a stirred environment.*

Definition: The Phase Space. The one-particle phase space is

$$\Gamma = \mathbb{Z}^2 \times \mathscr{V} \times S, \qquad S = \{1, 2, 3, 4\} \tag{2.10}$$

where $\mathscr{V}$ is defined in (2.1). An element $x = (q, e, \sigma) \in \Gamma$ is a one-particle state, q denotes the position in this state, e the velocity (hereafter called e-velocity), and σ the σ-velocity. This is the HPP one-particle phase space with an extra variable added, the σ-velocity.

The particle configuration space is

$$\Omega = \{0, 1\}^{\Gamma} \tag{2.11}$$

and we denote by $\eta = \{\eta(x), x \in \Gamma\}$ an element of Ω, $\eta(x)$ being the occupation number at x.

Definition: The Evolution. We fix a number $\nu \in (0, 1)$; then for $\varepsilon \in (0, 1]$ and such that $\varepsilon^{-\nu}$ is an integer,[4] we define the following updating rules. The updating rule at the times $k\varepsilon^{-\nu}$, $k \geqslant 0$, is given by the deterministic HPP rule for each value of σ. At $t \neq k\varepsilon^{-\nu}$, $k \geqslant 1$, the updating consists of two steps: denote by η the configuration before the updating, by η' that after the first step, and by η'' the final one. For all $(q, e, \sigma) \in \Gamma$ we then set $\hat{x} = (q, e)$ and

$$\eta'(q, e, \sigma) = \eta(q, e, \sigma - \sigma_{\hat{x}}) \tag{2.12}$$

where $\sigma - \sigma_{\hat{x}}$ is defined modulo 4 and the $\sigma_{\hat{x}}$ are i.i.d. variables with equiprobable values in S. The random variables relative to different updatings are mutually independent. We then set

$$\eta''(q, e, \sigma) = \eta'(q - c_{\sigma}, e, \sigma) \tag{2.13}$$

[4] We could consider as well the general case by taking the integer part of $\varepsilon^{-\nu}$; we did not do this, to have lighter notation.

Remarks. The stirring, also known as the symmetric simple exclusion process, is a slightly different process.[10] In the present form, it was introduced by Boghosian and Levermore[1] to simulate the Burgers equation (they considered for this purpose an "asymmetric" one-dimensional version of the model). The idea of inserting stirring updatings in between HPP (or more general cellular automata) updatings was emploied first in ref. 4 as a tool for proving convergence in the hydrodynamic limit. It might have also practical effects, as noticed recently by S. Chen, who found improved accuracy in viscosity measurements by computer simulations when adding stirring updatings.

In a stirring updating the e-velocity does not play any role; for each value of e we have an independent copy of the same system. The effect of the stirring updatings once the $\sigma_{\hat{x}}$ are given is just to exchange the contents of the states of Γ; hence, between an HPP updating and the successive one, the intermediate stirring updatings exchange stochastically the content of the single-particle states: this is the reason why we call our model an HPP model in a stirred environment.

We still need two definitions before stating our results on the present model.

Definition: The ρ Functions. A ρ function is a function on $\Gamma \times \mathbb{N}$ which satisfies the following two conditions:

1. For all $t \geqslant 0$, $t \neq k\varepsilon^{-\nu}$, and all x

$$\rho(x, t+1) = \sum_y P_{t+1,t}(x \to y)\, \rho(y, t) \tag{2.14}$$

where

$$P_{t+1,t}((q, e, \sigma) \to (q', e', \sigma')) = \begin{cases} 1/4 & \text{if } \; q' = q - c_\sigma, \quad e = e' \\ 0 & \text{otherwise} \end{cases} \tag{2.15}$$

[the same relation as (2.14) is obtained by averaging the expression giving η'' in terms of η in (2.13)].

2. For $t = k\varepsilon^{-\nu}$, $k \geqslant 1$,

$$\rho(x, t+1) = \rho(x', t) + \mathscr{C}_1 \rho(x', t) \tag{2.16}$$

where $x' = (q - e, e, \sigma)$ if $x = (q, e, \sigma)$ and $\mathscr{C}_1$ is, for each σ, defined as in (2.9).

Equation (2.16) is obtained by averaging (2.2) with a product measure. Given g on Γ and $s \geqslant 0$, we call $\rho(x, t \mid g, s)$, $t \geqslant s$, the solution of

(2.14) and (2.16) such that $\rho(x, s) = g(x)$. For $s = 0$ we simply write $\rho(x, t \mid g)$.

Definition: The v-Functions. Given $\eta \in \Omega$, $\underline{x} = (x_1, ..., x_n)$, the x_i being different states in Γ, we set

$$v_n(\underline{x}, t \mid \eta) = \mathbb{E}_\eta\left(\prod_{i=1}^{n} [\eta(x_i, t) - \rho(x_i, t \mid \eta)]\right) \tag{2.17a}$$

where $\mathbb{E}_\eta$ is the expectation with respect to the HPP process in a stirred environment starting from η. Analogously, given a product measure μ on Ω such that $g(x) = \mathbb{E}_\mu(\eta(x))$, we define

$$v_n(\underline{x}, t \mid \mu) = \mathbb{E}_\mu\left(\prod_{i=1}^{n} [\eta(x_i, t) - \rho(x_i, t \mid g)]\right) \tag{2.17b}$$

where $\mathbb{E}_\mu$ denotes the expectation starting from μ.

In Sections 4–6 we prove the following theorem.

Theorem 2.3 (The HPP model in a stirred environment). Let $\rho\colon \mathbb{R}^2 \times \mathscr{V} \to \mathbb{R}_+$ be a C^1-function bounded together with its derivative. For any $\varepsilon \in (0, 1)$ let μ^ε be the product measure in Ω such that for any (q, e, σ)

$$\mathbb{E}_{\mu^\varepsilon}(\eta(q, e, \sigma)) = \varepsilon\rho(\varepsilon q, e) \tag{2.18}$$

Let $T > 0$ be any time such that a unique bounded solution of the Broadwell equation (2.4) with initial datum ρ exists up to time T.

Then the following holds. There is $\delta > 0$ and for any $n \geq 1$ there is c so that for all $t \leq T$, for all sets of n distinct states $\underline{x} = (x_1, ..., x_n)$, $x_i = (q_i, e_i, \sigma_i)$, such that $|q_i| \leq \varepsilon^{-2}$ for all i,

$$|v_n(\underline{x}, [\varepsilon^{-1-\nu}t] \mid \mu^\varepsilon)| \leq c\varepsilon^{(1+\delta)n} \tag{2.19}$$

Furthermore, there is c so that for all $t \leq T$ and $x = (q, e, \sigma)$, $|q| \leq \varepsilon^{-2}$,

$$|\rho(q, e, \sigma, \varepsilon^{-1-\nu}t \mid \mu^\varepsilon(\cdot)) - \varepsilon f_t(\varepsilon q, e)| \leq c\varepsilon^{(1+\delta)} \tag{2.20}$$

where f_t solves the Broadwell equation (2.4) with initial datum ρ.

In particular, from (2.19) with $n = 1$ and (2.20) it follows that

$$\lim_{\varepsilon \to 0} \sup_{t \leq T} \sup_{(q,e,\sigma)} \varepsilon^{-1} |\mathbb{E}_{\mu^\varepsilon}(\eta(q, e, \sigma, [\varepsilon^{-1-\nu}t]) - \varepsilon f_t(\varepsilon q, e)| = 0$$

Equation (2.19) tells us that the finite distributions of the process at the time $\varepsilon^{-1-\nu}t$ are, when $\varepsilon \to 0$, close to those relative to a product measure with averages $\rho(\cdot \mid \mu^\varepsilon(\cdot))$. Then (2.20) shows that such averages are close to the solution of the Broadwell equation, multiplied by ε.

3. PROOF OF THEOREMS 2.1 AND 2.2

In this section we prove Theorems 2.1 and 2.2 by a straightforward extension of the Lanford technique.

Proof of Theorem 2.1. We define the *normalized correlation functions* as

$$\rho_n^\varepsilon(q_1, e_1,..., q_n, e_n; t) = \varepsilon^{-n}\mathbb{E}_{\mu^\varepsilon}\left(\prod_{i=1}^{n} \eta(q_i, e_i, t)\right) \tag{3.1}$$

We recall the notation introduced in Section 2: $x = (q, e)$, $\underline{x} = (x_1,..., x_n)$, $n \geqslant 1$, and we write $|\underline{x}| = n$. Moreover, we use $x^\perp = (q, e^\perp)$, $x^{-\perp} = (q, -e^\perp)$, $x^- = (q, -e)$, and $x' = (q - e, e)$. We want to derive a set of hierarchical equations for the functions $\rho_n(\underline{x})$. We have

$$\eta(q, e, t) = \sum_{\underline{y}} b(\underline{y}; (q-e, e))\, \eta(\underline{y}, t-1) \tag{3.2}$$

where

$$\eta(\underline{y}) = \prod_{i=1}^{|\underline{y}|} \eta(y_i)$$

and $b(\underline{y}, x) = 0$ if $|\underline{y}| > 3$. If $|\underline{y}| \leqslant 3$,

$$\begin{aligned} b(\underline{y}, x) = {}& \delta(|\underline{y}|, 1)\, \delta(y_1, x) \\ & + \delta(|\underline{y}|, 2)[\delta(y_1, x^\perp)\, \delta(y_2, x^{-\perp}) - \delta(y_1, x)\, \delta(y_2, x^-)] \\ & + \delta(|\underline{y}|, 3)[\delta(y_1, x)\, \delta(y_2, x^-)[\delta(y_3, x^\perp) + \delta(y_3, x^{-\perp})] \\ & - \delta(y_1, x^\perp)\, \delta(y_2, x^{-\perp})[\delta(y_3, x) + \delta(y_3, x^-)]] \end{aligned} \tag{3.3}$$

and $\delta(k, l)$ is the Kronecker delta. We have

$$\prod_{i=1}^{n} \eta(x_i, t) = \sum_{\underline{y}_1,..., \underline{y}_n} b(\underline{y}_1, x_1') \cdots b(\underline{y}_n, x_n') \prod_{i=1}^{n} \eta(\underline{y}_i, t-1) \tag{3.4}$$

Hence, denoting by ρ^ε the sequence $\{\rho_n^\varepsilon\}_{n=1}^\infty$, we have

$$\rho_n^\varepsilon(x_1,..., x_n; t) = (\mathscr{C}_\varepsilon \rho^\varepsilon)_n\, (x_1,..., x_n; t-1) \tag{3.5}$$

where, setting $\underline{Y} = \bigcup_{i=1}^n \underline{y}_i$,

$$\begin{aligned} &(\mathscr{C}_\varepsilon \rho^\varepsilon)_n\, (x_1,..., x_n) \\ &\quad = \sum_{\underline{y}_1,..., \underline{y}_n:\, |\underline{Y}| \geqslant n} b(\underline{y}_1, x_1') \cdots b(\underline{y}_n, x_n')\, \varepsilon^{|\underline{Y}|-n} \rho_{|\underline{Y}|}^\varepsilon(\underline{y}_1 \cup \cdots \cup \underline{y}_n) \end{aligned} \tag{3.6}$$

We note that some points may be in more than one of the sets $\underline{y}_1,..., \underline{y}_n$. Any such point has to be counted just once, since $\eta^2 = \eta$, and this is the reason why the set $\underline{y}_1 \cup \cdots \cup \underline{y}_n$ appears on the right-hand side of (3.6).

We now prove the following result.

Proposition 3.1. Suppose that $\sup_{r,v} \rho(r, v) < z_0$. Then there are $\tau^* > 0$ and $z > 0$ such that, if $\tau \leqslant \tau^*$, for any $k > 0$ we have, for all $\varepsilon \in (0, 1)$,

$$\sup_{(x_1,..., x_k) \in (\mathbb{Z}^2 \times \mathscr{V})^k} \rho_k^\varepsilon(x_1,..., x_k; \varepsilon^{-1}\tau) < z^k$$

Proof. The number of terms on the right-hand side of (3.5) can be bounded in the following way. Define the coefficients $a_{k,n}$ so that

$$c_\varepsilon^n = \sum_k a_{k,n} \varepsilon^k, \qquad c_\varepsilon \equiv 1 + 2\varepsilon + 4\varepsilon^2 \tag{3.7}$$

Then for each k, $a_{k,n}$ is an upper bound for the number of correlation functions of order $k + n$ present in (3.5). It is enough to verify this statement when all the particles are in the same position, thus with at most four particles; in fact, collisions at different positions are independent. We omit the details.

Therefore, defining

$$\rho^\varepsilon(n, t) = \sup_{(x_1,..., x_n) \in (\mathbb{Z}^2 \times \mathscr{V})^n} \rho_n^\varepsilon(x_1,..., x_n, t) \tag{3.8}$$

we have

$$\begin{aligned} \rho^\varepsilon(n, t) &\leqslant \sum_k a_{k,n} \varepsilon^k \rho^\varepsilon(n + k, t - 1) \\ &= c_\varepsilon^n \sum_k \frac{a_{k,n}\varepsilon^k}{c_\varepsilon^n} \rho^\varepsilon(n + k, t - 1) \end{aligned} \tag{3.9a}$$

In this way we can interpret the coefficients $a_{k,n}\varepsilon^k/c_\varepsilon^n$ as probabilities. They are in fact, as can be easily checked, the probabilities associated with a branching process, where before the branching, there are n particles. Then each one independently of the others may give birth to a new particle with probability $2\varepsilon/c_\varepsilon$. Another possible event is the creation of two new particles; this happens with probability $4\varepsilon^2/c_\varepsilon$. Finally, with complementary probability $1/c_\varepsilon$, no new particle is created. We denote by $\mathbb{E}_n$ the expectation with respect to this branching process starting with n particles. We can then rewrite (3.9a) as

$$\rho^\varepsilon(n, t) \leqslant \mathbb{E}_n[\rho^\varepsilon(N_1, t - 1)\, c_\varepsilon^n] \tag{3.9b}$$

where N_s, $s \geqslant 0$, denotes the number of particles existing at the sth step of the branching process. Iterating (3.9), we have

$$\rho^\varepsilon(n, t) \leqslant \mathbb{E}_n[\rho^\varepsilon(N_t, 0)\, c_\varepsilon^{n+N_1+\cdots+N_{t-1}}] \tag{3.10}$$

Since $\rho^\varepsilon(n, 0) \leqslant z_0^n$, (3.10) becomes

$$\rho^\varepsilon(n, t) \leqslant \mathbb{E}_n[z_0^{N_t} c_\varepsilon^{n+N_1+\cdots+N_{t-1}}] \tag{3.11}$$

By the Markov property we have

$$\rho^\varepsilon(n, t) \leqslant \mathbb{E}_n[c_\varepsilon^{n+N_1+\cdots+N_{t-1}} \mathbb{E}_n(z_0^{N_t} \mid N_{t-1})] \tag{3.12}$$

where $\mathbb{E}_n(\cdot \mid N_{t-1})$ denotes conditional expectation. Given that $N_{t-1} = k$, $N_t = m_1 + \cdots + m_k$, where m_l is the number of particles branching from the lth particle. Since these variables are independent, we have

$$\mathbb{E}_n(z_0^{N_t} \mid N_{t-1}) = (\mathscr{E}[z_0^m])^{N_{t-1}} \tag{3.13}$$

The law of the variable m is the following:

$$\mathscr{P}(m) = c_\varepsilon^{-1} \begin{cases} 1 & \text{if} \quad m = 1 \\ 2\varepsilon & \text{if} \quad m = 2 \\ 4\varepsilon^2 & \text{if} \quad m = 3 \end{cases} \tag{3.14}$$

Therefore

$$\mathscr{E}[z_0^m] = c_\varepsilon^{-1}[z_0 + 2\varepsilon z_0^2 + 4\varepsilon^2 z_0^3] \equiv c_\varepsilon^{-1} z_1 \tag{3.15}$$

By conditioning on the value of N_{t-2}, (3.12) becomes

$$\rho^\varepsilon(n, t) \leqslant \mathbb{E}_n[c_\varepsilon^{n+N_1+\cdots+N_{t-2}} \mathbb{E}_n(z_1^{N_{t-1}} \mid N_{t-2})] \tag{3.16}$$

Iterating the above procedure,. we define a sequence $\{z_s\}_{s=1,\ldots,t}$ with z_1 given by (3.15) and

$$z_k = z_{k-1}[1 + 2\varepsilon z_{k-1} + 4\varepsilon^2 z_{k-1}^2] \tag{3.17}$$

In terms of this sequence we have

$$\rho^\varepsilon(n, t) \leqslant z_t^n \tag{3.18}$$

Now we bound z_t by induction. We assume that there is a K such that $z_l < K$ for $l \leqslant s-1$ and prove that $z_s < K$ if $s \leqslant t$ with $t = [\tau\varepsilon^{-1}]$ and $\tau \leqslant \tau^*$

for τ^* small enough. In fact, letting $\lambda_k = \log z_k$,[5] since $\log z_k \leqslant \log z_{k-1} + \log(1+2\varepsilon z_{k-1})^2$, we get

$$\lambda_s < \lambda_{s-1} + 4\varepsilon K < \lambda_0 + 4\varepsilon s K < \lambda_0 + 4\tau^* K \tag{3.19}$$

If $K = z_0 e$ and $\tau^* < (4K)^{-1}$, $\lambda_0 + 4\tau^* K < \lambda_0 + 1 \equiv \log K$, i.e., $z_s < K$. Hence, by (3.18), Proposition 3.1 follows by taking $z = K$. ∎

We now introduce the decomposition of the operator $\mathscr{C}_\varepsilon$ in terms of the operators

$$(\mathscr{S}_\varepsilon \rho^\varepsilon)_n (x_1,..., x_n) = \sum_{\underline{y}_1,..., \underline{y}_n:\, |\underline{Y}| = n} b(\underline{y}_1, x_1') \cdots b(\underline{y}_n, x_n') \times \rho_n^\varepsilon(\underline{y}_1 \cup \cdots \cup \underline{y}_n) \tag{3.20}$$

$$(\mathscr{Q}_\varepsilon \rho^\varepsilon)_n (x_1,..., x_n) = \sum_{\underline{y}_1,..., \underline{y}_n:\, |\underline{Y}| = n+1} b(\underline{y}_1, x_1') \cdots b(\underline{y}_n, x_n') \times \rho_{n+1}^\varepsilon(\underline{y}_1 \cup \cdots \cup \underline{y}_n) \tag{3.21}$$

$$(\mathscr{R}_\varepsilon \rho^\varepsilon)_n (x_1,..., x_n) = \sum_{\underline{y}_1,..., \underline{y}_n:\, |\underline{Y}| > n+1} b(\underline{y}_1, x_1') \cdots b(\underline{y}_n, x_n') \times \varepsilon^{|\underline{Y}| - n - 2} \rho_{|\underline{Y}|}^\varepsilon(\underline{y}_1 \cup \cdots \cup \underline{y}_n) \tag{3.22}$$

Given $\underline{x}$, we call internal collisions with respect to $\underline{x}$ all the collisions which involve only particles in $\underline{x}$ and external collisions the other ones. The operator $\mathscr{S}_\varepsilon$ takes into account all the internal collisions, whereas the operators $\mathscr{Q}_\varepsilon$ and $\mathscr{R}_\varepsilon$ depend on the external collisions, and involve higher-order correlation functions, with some power of ε in front. Writing (3.5) as

$$\rho_n^\varepsilon(\underline{x}; t) = (\mathscr{S}_\varepsilon \rho^\varepsilon)_n (\underline{x}; t-1) + \varepsilon(\mathscr{Q}_\varepsilon \rho^3)_n (\underline{x}; t-1) + \varepsilon^2 (\mathscr{R}_\varepsilon \rho^\varepsilon)_n (\underline{x}; t-1) \tag{3.23}$$

and iterating it, we get the expansion

$$\rho_n^\varepsilon(\underline{x}; t) = \tilde{\rho}_n^\varepsilon(\underline{x}; t) + \bar{\rho}_n^\varepsilon(\underline{x}; t) \tag{3.24}$$

where

$$\tilde{\rho}_n^\varepsilon(\underline{x}; t) = \sum_{N \geqslant 0} \varepsilon^N \sum_{t_1 = 0}^{t-1} \cdots \sum_{t_N = 0}^{t_{N-1} - 1} (\mathscr{S}_\varepsilon^{(t-t_1)} \mathscr{Q}_\varepsilon \cdots \mathscr{S}_\varepsilon^{(t_{N-1} - t_N - 1)} \mathscr{Q}_\varepsilon \mathscr{S}_\varepsilon^{t_N} \rho^\varepsilon)_n (\underline{x}; 0) \tag{3.25}$$

[5] We assume $z_0 \geqslant 1$.

$$\bar{\rho}_n^{\varepsilon}(\underline{x}; t) = \sum_{N \geqslant 1} \varepsilon^{N+1} \sum_{t_1=0}^{t-1} \cdots \sum_{t_N=0}^{t_{N-1}-1} (\mathscr{S}_{\varepsilon}^{(t-t_1)} \mathscr{Q}_{\varepsilon} \cdots \mathscr{S}_{\varepsilon}^{(t_{N-1}-t_N-1)} \mathscr{R}_{\varepsilon} \rho^{\varepsilon})_n (\underline{x}; t_N) \tag{3.26}$$

and $\mathscr{S}_{\varepsilon}^m$ denotes the mth iteration of the operator $\mathscr{S}_{\varepsilon}$.

By Proposition 3.1, $\rho_k^{\varepsilon}(\cdot; t_N)$ is bounded by z^k, so that, if $\varepsilon < z^{-1}$, by (3.22) it follows that for any positive m

$$\begin{aligned}(\mathscr{R}_{\varepsilon} \rho^{\varepsilon})_m &\leqslant \varepsilon^{-2} z^m \sum_{n_1 + \cdots + n_m \geqslant 2} \prod_i (2\varepsilon z)^{n_i} \\ &\leqslant 4z^{m+2} \sum_{n_1 + \cdots + n_m \geqslant 0} \prod_i (1 + 2\varepsilon z)^{n_i} \leqslant 4z^{m+2} 13^m\end{aligned}$$

since, if $n_1 + \cdots + n_m \geqslant 2$, then $\prod_{i=1}^m (2\varepsilon z)^{n_i} \leqslant (2\varepsilon z)^2 \prod_{i=1}^m (1 + 2\varepsilon z)^{n_i}$. Since the $N-1$ operators $\mathscr{Q}_{\varepsilon}$ that appear in (3.26) create $N-1$ particles, we use the bound

$$(\mathscr{R}_{\varepsilon} \rho^{\varepsilon})_{n+N-1} \leqslant 4z^{n+N+1} 13^{N-1+n} \tag{3.27}$$

By (3.20) and (3.21) it follows that for any n

$$\sup \mathscr{S}_{\varepsilon} \rho_n^{\varepsilon} \leqslant \sup \rho_n^{\varepsilon} \tag{3.28}$$

$$\sup (\mathscr{Q}_{\varepsilon} \rho^{\varepsilon})_n \leqslant 4n \sup \rho_{n+1}^{\varepsilon} \tag{3.29}$$

Hence, for $t = [\varepsilon^{-1} \tau]$, taking advantage of the time ordering in the sum on $t_1, \ldots, t_N$ and using the bound $(N+n)!/n! \leqslant 2^{N+n} N!$, we get

$$\begin{aligned}\bar{\rho}_n^{\varepsilon}(\underline{x}; [\varepsilon^{-1}\tau]) &\leqslant \sum_{N \geqslant 1} \varepsilon^{N+1} \sum_{t_1=0}^{t-1} \cdots \sum_{t_N=0}^{t_{N-1}-1} n \cdot (n+1) \cdot \cdots \cdot (n+N-1) \\ &\quad \times 4^N z^{n+N+1} 13^{N-1+n} \\ &\leqslant (13z)^{n+1} \varepsilon^2 \sum_{N \geqslant 0} (104 z\tau)^N\end{aligned} \tag{3.30}$$

For $\tau < (104z)^{-1}$ the above sum converges and we conclude that ρ_n^{ε} differs from $\tilde{\rho}_n^{\varepsilon}$ by $C\varepsilon z^{n+1}$. Moreover, the same argument also shows that the series (3.25) is dominated by a convergent series. Since each term of the series in (3.25) has a limit, we conclude that the normalized correlation functions $\rho_k^{\varepsilon}([\varepsilon^{-1} q_1], e_1, \ldots, [\varepsilon^{-1} q_k], e_k; [\varepsilon^{-1}\tau])$ converge to some almost everywhere continuous functions $\rho_k(q_1, e_1, \ldots, q_k, e_k; \tau)$ for $\tau \leqslant \tau_0 < (104z)^{-1}$. Starting from time τ_0, the argument can be iterated up to the time τ^* defined in the proof of Proposition 3.1, showing the convergence up to this time.

The arguments of Uchiyama[13] can be repeated to show that if the n-body correlation functions at time zero are products, the limit correlation functions at time $t < \tau^*$ have this property with the exception of a set of zero measure. This, together with the law of large numbers, implies the relation (2.5) (see, for example, ref. 11, Sections 2.5 and 4.7).

To complete the proof of Theorem 2.1, we have to prove that ρ_1^ε does not converge to a solution of (2.4). This is related to the fact that the set of times and positions such that internal collisions are possible in the backward collision histories has nonvanishing measure in the limit $\varepsilon \to 0$. To prove this, we follow closely the Uchiyama argument, which compares the function $\rho_1(r, e, \tau)$ with the solution $f(r, v, \tau)$ of Eq. (2.4) with initial datum $\rho(r, e)$. Actually, we prove that

$$\lim_{\tau \to 0} \frac{1}{\tau^3} [f(r, e, \tau) - \rho_1(r, e, \tau)] = c[\rho(q, e^\perp)^2 \rho(q, -e^\perp)^2 - \rho(q, e)^2 \rho(q, -e)^2] \tag{3.31}$$

and the right-hand side of (3.31) vanishes only for initial data that are "local Maxwellians," i.e., distribution functions of the form $\exp[\alpha + \underline{\beta} \cdot e]$, $\alpha \in \mathbb{R}$, and $\underline{\beta} \in \mathbb{R}^2$. If the initial distribution is not a local Maxwellian, (3.31) shows that the limit density $\rho_1(r, e, \tau)$ does not solve (2.4). On the other hand, local Maxwellians are not generically solutions of (2.4). Namely, if a local Maxwellian [with $\alpha = \alpha(r, \tau)$ and $\underline{\beta} = \underline{\beta}(r, \tau)$] is a solution to (2.4), then necessarily $\partial_{r_1} \beta_1 = \partial_{r_2} \beta_2$ at time $t = 0$. In fact, on Maxwellians, (2.4) reduces to

$$\frac{\partial f(r, e, \tau)}{\partial \tau} = -e \cdot \nabla_r f(r, e, \tau) \tag{3.32}$$

Then $\rho(r, e, \tau) = \rho(r - e\tau, e, 0)$, and in terms of α and $\underline{\beta}$, $\alpha(r, \tau) + e \cdot \underline{\beta}(r, \tau) = \alpha(r - e\tau, 0) + e \cdot \underline{\beta}(r - e\tau, 0)$. Differentiating with respect to τ for $\tau = 0$, we get

$$\frac{\partial \alpha(r, 0)}{\partial \tau} + e \cdot \frac{\partial \underline{\beta}(r, 0)}{\partial \tau} = -e \cdot \nabla_r \alpha(r, 0) - e \cdot \nabla_r \underline{\beta}(r, 0) \cdot e \tag{3.33}$$

Summing the above relations for $e = c_1$ and $e = c_3$, we get

$$\frac{\partial \alpha(r, 0)}{\partial \tau} = -\frac{\partial \beta_1(r, 0)}{\partial r_1} \tag{3.34}$$

while summing them for $e = c_2$ and $e = c_4$, we have

$$\frac{\partial \alpha(r, 0)}{\partial \tau} = -\frac{\partial \beta_2(r, 0)}{\partial r_2} \tag{3.35}$$

Therefore (3.31) can be applied at any time $t_0 > 0$ to prove that for *generic* initial data $\rho_1^\varepsilon([\varepsilon^{-1}r], e, [\varepsilon^{-1}\tau])$ does not converge to the solution $f(r, e, t)$ of (2.4). ▮

Proof of (3.31). We simply sketch the argument. To compare f and ρ_1, it is useful to give a representation of f in terms of a sum over collision histories like ρ_1. We have

$$f(r, e, \tau) = \sum_{N \geqslant 0} \int_0^{\varepsilon^{-1}\tau} dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{N-1}} dt_N\, \varepsilon^N \\ \times \sum_{i_1=0}^{1} \sum_{\sigma_1=-,+} \cdots \sum_{i_N=0}^{N} \sum_{\sigma_N=-,+} \\ \times (\mathscr{S}_{\varepsilon^{-1}\tau - t_1} \mathscr{Q}_{i_1,\sigma_1} \mathscr{S}_{t_1 - t_2} \mathscr{Q}_{i_2,\sigma_2} \cdots \mathscr{S}_{t_{N-2} - t_{N-1}} \mathscr{Q}_{i_N,\sigma_N} \mathscr{S}_{t_N} f) \\ \times (q_1, v_1; 0) \tag{3.36}$$

with $(q_1, v_1) = ([\varepsilon^{-1}r], e)$, $f_n \equiv f^{\otimes n}$, and f denotes the sequence $\{f_n\}_{n=1}^{\infty}$. Moreover,

$$(\mathscr{S}_t f)_m (q_1, e_1,..., q_m, e_m) \\ = f_m(q_1 - e_1 t, e_1,..., q_m - e_m t, e_m) \tag{3.37}$$

$$(\mathscr{Q}_{i,+} f)_{m+1} (q_1, e_1,..., q_m, e_m) \\ = f_{m+1}(q_1, e_1,..., q_{i-1}, e_{i-1}, q_i, e_i^{\perp},..., q_m, e_m, q_i, -e_i^{\perp}) \tag{3.38}$$

$$(\mathscr{Q}_{i,-} f)_{m+1} (q_1, e_1,..., q_m, e_m) \\ = -f_{m+1}(q_1, e_1,..., q_i, e_i,..., q_m, e_m, q_i, -e_i) \tag{3.39}$$

We notice that the dependence on ε in (3.36) is fictitious and it has been introduced in order to compare (3.36) with the right-hand side of (3.25) with $n=1$. Since ρ_1^ε reduces to $\tilde{\rho}_1^\varepsilon$ in the limit $\varepsilon \to 0$ by (3.24) and (3.30), if recollisions could be ignored, as in the continuous velocities case, the convergence of ρ_1 to f would be achieved. In our case, it is easy to realize that the terms in (3.25) with $N < 3$ converge to the corresponding ones in (3.36). However, there are terms with $N = 3$ that do not converge to the corresponding terms in (3.36). They are

$$\sum_{t_1=0}^{\varepsilon^{-1}\tau} \sum_{t_2=0}^{t_1-1} \sum_{t_3=0}^{t_2-1} \varepsilon^3 (\mathscr{S}_\varepsilon^{(\varepsilon^{-1}\tau - t_1)} [\mathscr{Q}_{1,-} \mathscr{S}_\varepsilon^{(t_1-t_2)} \mathscr{Q}_{1,-} \mathscr{S}_\varepsilon^{(t_2-t_3)} \mathscr{Q}_{2,-} \\ + \mathscr{Q}_{1,-} \mathscr{S}_\varepsilon^{(t_1-t_2)} \mathscr{Q}_{2,-} \mathscr{S}_\varepsilon^{(t_2-t_3)} \mathscr{Q}_{1,-} + \mathscr{Q}_{1,+} \mathscr{S}_\varepsilon^{(t_1-t_2)} \mathscr{Q}_{1,-} \mathscr{S}_\varepsilon^{(t_2-t_3)} \mathscr{Q}_{2,-} \\ + \mathscr{Q}_{1,+} \mathscr{S}_\varepsilon^{(t_1-t_2)} \mathscr{Q}_{2,-} \mathscr{S}_\varepsilon^{(t_2-t_3)} \mathscr{Q}_{1,-}] \mathscr{S}_\varepsilon^{(t_3)} \rho^\varepsilon)_1 (q_1, e_1, 0) \tag{3.40}$$

The corresponding terms in (3.36) are given by the same expression with $\mathcal{S}_\varepsilon^{(t_i - t_{i+1})}$ substituted by $\mathcal{S}_{t_i - t_{i+1}}$. The main point is that each of the terms in (3.40) refers to situations such that for a proper choice of t_1, t_2, and t_3 an internal collision may happen in the interval $[\varepsilon^{-1}\tau - t_1, \varepsilon^{-1}\tau]$, and actually happens because the dynamics $\mathcal{S}_\varepsilon$ is deterministic and all the possible collision updatings are performed.

Before the application of $\mathcal{S}_\varepsilon^{(\varepsilon^{-1}\tau - t_1)}$, the velocity configurations corresponding to the four terms in (3.40), regardless of the positions, are the following:

$$(e, -e, -e, e), \quad (e, -e, e, -e), \quad (e^\perp, -e^\perp, -e^\perp, e^\perp), \quad (e^\perp, -e^\perp, e^\perp, -e^\perp)$$

where the first velocity is that of the original particle, while the others are added according to the operators $\mathcal{Q}_{i,\sigma}$. For a proper choice of the times the second couple of particles undergoes a collision at a time in the interval $[\varepsilon^{-1}\tau - t_1, \varepsilon^{-1}\tau]$, making all the velocity configurations the same up to a permutation. Therefore, in the limit $\tau \to 0$, the terms in (3.40) compensate each other and the expression (3.40) vanishes. The same is not true for the corresponding terms in (3.36). In this case the application of $\mathcal{S}_{\varepsilon^{-1}\tau - t_1}$ does not involve collisions and the velocity configurations remain different, contributing, in the limit, with the factor in the rhs of (3.31). It remains to prove that the cardinality of the set $\{(t_1, t_2, t_3)$ s.t. $\varepsilon^{-1}\tau > t_1 > t_2 > t_3 \geqslant 0$ and a collision happens$\}$, denoted by $\theta(\varepsilon, \tau)$, is such that

$$\lim_{\tau \to 0} \lim_{\varepsilon \to 0} \varepsilon^3 \tau^{-3} \theta(\varepsilon, \tau) > 0 \tag{3.41}$$

This follows from the fact that the following condition is sufficient to have a collision:

$$t_2 - t_3 < \varepsilon^{-1}\tau - t_1 \tag{3.42}$$

Namely the fourth particle is added at time t_3 in a position at distance $2(t_2 - t_3)$ from the third particle and they have opposite velocities. The distance decreases by $2(\varepsilon^{-1}\tau - t_1)$ in a time $\varepsilon^{-1}\tau - t_1$ and will become certainly zero at some time before $\varepsilon^{-1}\tau$ if (3.42) is satisfied. Taking the limit $\varepsilon \to 0$, we get

$$\lim_{\varepsilon \to 0} \varepsilon^3 \theta(\varepsilon, \tau) = \int_0^\tau d\tau_1 \int_0^{\tau_1} d\tau_2 \int_0^{\tau_2} d\tau_3\, \chi(\tau_2 - \tau_3 < \tau - \tau_1) = c\tau^3 \quad \blacksquare \tag{3.43}$$

Proof of Theorem 2.2. The proof of Theorem 2.2 is similar to the proof of Theorem 2.1. In this case we do not normalize the correlation functions and just consider the functions u_n^ε defined by (2.7).

The equations for the u_n^ε are again of the form (3.5), but the operator $\mathscr{C}_\varepsilon$ is defined in a slightly different way, namely, setting $\underline{x}=(x_1,\dots,x_n)$ and $\underline{Y}=\underline{y}_1\cup,\dots,\cup\underline{y}_n$, we define

$$(\mathscr{C}_\varepsilon u^\varepsilon)_n(\underline{x})=\sum_{\underline{Y}}\sum_{J\subset\{1,\dots,n\}}\varepsilon^{|J|}(1-\varepsilon)^{n-|J|}\\ \times\prod_{i\in J}b(\underline{y}_i,x_i')\prod_{i\notin J}\tilde{b}(\underline{y}_i,x_i')\,u^\varepsilon_{|\underline{Y}|}(\underline{Y}) \tag{3.44}$$

where u^ε denotes the sequence $\{u_n^\varepsilon\}_{n=1}^\infty$ and $\tilde{b}(\underline{y},x)=\delta(|\underline{y}|,1)\,\delta(y_1,x)$. We notice that from (2.7) it follows that

$$0\leqslant u_k^\varepsilon(x_1,\dots,x_n;t)\leqslant 1 \tag{3.45}$$

Therefore the analogue of Proposition 3.1 is an obvious consequence of the definition and it holds for any time. Using (3.45) instead of Proposition 3.1, we can extend the previous arguments to prove the convergence of the u_n^ε's. The only difference comes from the powers of ε involved. Indeed we modify the definitions (3.20)–(3.22) as follows:

$$(\mathscr{S}_\varepsilon u^\varepsilon)_n(\underline{x})=\sum_{\underline{Y}}(1-\varepsilon)^n\prod_{i=1}^n\tilde{b}(\underline{y}_i,x_i')\,u^\varepsilon_{|\underline{Y}|}(\underline{Y}) \tag{3.46}$$

$$(\mathscr{Q}_\varepsilon u^\varepsilon)_n(\underline{x})=\sum_{\underline{Y}}\sum_{i=1,\dots,n}(1-\varepsilon)^{n-1}\,b(\underline{y}_i,x_i')\prod_{j\neq i}\tilde{b}(\underline{y}_j,x_j')\,u^\varepsilon_{|\underline{Y}|}(\underline{Y}) \tag{3.47}$$

$$(\mathscr{R}_\varepsilon u^\varepsilon)_n(\underline{x})=\sum_{\underline{Y}}\sum_{J\subset\{1,\dots,n\},|J|>1}\varepsilon^{|J|-2}(1-\varepsilon)^{n-|J|}\\ \times\prod_{i\in J}b(\underline{y}_i,x_i')\prod_{i\notin J}\tilde{b}(\underline{y}_i,x_i')\,u^\varepsilon_{|\underline{Y}|}(\underline{Y}) \tag{3.48}$$

Introducing $\tilde{u}_n^\varepsilon$ and $\bar{u}_n^\varepsilon$ as in (3.25) and (3.26), the estimate (3.27) is still valid, since $\varepsilon<1$, but (3.30) has to be modified to take into account the fact that three-body collisions now have the same weight ε as two-body collisions. We have

$$\bar{u}_n^\varepsilon(\underline{x};[e^{-1}\tau])\leqslant\sum_{N\geqslant 0}\varepsilon^{N+1}\sum_{t_1=0}^{t-1}\cdots\sum_{t_N=0}^{t_{N-1}-1}\sum_{\sigma_1=1,2}\cdots\sum_{\sigma_N=1,2}n\cdot(n+\sigma_1)\cdot\cdots\\ \times\cdots\cdot(n+\sigma_1+\cdots+\sigma_N)\,7^{n+\sigma_1+\cdots+\sigma_N}\\ \leqslant 7^n\varepsilon\sum_{N\geqslant 0}(56\tau)^N \tag{3.49}$$

The last step follows from the fact that

$$n\cdot(n+\sigma_1)\cdot\cdots\cdot(n+\sigma_1+\cdots+\sigma_N)<2^N(n+N)!/n!$$

Therefore we have the convergence of u_n^ε as before, but for $\tau^*<1/56$.

As mentioned in Section 1, the uniform *a priori* bound (3.45) implies Theorem 2.2. In fact, as before, we deduce by the same arguments as refs. 8 and 13 the convergence of

$$u_k^\varepsilon([\varepsilon^{-1}q_1], e_1,..., [\varepsilon^{-1}q_k], e_k; \tau)$$

to a family of correlation functions that we denote by $u_k(q_1, e_1,..., q_k, e_k; \tau)$ for $\tau \leqslant \tau^*$ and for almost all $(q_1, e_1,..., q_k, e_k) \in (\mathbb{R}^2 \times \mathscr{V})^k$. Then the argument can be repeated in the time interval $[\tau^*, 2\tau^*]$ with initial datum $u_k(q_1, e_1,..., q_k, e_k; \tau^*)$, since, at time τ^*, the (3.45) still holds. Again the arguments of ref. 13 show the factorization of the correlation functions, but with no restriction on the set of configuration, because the probability of a given collision goes to zero in the limit $\varepsilon \to 0$. Namely, since internal collisions are also performed with probability ε, the arguments yielding (3.31) are not valid and $u_1(x; \tau)$ converges to the solution of (2.9). ∎

4. MULTISCALE ANALYSIS

We now go back to the HPP model in a stirred environment defined in Section 2. The evolution in this system is mostly made by stirring updatings. Since the stirring process has good smoothing properties (notice that a single stirring particle moves like a symmetric random walk), one may hope that such properties survive after inserting the HPP updatings. If this is so, some details of the initial configurations may not be relevant, e.g., only the averages of the initial configuration over regions of area t might have effective influence on the state at time t (recall that a random walk in two dimensions after a time t is spread out over an area of the order of t). Typical initial configurations averaged over regions of area t, $t \approx \varepsilon^{-1-\nu}$, namely at macroscopic times, reproduce the smooth initial macroscopic profile in some faithful way, as we shall see. If the above argument is correct, one may find essentially the same behavior when starting from most of the initial configurations as well as from the initial measure μ^ε. We could then hope to have convergence at short macroscopic times τ as in the traditional Lanford approach, but starting from single typical configurations. This could be the first step of an iterative procedure if at time τ the typical configurations keep the same structural properties as the initial ones.

There are many ifs in the above arguments, but the scheme will turn out to be essentially correct. It is, however, inadequate for an effective strategy of proof, as it looks quite difficult, at least directly, to single out the effects of the stirring from the complexity of the full evolution. Our strategy follows a somewhat modified pattern based on the arguments we present below. First notice that in the time interval $0 < t < \varepsilon^{-\nu}$ the problem disappears, simply because there are no HPP updatings, but only stirrings. By known probability estimates on the stirring process one can actually show that at time $\varepsilon^{-\nu}$ starting from "most" of the initial configurations the state looks like one with a density profile whose maximal value is ε^{ν}. This is still far from the macroscopic density values, which are of the order of ε; nonetheless a density profile with values in $[0, \varepsilon^{\nu}]$ can be studied using the techniques presented in the previous section up to a time t during which the number of HPP updatings is proportional to the inverse of the density, hence up to $t \approx \varepsilon^{-2\nu}$. Till this time one can argue that collisions are not so important and that mainly the process is made up by stirrings and streamings. Consequently the state at time $\varepsilon^{-2\nu}$ approximates a profile whose density is of the order of $\varepsilon^{2\nu}$: but then we can repeat the previous argument to reach times $\varepsilon^{-3\nu}$ and so on till macroscopic times.

This approach is closer to what we really do, but not literally: it already fails at the first step. It is indeed true that we can control the density, i.e., the one-body correlation functions, and prove that at time $\varepsilon^{-\nu}$ the density is at most of the order of ε^{ν}. The estimates on the stirring process show also that the n-body correlation functions are bounded by $c_n \varepsilon^{\nu n}$ and this gives the right dependence on ε: the bounds on the coefficients c_n are, however, too bad for applying the techniques used in the previous section. However, by conditioning at time $\varepsilon^{-\nu}$, we fix the configuration and avoid the buildup of the correlations. Then by using the good factorization properties of the measure at time $\varepsilon^{-\nu}$, as determined by the v-functions, we prove that the typical configurations at time $\varepsilon^{-\nu}$ have the same structural properties as the initial ones. We can then study as before the next time interval of length $\varepsilon^{-\nu}$; actually we can do this ε^{-a} times, if $a > 0$ is small enough (with respect to ν) and prove (i) that at time $\varepsilon^{-\nu-a}$ the one-body correlation functions are bounded by some constant times $\varepsilon^{\nu+a}$ and (ii) that the v-functions are bounded accordingly. We still do not have a good estimate on the size of the n-body correlations for $n \to \infty$, yet our estimates allow us to prove that the configurations at time $\varepsilon^{-\nu-a}$ satisfy the same conditions as those needed to apply the preceding procedure. We can then iterate ε^{-a} times this analysis of the process, which involves now a time interval of length $\varepsilon^{-\nu-a}$. In this way we reach time $\varepsilon^{-\nu-2a}$. By arguments similar to the previous ones, we can iterate ε^{-a} times the analysis of a time interval of length $\varepsilon^{-\nu-2a}$, and repeating this procedure, we eventually reach macroscopic times and prove Theorem 2.3.

The main point in the above scheme is a multiscale space-time analysis: a good control of the process at a given level, i.e., for a given space-time scale, allows us to control it for a slightly longer time. We can then exploit such an extra time and obtain, as an effect of the stirring, averages over larger spatial regions, hence improved bounds on the density. In this way we have access to the next space-time level. We shall now be more precise. Each level, or step, in our analysis is characterized by a time, a space, and a density units in terms of three positive parameters β, a, and ζ, which should be considered as fixed from now on. We shall also need, later on, another parameter, b, which is related to the previous ones. The values of all these parameters will be determined in the course of the proofs. To get an idea about them, consider that $v > \beta > b > a > \zeta$, $b > a + \zeta$, and $b < (1 - v)/2$. In particular, we can choose any β suitably small. Given such a value of β, we can take any b small enough and, again, given such a b, we take any a small enough and, given such an a, we can then take ζ small enough. It is convenient to choose $a > 0$ as the inverse of an integer, so we set

$$\bar{h}a = 1, \qquad \bar{h} \in \mathbb{N} \tag{4.1}$$

Definition: The Level *h*. For any integer $h \in [1, \bar{h} - 1]$ we define

$$T_\varepsilon(h) = \varepsilon^{-v}[\varepsilon^{-a}]^h \tag{4.2}$$

where $[r]$ denotes the integer prt of r. [Notice that $T_\varepsilon(h)$ for $h < \bar{h}$ is vanishingly small in macroscopic units.] In order to introduce the corresponding spatial scales, we first define h^* as the first integer such that

$$\varepsilon^{-v-h^*a+\beta} \geqslant \varepsilon^{-1} \tag{4.3}$$

(observe that since $v > \beta$, $\bar{h} > h^*$). We then let

$$A_\varepsilon(h) = \begin{cases} T_\varepsilon(h)\, \varepsilon^\beta & \text{if } h < h^* \\ \varepsilon^{-1} & \text{if } h \geqslant h^* \end{cases} \tag{4.4}$$

and

$$\rho_\varepsilon(h) = \frac{\varepsilon^{-\zeta}}{A_\varepsilon(h)} \tag{4.5}$$

Finally, we set

$$T_\varepsilon(0) = A_\varepsilon(0) = \rho_\varepsilon(0) = 1 \tag{4.6}$$

We say that $T_\varepsilon(h)$, $A_\varepsilon(h)$, and $\rho_\varepsilon(h)$ are respectively the time, space (area), and density units at the h level.

Notation. We denote by $\Lambda_{\varepsilon,t}$ the square in $\mathbb{Z}^2$ centered at the origin with area $(2[\varepsilon^{-4}] - 2t + 1)^2$. The important point here is that $\Lambda_{\varepsilon,t}$ contains the square centered at the origin with area ε^{-4} for all $t \leqslant \varepsilon^{-1-\nu}T$, T as in Theorem 2.3. We might have chosen $\Lambda_{\varepsilon,0}$ with side $\varepsilon^{-\alpha}$, $\alpha > 2$, and defining accordingly $\Lambda_{\varepsilon,t}$, and everything in the sequel would have worked as well.

We shall study the configurations at time t only in the region $\Lambda_{\varepsilon,t}$; the possibility of "large density fluctuations" prevents us from an analysis on the whole $\mathbb{Z}^2$. We shall therefore use seminorms rather than sup norms, as in the following definition.

Definition: Seminorms on the Particle Configurations. Given any positive integer $h \in [1, \bar{h}-1]$, we denote by Q_h a square in $\mathbb{Z}^2$ of area $A_\varepsilon(h)$. Furthermore, for any $h \in [1, \bar{h}-1]$, any $t \in [0, \varepsilon^{-1}T]$, T being the time mentioned in Theorem 2.3, and for any function $g\colon \Omega_\varepsilon \to \mathbb{R}$ we define

$$\|g\|_{\varepsilon,h,t} = \max_{Q_h \subset \Lambda_{\varepsilon,t}} \max_{e,\sigma} \sum_{q \in Q_h} |g(q, e, \sigma)| \tag{4.7}$$

A configuration η is "h-good" at time t with coefficient d if

$$\|\eta\|_{\varepsilon,h,t} \leqslant d\varepsilon^{-\zeta} \tag{4.8}$$

We shall simply say that η is h-good if it is so at time 0.

Remarks. The law of the random configurations at time t inside $\Lambda_{\varepsilon,t}$ conditioned on the value of the configuration at time $s < t$ depends only on the restriction to the region $\Lambda_{\varepsilon,s}$ of the configuration at time s. Hence if we limit ourselves, as we shall do, to studying the process only in the time-space region $(t, \Lambda_{\varepsilon,t})$, $t \in [0, [\varepsilon^{-1-\nu}T]]$, we have effectively reduced ourselves to finite volumes. Indeed our proofs extend quite straightforwardly to the case when the system is in a square with center the origin and with side $L\varepsilon^{-1}$. Assuming periodic boundary conditions, one can prove the analogue of Theorem 2.3 where the limiting equation is the Broadwell equation on the torus of side L. We shall not, however, discuss this case.

According to what has been said above, the seminorms depend on the configurations at time t only in the region $\Lambda_{\varepsilon,t}$. It should be noticed that the seminorms increase with h, giving a more accurate description of the configuration. They are strictly increasing (keeping ε and t fixed and varying h) for $h \leqslant h^*$, while for $h > h^*$ they remain constant: in fact, the squares of sides ε^{-1} have, in the average with respect to the initial measure μ^ε, a nonvanishing number of particles. The densities of particles in an h-good configuration are of the order of $\rho_\varepsilon(h)$ if we compute the densities by averaging over squares Q_h. Hence they have the right density value for the level h. The whole problem will be to show that the averages over smaller squares are not really relevant.

The next lemma shows that the measure μ^ε [cf. (2.16)] has support on h-good configurations for all $h \geqslant 1$.

Lemma 4.1. Let μ^ε be as in (2.18); then for any u and any $d > 0$ there is c so that

$$\mu^\varepsilon(\{\|\eta\|_{\varepsilon,h,0} \leqslant d\varepsilon^{-\zeta}; \forall h \geqslant 1\}) \geqslant 1 - c\varepsilon^u \tag{4.9}$$

Proof. According to the above Remark, it is enough to prove

$$\lim_{\varepsilon \to 0} \mu^\varepsilon(\{\|\eta\|_{\varepsilon,h^*,0} \leqslant d\varepsilon^{-\zeta}\}) = 1 \tag{4.10}$$

From the Chebychev inequality for any $n > 1$

$$\begin{aligned} &\mu^\varepsilon(\{\|\eta\|_{\varepsilon,h^*,0} > d\varepsilon^{-\zeta}\}) \\ &\leqslant \sum_{Q_{h^*} \subset A_{\varepsilon,0}} \sum_{e,\sigma} d^{-n}\varepsilon^{\zeta n}\mathbb{E}_{\mu^\varepsilon}\Big[\sum_{q \in Q_{h^*}} \eta(q,e,\sigma)\Big]^n \\ &\leqslant c d^{-n}\varepsilon^{\zeta n}\varepsilon^{-8} A_\varepsilon(h^*)^n \max_{e,\sigma,Q_{h^*}} \mathbb{E}_{\mu^\varepsilon}\Big[\frac{1}{A_\varepsilon(h^*)} \sum_{q \in Q_{h^*}} \eta(q,e,\sigma)\Big]^n \\ &\leqslant \bar{c} d^{-n}\varepsilon^{\zeta n}\varepsilon^{-8} A_\varepsilon(h^*)^n\, \varepsilon^n \end{aligned} \tag{4.11}$$

where c and $\bar{c}$ are suitable constants. Equation (4.11) easily follows from the definition of μ^ε. Since $A_\varepsilon(h^*)^n = \varepsilon^{-n}$, the lemma follows after choosing n sufficiently large. ▌

As already mentioned when outlining the strategy of the proof, one of the main points in the analysis of the level h is to prove that the one-body correlation functions starting from an h-good initial configuration are bounded at time $T_\varepsilon(h)$ by a constant times $\rho_\varepsilon(h)$. A first step is to prove the analogous property for the ρ-functions defined in Section 2.

Proposition 4.2. Let $h < \bar{h} - 1$, $d > 0$, $s \in [0, 2T_\varepsilon(h+1))$. Let η be such that $\|\eta\|_{\varepsilon,h,s} \leqslant d\varepsilon^{-\zeta}$. Then there is $c(d)$ such that for all $x = (q, e, \sigma)$ with $q \in A_{\varepsilon,t}$

$$\rho(x, t \mid \eta, s) \leqslant c(d) \begin{cases} \varepsilon^{-\zeta}/(t-s) & \text{for } \ 1 \leqslant t - s \leqslant A_\varepsilon(h) \\ \rho_\varepsilon(h) & \text{for } \ A_\varepsilon(h) \leqslant t - s \leqslant 2T_\varepsilon(h+1) - s \end{cases} \tag{4.12}$$

For notation see Definition: the ρ-Functions, in Section 2.

We shall prove Proposition 4.2 in Section 5. We next relate the actual trajectories of the process to the ρ functions.

Definition: The (h, b, d)-Good Trajectories. Let $h < \bar{h} - 1$; we then denote by $\eta^{(i)}$ the particle configuration at time $t_i \equiv iT_\varepsilon(h)$. For $b > a + \zeta$ and $d > 0$, the (h, b, d)-good set of trajectories $\mathscr{G}_\varepsilon(h, b, d)$ is the set of all trajectories $\underline{\eta} = \{\eta^{(i)}\}$, $i = 0,..., [\varepsilon^{-a}]$, such that:

1. For all $i = 0,..., [\varepsilon^{-a}]$, $\|\eta^{(i)}\|_{\varepsilon, h, t_i} \leqslant d\varepsilon^{-\zeta}$, where $t_i \equiv iT_\varepsilon(h)$.
2. For all $i = 0,..., [\varepsilon^{-a}]$

$$|||\eta^{(i)} - \rho(\cdot, t_i \mid \eta^{(i-1)}, t_{i-1})|||_{\varepsilon, h, t_i} \leqslant d\varepsilon^b \rho_\varepsilon(h) \tag{4.13}$$

where, for any function g,

$$|||g|||_{\varepsilon, h, t} = \max_{q \in \Lambda_{\varepsilon, t + T_\varepsilon^*(h)}} \max_{e, \sigma} \left| \sum_y P_{t + T_\varepsilon^*(h), t}((q, e, \sigma) \to y)\, g(y) \right| \tag{4.14a}$$

$$T_\varepsilon^*(h) = T_\varepsilon(h)\, \varepsilon^{\beta/2} \tag{4.14b}$$

For $t > s$ we have set

$$P_{t,s} = P_{t,t-1} \circ \cdots \circ P_{s+1,s} \tag{4.14c}$$

and

$$P_{s+1,s}(x \to y) = \begin{cases} \text{as in (2.15)} & \text{if } \quad s \notin \varepsilon^{-\nu}\mathbb{N} \\ \delta(q', q - e)\, \delta(e', e)\, \delta(\sigma', \sigma) & \text{otherwise} \end{cases} \tag{4.14d}$$

We shall prove a proposition, Proposition 4.3, which states that the (h, b, d)-good trajectories have large probability, under the assumption that the following estimate on the v-functions holds:

Definition: The Good Estimate on the v-Functions at the Level h. For any $\gamma < \beta/4$, any n, and any d there exists c so that for all η such that $\|\eta\|_{\varepsilon, h, 0} \leqslant d\varepsilon^{-\zeta}$, for all $t \in [T_\varepsilon(h), 2T_\varepsilon(h)]$, $t \in \varepsilon^{-\nu}\mathbb{N}$, and all n-tuples of distinct states $\underline{x} = (x_1,..., x_n)$ such that the site q_i of the state x_i is, for all i in $\Lambda_{\varepsilon, t}$,

$$|v_n^\varepsilon(\underline{x}, t \mid \eta)| \leqslant c\varepsilon^{\gamma n} \rho_\varepsilon(h)^n \tag{4.15}$$

Proposition 4.3. Assume the validity of the good estimate on the v-functions at the level $h < \bar{h} - 1$ and fix any $d_0 \geqslant 0$. Then there is d such that for any u there is a c so that

$$\mathbb{P}_\eta(\mathscr{G}_\varepsilon(h, b, d)) \geqslant 1 - c\varepsilon^u \tag{4.16}$$

for all the configurations η such that $\|\eta\|_{\varepsilon, h, 0} \leqslant d_0 \varepsilon^{-\zeta}$. Furthermore, the trajectories in $\mathscr{G}_\varepsilon(h, b, d)$ are such that for some c and for all $x = (q, e, \sigma)$ with $q \in \Lambda_{\varepsilon, t}$

$$|\rho(x, t \mid \eta^{(i)}, t_i) - \rho(x, t \mid \eta)| \leqslant ci\varepsilon^b \rho_\varepsilon(h) \tag{4.17}$$

where $t_i \equiv iT_\varepsilon(h)$ and, given t, let j be such that $t_j < t \leqslant t_{j+1}$. Then, if $t - t_j \geqslant T_\varepsilon^*(h)$, $i = j$; otherwise $i = j - 1$.

Proposition 4.3 will be proven in the next section. A consequence of Proposition 4.3 is the following lemma.

Lemma 4.4. Let $h < \bar{h} - 1$ and assume the validity of the good estimate on the v-functions at the level h. Then for any n and d_0 there is c so that for all configurations such that $\|\eta\|_{\varepsilon,h,0} \leqslant d_0 \varepsilon^{-\zeta}$, for all $t \in [T_\varepsilon(h), 2T_\varepsilon(h+1)]$, $t \in \varepsilon^{-\nu}\mathbb{N}$, and for all n-tuples of distinct states $\underline{x} = (x_1, ..., x_n)$ such that the site q_i of the state x_i is, for all i in $\Lambda_{\varepsilon,t}$,

$$|v_n^\varepsilon(\underline{x}, t \mid \eta)| \leqslant c\varepsilon^{(b-a)n} \rho_\varepsilon(h)^n \tag{4.18}$$

Proof. Let $\eta' \equiv \eta^{(i)}$ be the random configuration at time t_i, as in (4.17). Let $s \equiv t - t_i$; then, by the Markov property, we have

$$\begin{aligned} |v_n(\underline{x}, t \mid \eta)| &= \left| \mathbb{E}_\eta \left(\prod_{i=1}^n \{\eta(x_i, t) - \rho(x_i, t \mid \eta)\} \right) \right| \\ &= \left| \mathbb{E}_\eta \left(\mathbb{E}_{\eta'} \left(\prod_{i=1}^n [\eta(x_i, s) - \rho(x_i, s \mid \eta') \right.\right.\right. \\ &\qquad \left.\left.\left. + \rho(x_i, s \mid \eta') - \rho(x_i, t \mid \eta)] \right) \right) \right| \end{aligned} \tag{4.19}$$

We choose d so that (4.16) holds and we denote by χ_B the characteristic function of the set B and by $\mathscr{G}_i$ the set containing $\mathscr{G}_\varepsilon(h, b, d)$ determined by imposing the conditions 1 and 2 in the definition of the (h, b, d)-good trajectories only for $j \leqslant i$. Then, by using (4.16), we get from (4.19) that $|v_n(\underline{x}, t \mid \eta)|$ is bounded by

$$\begin{aligned} &\left| \mathbb{E}_\eta \left(\chi_{\mathscr{G}_i} \mathbb{E}_{\eta'} \left(\prod_{i=1}^n [\eta(x_i, s) - \rho(x_i, s \mid \eta') \right.\right.\right. \\ &\qquad \left.\left.\left. + \rho(x_i, s \mid \eta') - \rho(x_i, t \mid \eta)] \right) \right) \right| + c\varepsilon^u \\ &\leqslant \sum_{J \subset \{1, ..., n\}} \mathbb{E}_\eta \left(\chi_{\mathscr{G}_i} |v_{|J|}(\underline{x}_J, s \mid \eta') \right. \\ &\qquad \left. \times \prod_{i \notin J} |\rho(x_i, s \mid \eta') - \rho(x_i, t \mid \eta)| \right) + c\varepsilon^u \end{aligned} \tag{4.20}$$

where $|J|$ denotes the cardinality of the set J and $\underline{x}_J = \{x_j, j \in J\}$. We want to use (4.15) to estimate the v-function in (4.20). But this is not immediate, because η', being a configuration at time t_i, is controlled only in $\Lambda_{\varepsilon,t_i}$ and not in $\Lambda_{\varepsilon,0}$, as required for applying (4.15). However, for each $x=(q,e,\sigma)\in \underline{x}_J$, $q\in\Lambda_{\varepsilon,t}$, so that the v-function in (4.20) only depends on $\{\eta'(q',e',\sigma')\}$, $q'\in\Lambda_{\varepsilon,t_i}$. We therefore have

$$v_{|J|}(\underline{x}_J, s\,|\,\eta') = v_{|J|}(\underline{x}_J, s\,|\,\eta''), \qquad \eta''(q,e,\sigma) = \begin{cases} \eta'(q,e,\sigma) & \text{if} \quad q\in\Lambda_{\varepsilon,t_i} \\ 0 & \text{otherwise} \end{cases}$$

Since $\|\eta'\|_{\varepsilon,h,t_i} \leqslant d\varepsilon^{-\zeta}$, then $\|\eta''\|_{\varepsilon,h,0}\leqslant d\varepsilon^{-\zeta}$. We can now use the assumption that the good estimate on the v-function at the level h is valid, so that, by (4.15) and (4.17), we get that (4.20) is bounded by

$$c' \sum_{J\subset\{1,\ldots,n\}} (\varepsilon^{\gamma}\rho_\varepsilon(h))^{|J|}\,(\varepsilon^{b-a}\rho_\varepsilon(h))^{n-|J|} + c\varepsilon^{u}$$

From this the lemma follows. ∎

The estimate (4.18) is not the good estimate at the level $h+1$, but, as we shall see in Section 6, Lemma 4.4 provides one of the main ingredients for the proof of the following.

Theorem 4.5. The good estimate on the v-functions holds for all $h\leqslant\bar{h}-1$.

In Section 6 we shall prove that this holds for $h=1$ and that if it holds for all $h'<h\leqslant\bar{h}-1$, then it holds for h: this will prove Theorem 4.5.

Theorem 4.5 and Propositions 4.2 and 4.3 give us a good control of the process for a time $T_\varepsilon(\bar{h}-1)$. We need to iterate this at most ε^{-a} times in order to reach the macroscopic times. By imposing some further conditions on the initial configuration (which ensure that they are suitably close to the initial profile), we can easily extend the previous analysis and complete the proof of Theorem 2.3 as shown at the end of Section 5.

5. PROOF OF THEOREM 2.3

We start by proving Proposition 4.2. By the definition of the ρ-functions (see Section 2), we easily get for all $t>s\geqslant 0$

$$\rho(x,t) = \sum_y P_{t,s}(x\to y)\,\rho(y,s) + \sum_{\substack{s\leqslant s'<t\\ s'\in\varepsilon^{-\nu}\mathbb{N}}} \sum_y P_{t,s'}(x\to y)\,\mathscr{C}_1\rho(y,s) \tag{5.1}$$

where $\mathscr{C}_1\rho$ is defined in (2.16) and (2.9b) and $P_{t,s}$ in (4.14d).

Lemma 5.1. Let $s \geqslant 0$ and assume that $\rho(x, s)$ has values in $[0, 1]$ for all x; then $\rho(x, t)$, as given by (5.1), also has values in $[0, 1]$ for all x and all $t > s$. Furthermore, for all $t \geqslant s$,

$$\mathscr{C}_1 \rho(x, t) \leqslant \rho(x^{\perp}, t)\, \rho(x^{-\perp}, t) \tag{5.2}$$

where if $x = (q, e, \sigma)$, then $x^{\pm\perp} = (q, \pm e^{\perp}, \sigma)$; see below (2.2a) for notation.

Proof. Let $\tilde{\rho}$ be obtained from ρ by the action of either a stirring or a streaming or a collisional updating. If $0 \leqslant \rho(x) \leqslant 1$ for all x, then also $0 \leqslant \tilde{\rho}(x) \leqslant 1$. This is obvious for the stirring and the streaming updatings and it can be easily checked to hold in the case of a collisional updating. After this, (5.2) easily follows. ∎

We fix any two finite sequences $\underline{e}$ and $\underline{\tau}$ with the same number of elements, where $\underline{e}$ is a sequence of velocities and $\underline{\tau}$ is a decreasing sequence of times, all between t and s. Call $\tilde{P}_{t,s}(x \to y \mid \underline{e}, \underline{\tau})$ the transition probability obtained from $P_{t,s}(x \to y)$ by the following procedure. Let t_1 be the largest time in $\underline{\tau}$ and consider $P_{t,t_1}(x \to z)$; then change the e-velocity of z (which is the same as that in x) into e_1, the first element in $\underline{e}$. Call z' the resulting state. Consider then $P_{t_1,t_2}(z' \to w)$, where t_2 is the second largest time in $\underline{\tau}$. Change the e-velocity of w into e_2, the second element in $\underline{e}$, and so on. The resulting probability is $\tilde{P}_{t,s}(x \to y \mid \underline{e}, \underline{\tau})$. We have the following result.

Lemma 5.2. There is a constant c' for which the following is true. Given any nonnegative integer s we let $\|\eta\|_{\varepsilon,h,s} \leqslant d\varepsilon^{-\zeta}$. Then for all the integers $t > s$, for all $x = (q, e, \sigma)$, $q \in \Lambda_{\varepsilon,t}$, and for all $\underline{e}$ and $\underline{\tau}$:

$$\sum_y \tilde{P}_{t,s}(x \to y \mid \underline{e}, \underline{\tau})\, \eta(y) \leqslant c'd \begin{cases} \varepsilon^{-\zeta}/(t-s) & \text{if} \quad t-s \leqslant A_\varepsilon(h) \\ \rho_\varepsilon(h) & \text{otherwise} \end{cases} \tag{5.3}$$

Recall that $\rho_\varepsilon(h)$ is defined in (4.5).

Proof. Let $t - s \leqslant A_\varepsilon(h)$. We then introduce a partition of the space into squares of area $t - s$ and denote by B the restriction of this partition to $\Lambda_{\varepsilon,s}$. Then, setting

$$\hat{P}_{t,s}(x \to Q) = \max_{q \in Q} \max_{e,\sigma} \tilde{P}_{t,s}(x \to (q, e, \sigma) \mid \underline{e}, \underline{\tau})) \tag{5.4}$$

For $x = (q, e, \sigma)$, $q \in \Lambda_{\varepsilon,t}$, we get

$$\sum_y \tilde{P}_{t,s}(x \to y \mid \underline{e}, \underline{\tau})\, \eta(y) \leqslant d\varepsilon^{-\zeta} \sum_{Q \in B} \hat{P}_{t,s}(x \to Q)$$

We have used that $\tilde{P}_{t,s}(x \to y \mid \underline{e}, \underline{\tau}) = 0$ if $y = (q', e', \sigma')$ with $q' \notin \Lambda_{\varepsilon,s}$. The lemma follows then from the fact that the position displacement due to the stirring updatings in $\tilde{P}$ are those of a symmetric random walk. Therefore there are constants c and c' such that

$$\sum_{Q \in B} \hat{P}_{t,s}(x \to Q) \leqslant \frac{c}{t-s} \sum_{Q \in B} e^{-d(q_0, Q)^2/2}$$
$$\leqslant \frac{c'}{t-s}$$

where $d(q_0, Q)$ denotes the minimal number of squares in B separating q_0 from Q; q_0 is the site obtained by putting a particle initially in q $[x = (q, e, \sigma)]$ and then letting it move (backward) only at the collisional updatings by one step opposite to its e-velocity, initially equal to e, and then determined by $(\underline{e}, \underline{\tau})$. This proves (5.3) for $t - s \leqslant A_\varepsilon(h)$. In the other case we write $\tilde{P}_{t,s} = \tilde{P}_{t,s^*} \circ \tilde{P}_{s^*,s}$, with $s^* = s + A_\varepsilon(h)$. The first inequality in (5.3) (already proven) is now used for $\tilde{P}_{s^*,s}$. This and the fact that $\tilde{P}_{t,s^*}$ is a contradiction in L_∞ complete the proof of the lemma. ∎

Proof of Proposition 4.2. We write for simplicity $\rho(x, t')$ for $\rho(x, t' \mid \eta, s)$. We shall prove (4.12) with $c(d) = 4c'd$, where c' is the constant which appears in Lemma 5.2. We start by considering $t - s \leqslant A_\varepsilon(h)$ and in this case we prove (4.12) with $c(d) = 2c'd$ and for ε small enough. The proof is by induction, so that given $t' \in (s, t]$, we assume we have already proven the estimate for $s' < t'$. From (5.1), (5.2), Lemma 5.2, and the induction hypothesis, we get, for any $x = (q, e, \sigma)$, $q \in A_{\varepsilon, t'}$,

$$\rho(x, t') \leqslant c'd \frac{\varepsilon^{-\zeta}}{t'-s} + \sum_{\substack{s \leqslant s' < t' \\ s' \in \varepsilon^{-\nu}\mathbb{N}}} \sum_{y} P_{t',s'}(x \to y)\, \rho(y^\perp, s')\, 2c'd \frac{\varepsilon^{-\zeta}}{s'-s}$$

Since

$$\sum_{s' \in \varepsilon^{-\nu}\mathbb{N} \cap (s, t']} \frac{1}{s'-s} \leqslant c'' \varepsilon^{\nu} \log(t'-s)$$

iterating, we get

$$\rho(x, t') \leqslant c'd \frac{\varepsilon^{-\zeta}}{t'-s} \sum_{k \geqslant 0} \frac{1}{k!} [2c'dc''\varepsilon^{\nu-\zeta} \log(t'-s)]^k$$
$$\leqslant c'd \frac{\varepsilon^{-\zeta}}{t'-s} \exp[2c'dc''\varepsilon^{\nu-\zeta} \log A_\varepsilon(h)]$$

For ε small enough, the exponential becomes smaller than 2, hence the first inequality in (4.12) is proven with $c(d) = 2c'd$.

We prove the second inequality in (4.12) with $c(d)=4c'd$. The proof is by induction. Given $t' \in (s+A_\varepsilon(h), t]$, we assume that for all $s' \in (s+A_\varepsilon(h), t')$ and all $x \in A_{\varepsilon, s'}$

$$\rho(x, s') \leqslant 4c'd\rho_\varepsilon(h)$$

Using (5.1), we write

$$\rho(x, t') = \sum_y P_{t', s+A_\varepsilon(h)}(x \to y)\, \rho(y, s+A_\varepsilon(h)) + \sum_{\substack{s' \in \varepsilon^{-\nu}\mathbb{N} \\ s' \in (s+A_\varepsilon(h), t')}} \sum_y P_{t', s'}(x \to y)\, \mathscr{C}_1 \rho(y, s')$$

We then have, for all $x=(q, e, \sigma)$, $q \in A_{\varepsilon, t'}$, using the induction, (5.2), and the first inequality in (4.12), proven with $c(d)=2c'd$, and $t-s=A_\varepsilon(h)$,

$$\rho(x, t') \leqslant 2c'd\rho_\varepsilon(h) + \varepsilon^\nu[t' - A_\varepsilon(h)][4c'd\rho_\varepsilon(h)]^2 \tag{5.5}$$

We have that $t' - A_\varepsilon(h) \leqslant 2T_\varepsilon(h+1)$. If $h<h^*$ [see (4.3)], then by (4.5) and (4.2)

$$\rho_\varepsilon(h)\, T_\varepsilon(h+1) \leqslant \varepsilon^{-a}\rho_\varepsilon(h)\, T_\varepsilon(h) = \varepsilon^{-\beta-a-\zeta}$$

Hence the second term in (5.5) is bounded by

$$\varepsilon^{\nu-\beta-a-\zeta} 8c(d)^2\, \rho_\varepsilon(h) \leqslant 2c'd\rho_\varepsilon(h)$$

for ε small enough and choosing $\beta + a + \zeta < \nu$.

Assume now that $h \geqslant h^*$. Then by (4.1)

$$\rho_\varepsilon(h)\, T_\varepsilon(h+1) \leqslant \varepsilon^{-\zeta+1} T_\varepsilon(\bar{h}-1) \leqslant \varepsilon^{-\zeta+1}\varepsilon^{-1-\nu+a}$$

hence the second term in (5.5) is bounded by

$$\varepsilon^\nu \varepsilon^{-\nu+a-\zeta} 8c(d)^2\, \rho_\varepsilon(h) \leqslant 2c'd\rho_\varepsilon(h)$$

if ε is small enough and choosing $\zeta < a$.

Hence in all cases $\rho(x, t') \leqslant 4c'd\rho_\varepsilon(h)$ for ε small enough. By taking a suitably larger value for $c(d)$, we then prove the proposition for all $\varepsilon \in (0, 1]$. ∎

Corollary to the Proof of Proposition 4.2. Under the same assumptions of Proposition 4.2, there is a constant $c(d)$ so that for all the integers $t > s' > s$, for all $x=(q, e, \sigma)$, $q \in A_{\varepsilon, t}$, and for all $\underline{e}$ and $\underline{\tau}$:

$$\sum_y \tilde{P}_{t, s'}(x \to y \mid \underline{e}, \underline{\tau})\, \rho(y, s' \mid \eta, s) \leqslant c(d) \begin{cases} \varepsilon^{-\zeta}/(t-s) & \text{if } \; t-s \leqslant A_\varepsilon(h) \\ \rho_\varepsilon(h) & \text{otherwise} \end{cases}$$

We shall omit the proof of this corollary, which is essentially the same as that of Proposition 4.2.

Proof of Proposition 4.3. We shall prove (4.16) with $d = 2c(d_0)$; see Proposition 4.2 for the definition of $c(d_0)$. We first prove that if $\eta \in \mathscr{G}_\varepsilon(h, b, d)$, then (4.17) holds. We shall then use this result to prove (4.16), so that we start as follows.

Proof of (4.17). We define

$$R_\varepsilon(t) = \sup_{q \in \Lambda_{\varepsilon, t}} \sup_{e, \sigma} \bar{R}_\varepsilon(q, e, \sigma, t)$$

$$\bar{R}_\varepsilon(x, t) = |\rho(x, t \mid \eta^{(k-1)}, t_{k-1}) - \rho(x, t \mid \eta)|, \qquad t_{k-1} \equiv (k-1)\, T_\varepsilon(h) < t \leqslant t_k$$

Denoting by $B(x)$ the set of all the single-particle states with same q position as in x, we have

$$|\mathscr{C}_1 f(x) - \mathscr{C}_1 g(x)| \leqslant \sum_{z, z' \in B(x)} |f(z) - g(z)|\, |f(z') + g(z')| \tag{5.6}$$

We use in the sequel the following notation:

$$t_k = kT_\varepsilon(h), \qquad t^*_{k-1} = t_{k-1} + T^*_\varepsilon(h), \qquad T^*_\varepsilon(h) = T_\varepsilon(h)\, \varepsilon^{\beta/2} \tag{5.7}$$

For $t \in [t^*_{k-1}, t_k]$ and for $x = (q, e, \sigma)$, $q \in \Lambda_{\varepsilon, t}$, we write (5.1) with $s = t_{k-1}$, splitting the sum over s' smaller and larger than t^*_{k-1}. For the former we use (5.6) without exploiting the minus sign in the first factor on the right-hand side, while this is used for $s' \geqslant t^*_{k-1}$ to reconstruct R_ε. We then have, for a constant c which only depends on d,

$$\begin{aligned} \bar{R}_\varepsilon(x, t) \leqslant & \left| \sum_y P_{t, t_{k-1}}(x \to y)[\eta^{(k-1)}(y) - \rho(y, t_{k-1} \mid \eta)] \right| \\ & + \sum_{\substack{s \in \varepsilon^{-\nu}\mathbb{N} \\ s \in (t_{k-1}, t_{k-1} + A_\varepsilon(h)]}} c \frac{\varepsilon^{-\zeta}}{s - t_{k-1}} \rho_\varepsilon(h) + \sum_{\substack{s \in \varepsilon^{-\nu}\mathbb{N} \\ s \in (t_{k-1} + A_\varepsilon(h), t^*_{k-1}]}} c\rho_\varepsilon(h)^2 \\ & + \sum_{\substack{s \in \varepsilon^{-\nu}\mathbb{N} \\ s \in (t^*_{k-1}, t)}} cR_\varepsilon(s)\, \rho_\varepsilon(h) \end{aligned} \tag{5.8}$$

For the second and third terms on the right-hand side of (5.8) we have used Proposition 4.2 to bound one of the two factors in (5.6). For the other one we have used the Corollary to the proof of Proposition 4.2; hence

$$\sum_y P_{t, s}(x \to y) \max_{z \in B(y)} \rho(z, s \mid \eta^{(k-1)}, t_{k-1}) \leqslant c(d)\, \rho_\varepsilon(h) \tag{5.9}$$

the same bound holding for $\rho(z, s \mid \eta)$. For the fourth term in (5.8) we have used again (5.9).

The first term on the right-hand side of (5.8) is bounded by

$$\left| \sum P_{t, t_{k-1}}(x \to y)[\eta^{(k-1)}(y) - \rho(y, t_{k-1} \mid \eta^{(k-2)}, t_{k-2})] \right| + R_\varepsilon(t_{k-1})$$
$$\leqslant d\varepsilon^b \rho_\varepsilon(h) + R_\varepsilon(t_{k-1}) \tag{5.10}$$

We bound the second term on the right-hand side of (5.8) by

$$c'\varepsilon^{\nu-\zeta}\rho_\varepsilon(h) \log A_\varepsilon(h) \leqslant c''\varepsilon^b \rho_\varepsilon(h)\, \varepsilon^{\nu-\zeta-b} \log \varepsilon^{-1}$$

where c' and c'' are suitable constants and we choose $b + \zeta < \nu$.

The third term on the right-hand side of (5.8) is bounded by

$$c\varepsilon^\nu T_\varepsilon^*(h)[\rho_\varepsilon(h)]^2 \leqslant c\varepsilon^{\nu-\zeta}\rho_\varepsilon(h)\, \varepsilon^{\beta/2} \frac{T_\varepsilon(h)}{A_\varepsilon(h)} \leqslant c\varepsilon^b \rho_\varepsilon(h)\, \varepsilon^{-\zeta-b+\beta/2}\varepsilon^a$$

The last bound only arises if $h = \bar{h} - 1$; in the other cases we get a smaller one by choosing $b + \zeta < \beta/2$. Hence, for a suitable constant c, we get

$$R_\varepsilon(t) \leqslant R_\varepsilon(t_{k-1}) + \sum_{\substack{s \in \varepsilon^{-\nu}\mathbb{N} \\ s \in (t_{k-1}^*, t)}} R_\varepsilon(s)\, c\rho_\varepsilon(h) + \psi \tag{5.11a}$$

$$\psi = \varepsilon^b \rho_\varepsilon(h)(d + c\varepsilon^{\nu-\zeta-b} \log \varepsilon^{-1} + c\varepsilon^{-\zeta-b+\beta/2}) \tag{5.11b}$$

We denote by $\mathscr{I}$ the set of times s such that $s - t_l \geqslant T_\varepsilon^*(h)$, where $t_l < s \leqslant t_{l+1}$. By iteration we then get from (5.11)

$$R_\varepsilon(t) \leqslant \sum_{\substack{s \in \varepsilon^{-\nu}\mathbb{N} \\ s \in \mathscr{I}, s < t}} R_\varepsilon(s)\, c\rho_\varepsilon(h) + \frac{t}{T_\varepsilon(h)}\psi$$
$$\leqslant \sum_{n \geqslant 0} \frac{1}{n!}[t\varepsilon^\nu c\rho_\varepsilon(h)]^n \frac{t}{T_\varepsilon(h)}\psi \leqslant e^{t\varepsilon^\nu \rho_\varepsilon(h)} \frac{t}{T_\varepsilon(h)}\psi \tag{5.12}$$

This proves (4.17) for $t \in \mathscr{I}$. If instead $t_{k-1} < t < t_{k-1}^*$, we repeat the above proof with t_{k-2} instead of t_{k-1}. Thus, (4.17) holds for all (h, b, d)-good trajectories.

Proof of (4.16). Denote by $\mathscr{G}_i$ the set of trajectories η which satisfy the conditions 1 and 2 in the definition of the (h, b, d)-good trajectories, but only for $\eta^{(j)}$ with $j \leqslant i$. We have chosen $d = 2c(d_0)$. We shall now estimate the conditional probability of $\mathcal{G}_i$ given $\mathcal{G}_{i-1}$. Denoting by $\|\cdot\|$ the seminorm $\|\cdot\|_{\varepsilon,h,t_i}$, we have

$$\|\eta^{(i)}\| \leqslant \|\eta^{(i)} - \rho(\cdot, t_i \mid \eta^{(i-1)}, t_{i-1})\| + \|\rho(\cdot, t_i \mid \eta^{(i-1)}, t_{i-1}) - \rho(\cdot, t_i \mid \eta^{(0)})\| + \|\rho(\cdot, t_i \mid \eta^{(0)})\| \tag{5.13}$$

By Proposition 4.2 the last term is bounded by $c(d_0)\,\varepsilon^{-\zeta}$. By (4.17) the second term is bounded by $2c\varepsilon^{-a}\varepsilon^{b}\varepsilon^{-\zeta}$, hence for ε small enough this is smaller than $\varepsilon^{-\zeta}c(d_0)/2$. For the first one we shall prove later that for any u there is c so that

$$\mathbb{P}_\eta\left(\|\eta^{(i)} - \rho(\cdot, t_i \mid \eta^{(i-1)}, t_{i-1})\| \geqslant \frac{c(d_0)}{2}\,\varepsilon^{-\zeta} \,\middle|\, \mathcal{G}_{i-1} \right) \leqslant c\varepsilon^{u} \tag{5.14}$$

where $P_\eta(\cdot \mid \mathcal{G}_{i-1})$ denotes the conditional expectation with respect to $\mathcal{G}_{i-1}$. We now complete the proof of (4.16). We have so far seen that, conditioned on $\mathcal{G}_{i-1}$, $\eta^{(i)}$ satisfies the condition 1 [of the definition of the (h, b, d)-good trajectories] with probability not smaller than $1 - c\varepsilon^{u}$. We shall also prove later that, if $b < \gamma$ [see (4.15)], then for any u there is c so that

$$\mathbb{P}_\eta(|||\eta^{(i)} - \rho(\cdot, t_i \mid \eta^{(i-1)}, t_{i-1}|||_{\varepsilon,h,t_i} \geqslant d\varepsilon^{b}\rho_\varepsilon(h) \mid \mathcal{G}_{i-1}) \leqslant c\varepsilon^{u} \tag{5.15}$$

Hece for any $\bar{u}$ there is c so that

$$\mathbb{P}(\mathcal{G}_i \mid \mathcal{G}_{i-1}) \geqslant 1 - c\varepsilon^{\bar{u}}$$

We then have that if η is an h-good configuration with coefficient d_0, then

$$\mathbb{P}_\eta(\mathcal{G}_\varepsilon(h, b, 2c(d_0))) \geqslant 1 - c\varepsilon^{\bar{u}}\varepsilon^{-a}$$

for any $b > a + \zeta$, $b < \gamma$ [see (4.15)] [hence for any $b \in (a + \zeta, \beta/4)$]. This will prove (4.16) with $u = \bar{u} - a$ once we show the validity of (5.14) and (5.15).

Proof of (5.14). The left-hand side of (5.14) is bounded by $[\eta' \equiv \eta^{(i-1)},\ c' \equiv c(d_0)/2]$

$$\sum_{Q_h \subset \Lambda_{\varepsilon,t_i}} (c'\varepsilon^{-\zeta})^{-2n} \sum_{\underline{y}} \mathbb{E}_{\eta'}\left(\prod_{i=1}^{2n} [\eta(y_i, T_\varepsilon(h)) - \rho(y_i, T_\varepsilon(h) \mid \eta')] \right) \tag{5.16}$$

where the sum is over all not necessarily different states $\underline{y} = (y_1, \ldots, y_{2n})$ with $y_i = (q_i, e_i, \sigma_i)$ and $q_i \in Q_h$.

To estimate (5.16), we introduce a partition π of $\{1,..., 2n\}$ and let $Y(\pi)$ be the set of all the $\underline{y}$ such that $y_i = y_j$ if and only if i and j are in the same atom of π. For $k > 1$ we can write

$$[\eta(y, T_\varepsilon(h)) - \rho(y, T_\varepsilon(h) \mid \eta')]^k = F_0 + [\eta(y, T_\varepsilon(h)) - \rho(y, T_\varepsilon(h) \mid \eta')] F_1$$

where the F_i are polynomials of the variable $\rho(y, T_\varepsilon(h) \mid \eta')$ whose coefficients are functions of k. In particular, in F_0 the constant term is missing. Given π, denote by $l(\pi)$ the number of atoms of π and by $j(\pi)$ the number of atoms with only one element. Given $\underline{y} \in Y(\pi)$, let $z_1,..., z_{j(\pi)}$ be the states of the singletons, and z_i, $j(\pi) < i \leqslant l(\pi)$, the states in each of the other atoms. We have, writing j for $j(\pi)$ and l for $l(\pi)$,

$$\left| \mathbb{E}_{\eta'} \left(\prod_{i=1}^{2n} [\eta(y_i, T_\varepsilon(h)) - \rho(y_i, T_\varepsilon(h) \mid \eta')] \right) \right| \leqslant \sum_{\sigma_{j+1} \cdots \sigma_l} \left[\prod_{i=j+1}^{l} \hat{F}^i_{\sigma_i} \right] |v(\{\underline{z}^\sigma\}, T_\varepsilon(h) \mid \eta')| \tag{5.17}$$

where the σ_i have values 0 and 1. $\hat{F}^i_{\sigma_i}$ denotes the absolute value of the polynomial F_{σ_i} associated to z_i and $v(\{\underline{z}^\sigma\}, T_\varepsilon(h) \mid \eta')$ is the v-function v_n with n the cardinality of $\{\underline{z}^\sigma\}$ [cf. (2.17)]. Finally, $\{\underline{z}^\sigma\}$ is the configuration made by all the states z_i with $i \leqslant j$ and all z_i with $i > j$ and such that $\sigma_i = 1$. We use (4.15) and (4.12), which can be applied because of the assumptions on η' and by the same argument used in the proof of Lemma 4.4. We then obtain the bound

$$c \sum_{\sigma_{j+1} \cdots \sigma_l} [\rho_\varepsilon(h)]^{\Sigma(1-\sigma_i)} [\rho_\varepsilon(h)\, \varepsilon^\gamma]^{j + \Sigma \sigma_i} \leqslant \hat{c} [\rho_\varepsilon(h)\, \varepsilon^\gamma]^j \rho_\varepsilon(h)^{l-j} = \hat{c} \rho_\varepsilon(h)^l \varepsilon^{\gamma j}$$

for suitable constants c and $\hat{c}$ (which depend on n and d_0, but not on ε). We then have that (5.16) is bounded by

$$c\varepsilon^{-8} (c'\varepsilon^{-\zeta})^{-2n} \sum_\pi A_\varepsilon(h)^{l(\pi)} \rho_\varepsilon(h)^{l(\pi)} \varepsilon^{\gamma j(\pi)} = c'' \varepsilon^{-8} \sum_\pi \varepsilon^{\zeta(2n - l(\pi)) + \gamma j(\pi)}$$

for a suitable constants c, c', and c''. We observe that for any n the number of possible partitions π is finite. For any given π, we denote by $\kappa(\pi)$ the number of atoms of π that are not singletons. Then $2\kappa(\pi) \leqslant 2n - j(\pi)$, hence $l(\pi) = \kappa(\pi) + j(\pi) \leqslant n + j(\pi)/2$. In consequence $2n - l(\pi) \geqslant n - j(\pi)/2$ and

$$c'' \varepsilon^{-8} \sum_\pi \varepsilon^{\zeta(2n - l(\pi)) + \gamma j(\pi)} \leqslant c'' \varepsilon^{-8} \sum_\pi \varepsilon^{\zeta n + (\gamma - \zeta/2) j(\pi)} < c\varepsilon^{-8 + \zeta n}$$

for a suitable constant c and choosing $\zeta < 2\gamma$. For any $u > 0$, after choosing n large enough, we see that the last expression is bounded by $c\varepsilon^{u}$ for a suitable c. This completes the proof of (5.14).

Proof of (5.15). We use the same argument as before. The left-hand side of (5.15) is bounded by $(\eta' = \eta^{(i-1)})$

$$\sum_{Q_h \subset \Lambda_{\varepsilon,t_i}} [d\varepsilon^b \rho_\varepsilon(h)]^{-2n} \sum_{\underline{y}} \prod_j \mathbb{P}_{t_i + T_\varepsilon^*(h), t_i}(0 \to y_j)$$
$$\times \mathbb{E}_{\eta'}\left(\prod_{j=1}^{2n} [\eta(y_j, T_\varepsilon(h)) - \rho(y_j, T_\varepsilon(h) \mid \eta')] \right)$$

We use again (5.17) to bound the above expectation. Consider now a fixed atom π of the partition. There are $2n - l(\pi)$ of the variables y_j which have to coincide with some other variable y_k. Therefore the left-hand side of (5.15) is bounded by

$$c\varepsilon^{-8}[d\varepsilon^b \rho_\varepsilon(h)]^{-2n} \sum_\pi \left(\frac{1}{T_\varepsilon^*(h)}\right)^{2n - l(\pi)} \rho_\varepsilon(h)^{l(\pi)}\, \varepsilon^{\gamma j(\pi)}$$
$$\leqslant c' \sum_\pi \varepsilon^{-2nb + \beta/2[2n - l(\pi)] + \gamma j(\pi)}$$

Using again the estimate $2n - l(\pi) \geqslant n - j(\pi)/2$, we have

$$\varepsilon^{-2nb + \beta/2[2n - l(\pi)] + \gamma j(\pi)} \leqslant \varepsilon^{-2nb + n\beta/2 + (\gamma - \beta/4) j(\pi)} \leqslant c\varepsilon^{(\gamma - b)n}$$

The last step follows by the condition $\gamma < \beta/4$, so that the largest value of the bound is obtained for the maximum value of $j(\pi)$, i.e., $j(\pi) = 2n$. The rest of the proof is the same as before. ∎

Proof of Theorem 2.3 (Conclusion). We shall extend the previous analysis to $h = \bar{h} - 1$, but several modifications will be needed. The first one arises when extending Proposition 4.2; the assumption that $\|\eta\|_{\varepsilon, \bar{h}-1, 0} \leqslant d\varepsilon^{-\zeta}$ is no longer sufficient, and in fact the bounds we get by repeating the proof of Proposition 4.2 in this case diverge when $\varepsilon \to 0$. We need extra assumptions on η, namely, that η is "close" to the initial density $\rho(r, e)$ (see Theorem 2.3). We start by defining

$$F_\varepsilon(q, e, \sigma, t) = \varepsilon f_{\varepsilon^{1+\nu}t}(\varepsilon q, e) \tag{5.18a}$$

where f_t solves (2.4). In particular, by (2.18),

$$F_\varepsilon(q, e, \sigma, 0) = \varepsilon \rho(\varepsilon q, e) = \mathbb{E}_{\mu^\varepsilon}(\eta(q, e, \sigma)) \tag{5.18b}$$

Lemma 5.3. Let T be as in Theorem 2.3. There is c such that for all $t \leqslant \varepsilon^{-1-\nu} T$ and for all $x \in \Gamma$ [see (2.10)]

$$\left| F_\varepsilon(x, t) - \sum_y P_{t,0}(x \to y)\, F_\varepsilon(y, 0) - \sum_{\substack{0 < s < t \\ s \in \varepsilon^{-\nu}\mathbb{N}}} \sum_y P_{t,s}(x \to y)\, \mathscr{C}_1 F_\varepsilon(y, s) \right| \leqslant c\varepsilon(\varepsilon \sqrt{t}) \tag{5.19}$$

Proof. Since f solves (2.4), we can write for $t\varepsilon^{1+\nu} \leqslant T$

$$\varepsilon f_{t\varepsilon^{1+\nu}}(\varepsilon q, e) - \varepsilon f_0(\varepsilon q - et\varepsilon^{1+\nu}, e) - \int_0^{t\varepsilon^{1+\nu}} ds\, \varepsilon \mathscr{C} f_s(\varepsilon q - e(t\varepsilon^{1+\nu} - s), e) = 0 \tag{5.20}$$

We compare the three terms on the left-hand side of this equation with the three terms on the left of (5.19). The first ones are equal by (5.18). For the second ones we have, setting $x = (q, e, \sigma)$,

$$\left| \sum_y P_{t,0}(x \to y)\, F_\varepsilon(y, 0) - F_\varepsilon(q - e\varepsilon^\nu t, e, \sigma, 0) \right| \leqslant c\varepsilon^2 \sqrt{t}\, \|\nabla f\|_\infty$$

where we have used the central limit theorem for estimating $P_{t,0}(x \to y)$. Moreover, by standard arguments one can show that if $\|\nabla f_0\|_\infty$ is finite, there is a constat c such that $\sup_{t \leqslant T} \|\nabla f_t\|_\infty \leqslant c$.

Since there is c such that $|\mathscr{C}_1 F_\varepsilon - \mathscr{C} F_\varepsilon| \leqslant c\varepsilon^3$, by the same argument as before, we have

$$\left| \sum_{\substack{s \in \varepsilon^{-\nu}\mathbb{N} \\ 0 < s < t}} \left[\sum_y P_{t,s}(x \to y)\, \mathscr{C}_1 F_\varepsilon(y, s) - \mathscr{C} F_\varepsilon(q - e\varepsilon^\nu(t - s), e, \sigma, s) \right] \right| \leqslant c e^\nu t \varepsilon^3 + c\varepsilon^\nu \varepsilon t^{3/2} \varepsilon^2$$

where c is a suitable constant. We have

$$\sum_{\substack{0 < s < t \\ s \in \varepsilon^{-\nu}\mathbb{N}}} \mathscr{C} F_\varepsilon(q - e\varepsilon^\nu(t - s), e, \sigma, s) = \sum_{\substack{0 < s < t \\ s \in \varepsilon^{-\nu}\mathbb{N}}} \varepsilon^2 \mathscr{C} f_{\varepsilon^{1+\nu} s}(\varepsilon q - \varepsilon^{1+\nu}(t - s) e, e)$$

which is the Riemann sum of the integral in (5.20); hence the error is bounded by $\varepsilon^2(\|\partial f/\partial t\|_\infty + \|\nabla f\|_\infty)$. This concludes the proof. ∎

We choose η close to F_ε in the $|||\cdot|||$ sense. We can do this because of the following result.

Lemma 5.4. For any $d>0$ and u there is c such that

$$\mu^{\varepsilon}(\{|||\eta - F_{\varepsilon}|||_{\varepsilon,\bar{h}-1,0} \leqslant d\varepsilon^{1+b}\}) \geqslant 1 - c\varepsilon^{u} \tag{5.21}$$

The lemma is proven using the Chebychev inequality with arguments similar to those used when proving Proposition 4.3; details are omitted.

Proposition 4.2 is replaced by the following.

Proposition 5.5. Let $0 \leqslant s < t \leqslant \varepsilon^{-1-\nu}T$ and η be such that

$$\|\eta\|_{\varepsilon,\bar{h}-1,s} \leqslant d\varepsilon^{-\zeta}, \qquad |||\eta - F_{\varepsilon}(\cdot, s)|||_{\varepsilon,\bar{h}-1,s} \leqslant d\varepsilon^{1+b} \tag{5.22}$$

Then there is c so that for all $x=(q, e, \sigma)$, $q \in \Lambda_{\varepsilon,t}$,

$$|\rho(x, t \mid \eta, s) - F_{\varepsilon}(x, t)| \leqslant c \begin{cases} \varepsilon^{-\zeta}/(t-s) & \text{if } \ 1 \leqslant t-s \leqslant A_{\varepsilon}(\bar{h}-1) \\ \rho_{\varepsilon}(\bar{h}-1) & \text{if } \ A_{\varepsilon}(\bar{h}-1) \leqslant t-s \leqslant T_{\varepsilon}^{*}(\bar{h}-1) \\ \varepsilon(\varepsilon^{b} + \varepsilon^{(1-\nu)/2}) & \text{otherwise} \end{cases} \tag{5.23a}$$

Furthermore,

$$|\rho(x, t \mid F_{\varepsilon}(\cdot, 0)) - F_{\varepsilon}(x, t)| \leqslant c\varepsilon\varepsilon^{(1-\nu)/2} \tag{5.23b}$$

where $\rho(x, t \mid F_{\varepsilon}(\cdot, 0))$ is the ρ function with initial datum $F_{\varepsilon}(\cdot, 0)$.

Proof. The first two inequalities in (5.23a) are proven in the same way as in Proposition 4.3, without using the second of the inequalities (5.22). In analogy to the proof of Proposition 4.3, we introduce

$$R_{\varepsilon}(t) = \sup_{q \in \Lambda_{\varepsilon,t}} \sup_{e,\sigma} \bar{R}_{\varepsilon}(q, e, \sigma, t)$$

$$\bar{R}_{\varepsilon}(x, t) = |\rho(x, t \mid \eta, s) - F_{\varepsilon}(x, t)|$$

Then we have as in (5.8), with $h = \bar{h} - 1$,

$$\begin{aligned} \bar{R}_{\varepsilon}(x, t) \leqslant{}& \sum_{y} P_{t,s}(x \to y)\, \bar{R}_{\varepsilon}(y, s) + c\varepsilon\varepsilon^{(1-\nu)/2} \sum_{\substack{s' \in \varepsilon^{-\nu}\mathbb{N} \\ s' \in (s, s+A_{\varepsilon}(h)]}} c\frac{\varepsilon^{-\zeta}}{s'-s}\varepsilon^{1-\zeta} \\ &+ \sum_{\substack{s' \in \varepsilon^{-\nu}\mathbb{N} \\ s' \in (s+A_{\varepsilon}(h), T_{\varepsilon}^{*}(h)]}} c\varepsilon^{2-2\zeta} + \sum_{\substack{s' \in \varepsilon^{-\nu}\mathbb{N} \\ s' \in (T_{\varepsilon}^{*}(h), t)}} cR_{\varepsilon}(s')(R_{\varepsilon}(s') + c'\varepsilon) \end{aligned} \tag{5.24}$$

where we have used the first two inequalities in (5.23a), (5.19), and the corollary to the proof of Proposition 4.2 and we have written $\rho(x, s' \mid \eta) \leqslant R_\varepsilon(s') + c'\varepsilon$. We have, by (5.22), for $t > T_\varepsilon^*(h)$

$$\sum_y P_{t,s}(x \to y)\, \bar{R}_\varepsilon(y, s) \leqslant d\varepsilon^{1+b}$$

Given $C > 0$, let S be the last time after $T_\varepsilon^*(h)$ such that $R_\varepsilon(s') \leqslant C\varepsilon$. We assume $t \leqslant S$ and we replace one of the two $R_\varepsilon(s')$ in the last term in (5.24) by $C\varepsilon$. Hence we have the bound

$$\begin{aligned} R_\varepsilon(t) &\leqslant d\varepsilon^{1+b} + c\varepsilon^{1-2\zeta}\varepsilon^{\nu} \log[A_\varepsilon(h)] + c\varepsilon^{2-2\zeta}\varepsilon^{\nu}\varepsilon^{-1-\nu+\beta/2} + \bar{c}\varepsilon\varepsilon^{(1-\nu)/2} \\ &\quad + c\varepsilon \sum_{\substack{s' \in \varepsilon^{-\nu}\mathbb{N} \\ s' \in (T_\varepsilon^*(h), t)}} R_\varepsilon(s') \\ &\leqslant c'\varepsilon(\varepsilon^b + \varepsilon^{(1-\nu)/2}) + c\varepsilon \sum_{\substack{s' \in \varepsilon^{-\nu}\mathbb{N} \\ s' \in (T_\varepsilon^*(h), t)}} R_\varepsilon(s') \end{aligned}$$

By iteration we then have

$$R_\varepsilon(t) \leqslant C\varepsilon(\varepsilon^b + \varepsilon^{(1-\nu)/2}) \exp(c' t \varepsilon^{1+\nu})$$

which also shows that $S > T\varepsilon^{-1-\nu}$. In an analogous way we prove (5.23b). ∎

We put

$$t_i = iT_\varepsilon(\bar{h} - 1) \qquad \text{for all } i \text{ such that} \quad [\varepsilon^{-1-\nu}T] - t_i \geqslant T^*(\bar{h} - 1) \tag{5.25a}$$

We denote by $l - 1$ the largest i in (5.25a) and define

$$t_l = [\varepsilon^{-1-\nu}T] \tag{5.25b}$$

Definition: the Good Set $\mathscr{F}_\varepsilon(d, b)$. Let $d > 0$ and $(1-\nu)/2 > b > a + \zeta$. Then $\mathscr{F}_\varepsilon(d, b)$ is the set of all trajectories $\eta = \{\eta^{(i)}\}$, where $\eta^{(i)}$ is the configuration at the times t_i, $i \leqslant l$, defined in (5.25) which satisfy the conditions 1 and 2 in the definition of the (d, b, h)-good trajectories and moreover $\eta \equiv \eta_0$ satisfies (5.22) and for all $0 \leqslant i \leqslant l$

$$|||\rho(\cdot, t_i \mid \eta^{(i-1)}, t_{i-1}) - \eta^{(i-1)}(\cdot)|||_{\varepsilon, h, t_i} \leqslant d\varepsilon^{1+b} \tag{5.26}$$

The analogue of Proposition 4.3 is the following.

Proposition 5.6. For any $d_0 > 0$ there is d such that the following holds. Given any u, there is c so that

$$\mathbb{P}_\eta(\mathscr{F}_\varepsilon(b, d)) \geqslant 1 - c\varepsilon^u \tag{5.27}$$

for all η which satisfy (5.22) with coefficient d_0. Furthermore, using the same notation as in (4.17), for some c and for all $x=(q, e, \sigma)$, $q \in \Lambda_{\varepsilon,t}$,

$$|\rho(\cdot, t_i | \eta^{(i-1)}, t_{i-1}) - \rho(\cdot, t | \eta)| \leqslant ci\varepsilon^{1+b} \tag{5.28}$$

The proof is just the same as that of Proposition 4.3, using Proposition 5.5 instead of Proposition 4.2; we omit the details. The analogue of Lemma 4.4 also holds for configurations η which satisfy (5.22), so that (4.18) holds for $t=[\varepsilon^{-1-\nu}T]$ and $h=\bar{h}-1$. By Proposition 5.5 we have

$$|\rho(x, t | \eta) - F_\varepsilon(x, t)| \leqslant c\varepsilon^{1+b-a-\zeta},$$

$$|\rho(x, t | \eta) - \rho(x, t | F_\varepsilon(\cdot, 0))| \leqslant c\varepsilon^{1+b-a-\zeta}$$

Using this and (5.18a), we prove (2.19) and conclude the proof of Theorem 2.3. ▮

6. THE BASIC ESTIMATE ON THE v-FUNCTIONS

We prove Theorem 4.5 by iteration. We let $h \leqslant \bar{h}-1$, $t \in \varepsilon^{-\nu}\mathbb{N} \cap [T_\varepsilon(h), 2T_\varepsilon(h)]$, and we fix an h-good configuration η; to have lighter notation, in this section we drop writing the dependence on η in the argument of the ρ- and v-functions.

We first express $v_n(\underline{x}, s)$ in terms of v-functions at time $s-1$. By iteration we shall then prove that $v_n(\underline{x}, t)$ is a finite sum of terms which can be interpreted in terms of a branching process. Theorem 4.5 will then be a consequence of some probability estimates on this branching process.

Let $\underline{x}=(x_1,..., x_n)$, $x_i=(q_i, e_i, \sigma_i)$; if the time step $(t-1, t)$ corresponds to a stirring updating, we set

$$\underline{x}^* = \{\hat{x}_i, \sigma_i - \sigma_{\hat{x}_i}, i=1,..., n\}, \qquad \hat{x}_i = (q_i - c_{\sigma_i}, e_i) \tag{6.1a}$$

where the random variables $\sigma_{q,e}$ (which also depend on t) are i.i.d. with values in $\{1, 2, 3, 4\}$, each having probability 1/4. The sum of the σ's is defined modulo 4. For an HPP updating we set

$$\underline{x}^* = \{(q_i - e_i, e_i, \sigma_i), i=1,..., n\} \tag{6.1b}$$

Definition of Clusters. We say that $(\mathscr{C}_1,..., \mathscr{C}_N)$ are the clusters of $\underline{x}^*$ if $(\mathscr{C}_1,..., \mathscr{C}_N)$ is a partition of $\{1,..., n\}$ such that i and j are in the same cluster if and only if $\hat{x}_i = \hat{x}_j$, for stirring updatings; while, for HPP updatings, if and only if $(q_i - e_i, \sigma_i) = (q_j - e_j, \sigma_j)$.

Notice that the clusters of $\underline{x}^*$ are completely determined by $\underline{x}$, both in the HPP and stirring updatings; in particular, in this latter case they do not depend on the choice of the random variables $\sigma_{q,e}$.

Notation. For an HPP updating we write

$$\underset{\sim}{x}^* = \{\underset{\sim}{x}_i^*, i=1,..., N\}, \qquad \underset{\sim}{x}_i^* = (q_i^*, \underset{\sim}{e}_i^*, \sigma_i^*), \qquad \underset{\sim}{e}_i^* = (e_{i,1}^*,..., e_{i,h_i}^*)$$

where $\underset{\sim}{x}_i^*$ is the configuration in the ith cluster of $\underset{\sim}{x}^*$, $h_i \geqslant 1$ denotes the number of particles in the ith cluster, q_i^* are their common positions, σ_i^* their common σ values, and $\underset{\sim}{e}_i^*$ the set of their e-velocities. For a stirring updating we write

$$\underset{\sim}{x}^* = \{\underset{\sim}{x}_i^*, i=1,..., N\}, \qquad \underset{\sim}{x}_i^* = (q_i^*, e_i^*, \underset{\sim}{\sigma}_i^*), \qquad \underset{\sim}{\sigma}_i^* = (\sigma_{i,1}^*,..., \sigma_{i,h_i}^*)$$

As before, h_i denotes the number of particles in the ith cluster, (q_i^*, e_i^*) the common values of the (q, e) state in that cluster, while $\underset{\sim}{\sigma}_i^*$ is the set of values of the σ-velocities. Finally, if $\underset{\sim}{x}_i$, $i \in J$, are disjoint sets of states, we denote by $\{\underset{\sim}{x}_i, i \in J\}$ the collection of all the states $\underset{\sim}{x}_i$.

We have the following results.

Proposition 6.1. Given $\underset{\sim}{x}$ and t, we denote below by J any subset of $\{1,..., N\}$, where N is the number of clusters in $\underset{\sim}{x}^*$. Then, if at $(t-1, t)$ there is an HPP updating, we have

$$v_n(\underset{\sim}{x}, t) = v_n(\underset{\sim}{x}^*, t-1) + \sum_{J \neq \varnothing} \sum_{\{\underset{\sim}{x}_i, i \in J\}} \left[\prod_{i \in J} C_{h_i,k_i}(\underset{\sim}{x}_i, t \mid \underset{\sim}{x}_i^*)\right] \times v_k(\{\underset{\sim}{x}_i, i \in J, \underset{\sim}{x}_i^*, i \notin J\}, t-1) \tag{6.2a}$$

while for a stirring updating we have

$$v_n(\underset{\sim}{x}, t) = \left(\frac{1}{4}\right)^N \sum_{\underset{\sim}{x}^*} \Bigg(v_n(\underset{\sim}{x}^*, t-1) + \sum_{J \neq \varnothing} \sum_{\{\underset{\sim}{x}_i, i \in J\}} \times \left[\prod_{i \in J} C_{h_i,k_i}(\underset{\sim}{x}_i, t \mid \underset{\sim}{x}_i^*)\right] v_k(\{\underset{\sim}{x}_i, i \in J, \underset{\sim}{x}_i^*, i \notin J\}, t-1)\Bigg) \tag{6.2b}$$

the sum over $\underset{\sim}{x}^*$ being over all the possible 4^N values of $\underset{\sim}{x}^*$ [see (6.1a)]; $h_i = |\underset{\sim}{x}_i^*|$ and $0 \leqslant k_i \leqslant 4$;

$$k = \sum_{i \in J} k_i + \sum_{i \notin J} h_i$$

$\underset{\sim}{x}_i$ is any set of k_i distinct states having all the same (q, e) [(q, σ)] values as those in $\underset{\sim}{x}_i^*$ if the updating is of stirring [HPP] type.

For stirring updatings, $C_{h_i,k_i}(\underset{\sim}{x}_i, t \mid \underset{\sim}{x}_i^*) = 0$ if $h_i = 1$ and if $\underset{\sim}{x}_i$ is not a proper subset of $\underset{\sim}{x}_i^*$; in the other cases

$$C_{h_i,k_i}(\underset{\sim}{x}_i, t \mid \underset{\sim}{x}_i^*) = \prod_{x' \in \underset{\sim}{x}_i^*/\underset{\sim}{x}_i} [\rho(x', t-1) - \rho(x, t)] \tag{6.3}$$

where $x = (q + c_\sigma, e, \sigma)$ if $x' = (q, e, \sigma)$.

Also for the HPP updatings the coefficients $C_{h,k}$ are uniformly bounded. Furthermore, $C_{1,0} \equiv 0$; $C_{1,1}(x_i, t \mid x_i^*)$ is a polynomial of degree 3 in the variables $\rho(q_i^*, \cdot, \sigma_i^*, t-1)$ with the constant term missing. For $h \geqslant 2$, $k < h$, then $C_{h,k}(\underline{x}_i, t \mid \underline{x}_i^*)$ is a polynomial of degree $\leqslant 4$ in the variables $\rho(q_i^*, \cdot, \sigma_i^*, s)$, $s = t-1, t$, with minimal degree $h-k$ and with all the coefficients uniformly bounded.

Proof. *Stirring updating.* We have, given the $\sigma_{\hat{x}}$,

$$\eta(q, e, \sigma, t) = \eta(q - c_\sigma, e, \sigma - \sigma_{q - c_\sigma, e}, t-1) \tag{6.4}$$

See (6.1a) for notation. Set

$$\underline{x}^* = \{\tilde{x}_1, ..., \tilde{x}_n\}, \qquad \tilde{x}_i = (\hat{x}_i, \sigma_i - \sigma_{\hat{x}_i})$$

Then

$$v_n(\underline{x}, t) = \left(\frac{1}{4}\right)^N \sum_{\{\sigma_{\hat{x}_i}\}} \mathbb{E}\left(\prod_{i=1}^n [\eta(\tilde{x}_i, t-1) - \rho(x_i, t)]\right) \tag{6.5}$$

The sum in (6.5) is not over independent variables, because if i and j are in the same cluster, then $\sigma_{\hat{x}_j} \equiv \sigma_{\hat{x}_i}$. When all the $\hat{x}_i$ are distinct, $N = n$ and, as we are going to see, (6.5) becomes

$$v_n(\underline{x}, t) = \left(\frac{1}{4}\right)^n \sum_{\{\sigma_{\hat{x}_i}\}} v_n(\underline{x}^*, t-1) \tag{6.6a}$$

which proves (6.2b) in this particular case.

We are going to show that (6.6a) is obtained from (6.5) by induction. Recall that

$$\rho(x_i, t) = \frac{1}{4} \sum_\sigma \rho(\hat{x}_i, \sigma, t-1) \tag{6.6b}$$

Let

$$\eta_i = \eta(\tilde{x}_i, t-1), \qquad \rho_i = \rho(\tilde{x}_i, t-1), \qquad \rho_i' = \rho(x_i, t)$$

Under the assumption that the $\hat{x}_i$ are distinct, we will show that for any $k \in \{0, ..., n\}$

$$v_n(\underline{x}, t) = \left(\frac{1}{4}\right)^n \sum_{\{\sigma_{\hat{x}_i}\}} \mathbb{E}\left(\left\{\prod_{i=1}^k [\eta_i - \rho_i]\right\}\left\{\prod_{i=k+1}^n [\eta_i - \rho_i']\right\}\right) \tag{6.6c}$$

which gives (6.6a) for $k = n$. For $k = 0$ the first product on the right-hand side is missing, hence in this case (6.6c) becomes (6.5) and the equality is true. We now assume it for k and want to prove it for $k+1$. We write

$$\eta_{k+1} - \rho_{k+1}' = (\eta_{k+1} - \rho_{k+1}) + (\rho_{k+1} - \rho_{k+1}')$$

and we insert this equality in (6.6c). The first term reproduces (6.6c) with k replaced by $k+1$; the contribution of the other term vanishes, as seen by performing first the sum over $\hat{\sigma}_{k+1}$ and using (6.6b). The induction assumption is therefore proven and the validity of (6.6a) is established.

For the general case we add and subtract in (6.5) from each η_i the corresponding ρ_i. Denote by $\Lambda \subset \{1,..., n\}$ any subset of labels each one belonging to some of those clusters which have more than one element, namely if $i \in \Lambda$, then there is $j \neq i$ such that $\hat{x}_i = \hat{x}_j$. The sum over Λ will denote the sum over all such subsets (if $N = n$, there are no such subsets and the sum is absent). We then have

$$v_n(\underline{x}, t) = \left(\frac{1}{4}\right)^N \sum_{\{\sigma_{\hat{x}_i}\}} \sum_{\Lambda} v_{n-|\Lambda|}(\{\hat{x}_i, \sigma - \sigma_{\hat{x}_i}\}_{\{i \notin \Lambda\}}, t-1) \times \prod_{i \in \Lambda} [\rho(\hat{x}_i, \sigma_i - \sigma_{\hat{x}_i}, t-1) - \rho(x_i, t)]$$

which can be easily seen to be the same as (6.2); we have therefore completed the proof of Proposition 6.1 in the case of stirring updatings.

HPP updating. For $l = 1, 2, 3, 4$ let $\eta_l \equiv \eta(q_i^*, c_l, \sigma_i^*, t-1)$; then it is not difficult to see that

$$\begin{aligned}
&\prod_{j=1}^{h_i} [\eta(q_i^* + e_{i,j}^*, e_{i,j}^*, \sigma_i^*, t) - \rho(q_i^* + e_{i,j}^*, e_{i,j}^*, \sigma_i^*, t)] \\
&\quad - \prod_{j=1}^{h_i} [\eta(q_i^*, e_{i,j}^*, \sigma_i^*, t-1) - \rho(q_i^* + e_{i,j}^*, e_{i,j}^*, \sigma_i^*, t)] \\
&\quad = \eta_1\eta_3(1-\eta_2)(1-\eta_4) \left\{ \prod_{j=1}^{h_i} [\delta_{e_{i,j}^*, c_2} + \delta_{e_{i,j}^*, c_4} - \rho(q_i^* + e_{i,j}^*, e_{i,j}^*, \sigma_i^*, t)] \right. \\
&\quad \left. - [\delta_{e_{i,j}^*, c_1} + \delta_{e_{i,j}^*, c_3} - \rho(q_i^* + e_{i,j}^*, e_{i,j}^*, \sigma_i^*, t)] \right\} \\
&\quad + \eta_2\eta_4(1-\eta_1)(1-\eta_3) \left\{ \prod_{j=1}^{h_i} [\delta_{e_{i,j}^*, c_1} + \delta_{e_{i,j}^*, c_3} - \rho(q_i^* + e_{i,j}^*, e_{i,j}^*, \sigma_i^*, t)] \right. \\
&\quad \left. - [\delta_{e_{i,j}^*, c_2} + \delta_{e_{i,j}^*, c_4} - \rho(q_i^* + e_{i,j}^*, e_{i,j}^*, \sigma_i^*, t)] \right\}
\end{aligned} \tag{6.7a}$$

$$\begin{aligned}
&\prod_{j=1}^{h_i} [\eta(q_i^* + e_{i,j}^*, e_{i,j}^*, \sigma_i^*, t) - \rho(q_i^* + e_{i,j}^*, e_{i,j}^*, \sigma_i^*, t)] \\
&\quad = \sum_{k_i} \sum_{\underline{e}_i} C_{h_i, k_i}(e_i, t \mid \underline{x}_i^*) \prod_{e \in \underline{e}_i^*} [\eta(q_i^*, e, \sigma_i^*, t-1) - \rho(q_i^*, e, \sigma_i^*, t-1)] \\
&\quad + \prod_{j=1}^{h_i} [\eta(q_i^*, e_{i,j}^*, \sigma_i^*, t-1) - \rho(q_i^*, e_{i,j}^*, \sigma_i^*, t-1)]
\end{aligned} \tag{6.7b}$$

where $\delta_{e,e'} = 1$ if $e = e'$ and 0 otherwise. The numerical coefficients $C_{h,k}$ are then defined so that (6.7b) holds. To obtain them explicitly, one has to add and subtract from the η_i in (6.7a) their corresponding ρ's and expand the products. After some simple algebra one can then check that the properties stated in the proposition hold, so that the proposition is proved; we omit the details. ▮

Definition: The Backward Streaming + Stirring Process.[6] Given $t > 0$, we call $s \in (0, t)$ a stirring time if $t - s - 1 \neq k\varepsilon^{-v}$ for some integer k; otherwise we call it an HPP time. Then the backward streaming + stirring process relative to t is defined so that at any stirring time s the updating is made first by a unit shift along the direction opposite to that indicated by σ and then by uniform independent rotations of the σ's as in the direct stirring process. If instead s is an HPP time, then the particles move one step in the direction opposite to their e-velocity.

Let us agree that the v-function to estimate is computed at a given and from now on fixed time $t \in \varepsilon^{-v}\mathbb{N}$ such that $2T_\varepsilon(h) \geqslant t \geqslant T_\varepsilon(h)$, $\bar{h} - 1 \geqslant b$ [cf. (4.15)]. We shall then consider, most of the times in the sequel, the process relative to this time, and we shall refer to this simply as the backward streaming + stirring process.

Notice that $\underline{x}^*$ in (6.1) is obtained from $\underline{x}$ by applying the evolution rules of the backward streaming + stirring process. Therefore, iterating (6.2), we get a series of terms which are characterized by time intervals where there are particles moving with the law of the backward streaming + stirring process, while at the end of these intervals there are branchings where the number of particles may change and their state vary. We shall first take care of the stirring + streaming part by introducing a coupling with an independent process: this will give us the main probability estimates needed for proving the desired bounds on the v-functions.

Definition: The Coupling. We introduce an auxiliary process, the Bernoulli process $(\Sigma^*, \mathbb{P})$, $\Sigma^* \equiv \{\sigma_i^*(s),\ i \geqslant 1,\ s \geqslant 0\}$, where the $\sigma_i^*(s)$ are i.i.d. variables with $\mathbb{P}(\sigma_i^*(s) = j) = 1/4$, $j \in \{1, 2, 3, 4\}$. Then for any time interval $[s_1, s_2]$, $s_1 < s_2 \leqslant t$, for any finite set of labels $J \subset \mathbb{N}$, and for any configuration $\underline{x} = \{x_i,\ i \in J\}$ of distinct states, we introduce a map $\mathcal{M}$ with specifications $\{[s_1, s_2], \underline{x}\}$ defined on the set of trajectories $\{\sigma_i^*(s),\ i \in J,\ s \in [s_1, s_2]\}$ and with values in the set $\{x_i(s),\ i \in J,\ s \in [s_1, s_2]\}$. We set $\underline{x}(s_1) = \underline{x}$ and $e_i(s) = e_i(s_1)$ for all $i \in J$ and $s \in [s_1, s_2]$. Then if $s - 1$ is a stirring time, we let $q_i(s) = q_i(s-1) - c_{\sigma_i(s-1)}$ and

$$\sigma_i(s) = \sigma_j^*(s) - \sigma_j(s-1) + \sigma_i(s-1) \quad \text{(modulo 4)}$$

[6] We shall describe the particle configurations by specifying the states the particles occupy; this is evidently equivalent to the occupation number description used previously.

where $j \leqslant i$ is the smallest label in J such that $\hat{x}_j(s) = \hat{x}_i(s)$, i.e., such that $q_j(s) = q_i(s)$ and $e_j(s) = e_i(s)$.

Notice that $\sigma_i(s) = \sigma_i^*(s)$ when $\hat{x}_l(s) = \hat{x}_i(s)$ implies $l \geqslant i$. Then the particles with labels $l > i$ and such that $\hat{x}_l(s) = \hat{x}_i(s)$ undergo the same σ-rotation as particle i. For this reason we say that particles with higher labels are of second class. If s_1 is an HPP time, then $\sigma_i(s) = \sigma_i(s-1)$ and $q_i(s) = q_i(s-1) - e_i$.

We omit the easy proof of the following lemma:

Lemma 6.2. For any given specification $\{[s_1, s_2], \underset{\sim}{x}\}$ the set of trajectories $\{\underset{\sim}{x}(s)), s \in [s_1, s_2]\}$ inherits from $(\Sigma^*, \mathbb{P})$ via $\mathscr{M}$ the same law as the backward stirring + streaming process in the same time interval and conditioned to $\{\underset{\sim}{x}(s_1) = \underset{\sim}{x}\}$.

Notice that the above realization of the backward stirring + streaming process allows us to identify the particles during their evolution, and therefore from now on we consider a labeled configuration of particles.

The main reason for realizing the process as above is to have a natural way to couple it to the independent process $\underset{\sim}{x}^0(s)$. This is the image under $\mathscr{M}^0$ of the trajectories in Σ^*: $\mathscr{M}^0$ is defined as $\mathscr{M}$ except for the updating of the σ's; we simply set, at the stirring times s, $\sigma_i^0(s) = \sigma_i^*(s)$, and leave all the other rules unchanged. In this way it becomes clear that the displacements of a "stirring" particle, i.e., a particle in $\underset{\sim}{x}(\cdot)$, are the same as those of the corresponding "independent" particle, i.e., the particle in $\underset{\sim}{x}^0(s)$ with the same label, at all the times when the stirring particle is alone [no other stirring particle having the same (q, e) at the moment of the σ rotational updating]. Consequently the following result holds:

Proposition 6.3. For any $n > 1$, $u > 0$, and $\lambda > 0$ there is c so that for any set J of n distinct labels, any $\underset{\sim}{x}$, consisting of n different states, any t, and any $[s_1, s_2] \subset [0, t]$,

$$\mathbb{P}\Big(\sup_{s_2 \geqslant s \geqslant T \geqslant s_1} \sup_{i \in J} |q_i(s) - q_i^0(s)| \, (s - s_1)^{-\lambda} \leqslant 1 \Big) \geqslant 1 - c(T - s_1)^{-u} \tag{6.8}$$

where $\mathbb{P}$ is the probability law in Σ^*, $\underset{\sim}{x}(s_1) = \underset{\sim}{x}^0(s_1) = \underset{\sim}{x}$, and q_i and q_i^0 are the positions of the ith stirring and independent particles, respectively.

Proof. As already mentioned, the displacements of the stirring and independent particles with the same label i are the same except when, at the moment of the σ-rotational updating, there is another stirring particle j, $j < i$, with $e_j = e_i$ and $q_j = q_i$. We therefore have to count how many times this happens. The difference between the positions of the two stirring particles i and j has the law of a symmetric random walk between the successive return times to the origin. The probability to have s^λ returns in a time interval s goes to zero like s^{-k} (for any given λ and k positive) as $s \to \infty$. ▮

The Backward Branching Process. When we iterate the equation for the v-functions we get besides stirring + streaming a more complex structure, which can be interpreted in terms of a branching process with births and deaths of particles.[7] The branching process that we consider is denoted by $(\underline{x}(s), I(s))_{0 \leqslant s \leqslant t}$, where $I(s)$ has values 0 and 1. When s is such that $I(s) = 1$ we say that s is an interaction time, a collision interacting time if s is an HPP time, a stirring interacting time otherwise. $\underline{x}(s)$ denotes the set of states (q, e, σ) occupied by the labeled particles present at time s; their number might vary with time.

Not all the trajectories are allowed, as we are going to see. Consider for each i the map $\mathscr{M}$ with specifications $\{[s_{i-1}+1, s_i+1], \underline{x}(s_{i-1}+1)\}$ and let

$$\underline{x}^*(\cdot) = \mathscr{M}(\underline{\sigma}^*(\cdot)) \tag{6.9}$$

We shall say that $\underline{x}(\cdot)$ is allowed if there is $\underline{\sigma}^*(\cdot)$ which is "compatible" with it: $\underline{x}(\cdot)$ and $\underline{\sigma}^*(\cdot)$ are compatible if, first, $\underline{x}(s) = \underline{x}^*(s)$ for $s \in [s_{i-1}+1, s_i]$ and, second, $\underline{x}(s_i+1)$ is obtained from $\underline{x}^*(s_i+1)$ according to the rules we state below. It will be sufficient to specify them at s_1. Assume first that s_1 is a stirring time. Then: (1) in $\underline{x}^*(s_1+1)$ there is a cluster with multiple occupancy, i.e., a set of particles (more than 1) with the same q and e, and (2) $\underline{x}(s_1+1)$ is obtained from $\underline{x}^*(s_1+1)$ by looking at all the clusters with multiple occupancy and by deleting in some of them (or all of them, but not in none of them) some or all of the particles in there and relabeling those remaining in such a way that the labels of those missing are the highest labels among those previously present in the cluster.

Assume now that s_1 is a collision time. Divide the particles of $\underline{x}^*(s_1+1)$ into clusters; recall that particles are in the same cluster if and only if they have the same (q, σ) state. The clusters are then ordered: cluster 1 contains the particle with the lowest label (in this case particle 1, but this is not necessarily so at successive steps, because particle 1 might in the mean time have disappeared). In cluster 2 there is the particle with lowest label among the remaining ones, i.e., all those which are not in cluster 1. The ordering of the other clusters is done similarly. Then $\underline{x}(s_1+1)$ should be obtained from $\underline{x}^*(s_1+1)$ in one of the following ways.

[7] In Section 3 we introduced a branching process to study the correlation functions of the HPP deterministic model. In that case there were no deaths, while here we have to take them into account because we are considering the v-functions and not the correlation functions.

Start from cluster 1 and call (q, σ) their common values of position and σ-velocity. Then change this cluster into any other one, where the e-velocities are distinct, while all the particles have the same (q, σ) as before. This new cluster might be the same as the old one or might even be empty, except when the old cluster had only 1 particle. If the new cluster has fewer particles than the old one, the labeling of the particles in the new cluster satisfies the same convention used before in the case of stirring deaths. If there are more particles—say there are k more particles than in the old cluster—then all the old labels are used to label the particles in the new cluster, the extra k particles being labeled by $N+1$, $N+2$,..., $N+k$, if N was the maximal label used previously; in this specific case $N=n$, of course. The same holds for cluster 2, taking into account the extra labels possibly used for the first cluster and so on, iteratively, till all the clusters are updated.

Notice that in this way we have covered all the cases occurring in (6.2).

Definition: A Measure on the Backward Branching Process. We introduce a positive measure on the product space of the backward branching process times Σ^* as follows. Let $(\underline{x}(\cdot), I(\cdot))$ be a trajectory in the backward branching process and let S be a measurable set in Σ^*; then the measure of the product of these two sets is the measure of all the trajectories $\sigma^*(\cdot) \in S$ which are compatible with $(\underline{x}(\cdot), I(\cdot))$, namely they are compatible, according to the definition given above, in all the time intervals $[s_{i-1}+1, s_i+1]$. In particular, the measure of $(\underline{x}(\cdot), I(\cdot))$ is the measure of all the trajectories $\sigma^*(\cdot)$ compatible with it. Since the same $\underline{\sigma}^*$ might give rise to different trajectories of the backward branching process, the measure we have defined is not normalized to 1 and it is not a probability measure.

Denote by $\mathbb{E}$ the expectation with respect to the measure defined above. We then have from (6.5) and (6.6) for any $0 \leqslant s \leqslant t$

$$|v_n(\underline{x}, t)| \leqslant \mathbb{E}(|v_{|\underline{x}(t-s)|}(\underline{x}(t-s), s)|\, D(\{(\underline{x}(t'), I(t')), t' \leqslant t-s\})) \qquad (6.10)$$

where the functional D is a product of several factors, as we shall see in a while, and $v_0 = 1$. We first specify the initial configuration η; recall that η is not explicitly mentioned in the v-function and that its full expression is given in (2.17a).

Assumptions on the Initial Configuration η. We assume that the initial configuration η is h-good [cf. (4.8)]. Then the solution $\rho(\cdot, \cdot \mid \eta)$ of (5.1) satisfies the three conditions (6.11)–(6.13) stated in the following lemma. While (6.11) and (6.13) are proven in the previous section,[8] (6.12) will be proven in an appendix.

Lemma 6.4. Given any $d>0$ and $h\leqslant \bar{h}-1$, there is c so that for all $\|\eta\|_{\varepsilon,h,0}\leqslant d\varepsilon^{-\zeta}$, for all $s\leqslant 2T_\varepsilon(h)$, and all $x=(q,e,\sigma)$, $q\in\Lambda_{\varepsilon,s}$,

$$\rho(x,s\,|\,\eta)\leqslant c\varepsilon^{-\zeta}\begin{cases}\varepsilon^{-\beta}/s & \text{for}\quad s<\varepsilon^{-1-\beta}\\ \varepsilon & \text{otherwise}\end{cases} \tag{6.11}$$

Furthermore, if $(s,s+1)$ corresponds to a stirring updating,

$$|\rho(q+c_\sigma,e,\sigma,s+1\,|\,\eta)-\rho(q,e,\sigma,s\,|\,\eta)|$$
$$\leqslant\frac{c}{\sqrt{s}}\begin{cases}\min\{\varepsilon^{-\beta}/s,1\} & \text{for}\quad s<\varepsilon^{-1-\beta}\\ \varepsilon & \text{otherwise}\end{cases} \tag{6.12}$$

Finally, let

$$\rho_{(q,s)}=\max_{(e,\sigma)}\rho(q,e,\sigma,s\,|\,\eta)\leqslant 1,\qquad \rho_s=\max_{q\in\Lambda_{\varepsilon,s}}\rho_{(q,s)}$$

and let $P_{s,s'}$ be the transition probability defined in (4.14). Then there is a c so that for all x and all $s'<s$

$$\sum_{(q',e',\sigma')}P_{s,s'}((q,e,\sigma)\to(q',e',\sigma'))\,\rho_{(q',s')}$$
$$\leqslant c\varepsilon^{-\zeta}\begin{cases}\varepsilon^{-\beta}/s & \text{for}\quad s<\varepsilon^{-1-\beta}\\ \varepsilon & \text{otherwise}\end{cases} \tag{6.13}$$

where $q(t')$ denotes the position that the particle has at time t'.

Definition: The Function D. If η satisfies the above assumption so that the bounds in Lemma 6.4 hold, then the function D in (6.10) is the product of the following factors. For each stirring death there is a factor $c\rho_{(q,s)}/\sqrt{s}$ in D, if the death occurs at time $t-s$ in the backward process; q is the position of the missing particle at the moment of its death. Each death for collision contributes a factor $c\rho_{(q,s)}$; recall that deaths for collision as well as for stirring only occur in clusters with multiple occupancy. Finally, it might be that at a collision interaction time there is no cluster with multiple occupancy; then for each of these times there is a factor $c\rho_s$ (see Lemma 6.4 for the origin of this factor) contributing to D.

We shall now rewrite (6.10) in a more convenient way, using a sort of strong Markov property. We are in fact going to choose s as implicitly

[8] The estimate (6.11) is worse than that proven in the previous section, but it allows a more unified analysis of the various cases, as will become clear in the sequel.

determined by the trajectory (of the backward branching process) till time s itself. We say that the function

$$T: \quad \{\underset{\sim}{x}(s), I(s), \sigma^*(s)\}_{s \in [0,t]} \to \mathbb{N}$$

is a stopping time if for any $s \in [0, t]$ the set $\{T = s\}$ does not depend on the values of the trajectory after time s. It is easy to see that (6.10) holds when s is replaced by any stopping time T.

We choose

$$T = \min\{T_1, T_2\}$$

where T_1 and T_2 are defined below. $\{T_1 = s\}$ is the event where s is the first time when all the particles have disappeared. If such a time does not exist, we set $\{T_1 = t\}$; notice that $v(\underset{\sim}{x}, 0) = 0$, $\underset{\sim}{x} \neq \emptyset$, no matter what $\underset{\sim}{x}$ is, so that the case when the particles survive till the final time t will not contribute to $v_n(\underset{\sim}{x}, t)$. The time T_2 is defined by means of a finite sequence of increasing integers $N_1, N_2, \ldots$, fixed once for all independently of ε; its values will be specified later on [see (6.41)]. Assume that $2T_\varepsilon(h) \geqslant t \geqslant T_\varepsilon(h)$ and that $t \in \varepsilon^{-\nu}\mathbb{N}$. Then for s such that $t - s \geqslant T_\varepsilon(h-1)$, $\{T_2 = s\}$ implies that s is the first time up to which there have been exactly N_1 collision interaction times. For $T_\varepsilon(h-1) > t - s \geqslant T_\varepsilon(h-2)$; $\{T_2 = s\}$ implies that we are not in the preceding case and that s is the first time such that in $[t - T_\varepsilon(h-1), s]$ there have been exactly N_2 collision interaction times. An analogous definition is given when $T_\varepsilon(k) > t - s \geqslant T_\varepsilon(k-1)$ $(h > k \geqslant 1)$, $\{T_2 = s\}$. If none of the above conditions is satisfied we set $T_2 = \infty$.

It is easy to see that T is a stopping time, hence that we can replace s in (6.10) by T. We then have

$$\begin{aligned} |v_n(\underset{\sim}{x}, t)| \leqslant{} & \mathbb{E}(1_{\{T = T_1\}} D(\{\underset{\sim}{x}(t'), I(t')), t' \leqslant T_1\})) \\ & + \mathbb{E}(1_{\{T = T_2\}} |v_{|\underset{\sim}{x}(T_2)|}(\underset{\sim}{x}(T_2), t - T_2)| \, D(\{(\underset{\sim}{x}(t'), I(t')), t' \leqslant T_2\})) \end{aligned} \tag{6.14}$$

where 1_A is the characteristic function of the set A. By using the induction hypothesis in Theorem 4.5 we shall bound the v-function appearing in the last term of (6.14) uniformly in $A_{\varepsilon,t}$, so that the estimates of the two expectations in (6.14) become similar to each other.

The expectation in (6.14) is done with respect to a measure defined on a space which is the product of Σ^* times the set of all the trajectories in the backward branching process. On such a space we introduce the "skeleton trajectories" as follows:

Definition: The Skeleton of a Trajectory. Given a trajectory $(\underset{\sim}{x}, I, \underset{\sim}{\sigma}^*)$, its skeleton is the atom π of the partition defined below which contains $(\underline{x}, I, \underline{\sigma}^*)$. The partition is determined by the following equivalence relation: $(\underline{x}(\cdot), I(\cdot), \sigma^*(\cdot))$ and $(\underline{x}'(\cdot), I'(\cdot), \sigma'^*(\cdot))$ are equivalent if and only if the following happens. (1) They have the same number of interaction times, say m; (2) the number of particles, their labels, and their e-velocities are the same in $\underline{x}(s_i+1)$ and $\underline{x}'(s_i'+1)$, $i=1,..., m$; (3) the clusters of $\underline{x}^*(s_i+1)$ and of $\underline{x}'^*(s_i+1)$ are the same (cf. the definition of the backward branching process for notation).

Because of the definition of T there is a bounded number of collision times (i.e., $\leqslant N_1 + N_2 + \cdots$), hence a finite number of births, and since the stirring interactions only produce deaths, each trajectory has a bounded number of interaction times. It will therefore suffice to bound the contribution of the generic skeleton π, because there are finitely many skeletons.

The $\underline{\xi}$ and $\underline{\xi}^0$ Processes. Let us fix a skeleton π and let m be the number of interaction times in π. For any set $\underline{s} = (s_1,..., s_m)$ of increasing times between 0 and t and for any trajectory $\sigma^*(\cdot) \in \Sigma^*$ we define the trajectories $\underline{\xi}(\cdot)$, $\underline{\xi}^0(\cdot)$ as follows. The number of particles, their labels, and their e-velocities change in $\underline{\xi}(\cdot)$ and $\underline{\xi}^0(\cdot)$ as in any of the trajectories of the skeleton: namely, if a particle dies at the ith interaction time in the skeleton, then it dies at time s_i+1 both in $\underline{\xi}(\cdot)$ and $\underline{\xi}^0(\cdot)$; if a particle changes its e-velocity in the skeleton at the ith interaction time, then it does so at time s_i+1 also in $\underline{\xi}(\cdot)$ and $\underline{\xi}^0(\cdot)$; similarly, if a new particle is created in π, then it is also created in $\underline{\xi}$ and $\underline{\xi}^0$ with the same e-velocity. Since the e-velocities remain unchanged between the interaction times, to complete the definition of $\underline{\xi}$ and $\underline{\xi}^0$, we need only specify the positions of the particles and their σ-velocities. Consider the map $\mathscr{M}$ and $\mathscr{M}^0$ with specifications $\{[s_{i-1}+1, s_i+1], \underline{\xi}(s_{i-1}+1)\}$ and $\{[s_{i-1}+1, s_i+1], \underline{\xi}^0(s_{i-1}+1)\}$, respectively. We then set

$$\underline{\xi}^*(\cdot) = \mathscr{M}(\underline{\sigma}^*(\cdot)), \qquad \underline{\xi}^{*,0}(\cdot) = \mathscr{M}^0(\underline{\sigma}^*(\cdot))$$

Then

$$\underline{\xi}(s) = \underline{\xi}^*(s), \qquad \underline{\xi}^0(s) = \underline{\xi}^{*,0}(s); \qquad s_i + 1 \leqslant s \leqslant s_i$$

We now define $\underline{\xi}(s_i+1)$ and $\underline{\xi}^0(s_i+1)$ in the following way. First we remove from $\underline{\xi}^*(s_i+1)$ and $\underline{\xi}^{*,0}(s_i+1)$ all the particles which die at the ith interaction time in the skeleton and change the e-velocities as specified by the skeleton; call $\tilde{\underline{\xi}}$ and $\tilde{\underline{\xi}}^0$ the configurations obtained in this way. In order to complete the definition of $\underline{\xi}(s_i+1)$ and $\underline{\xi}^0(s_i+1)$, we then need to specify the states of the particles born at time s_i+1. Call these new states $\underline{z}$ and $\underline{z}^0$, so that $\underline{\xi}(s_i+1) = \tilde{\underline{\xi}} \cup \underline{z}$, $\underline{\xi}^0(s_i+1) = \tilde{\underline{\xi}}^0 \cup \underline{z}^0$. We set $\underline{z} = \underline{z}^0$; thus, it only remains to specify $\underline{z}$. We use the following convention. Assume that in π, k is the minimal label of the particles created at the ith interaction time; let e be its velocity and l the minimal label of the particles in the cluster to which k belongs, at the moment of its creation. We shall now specify the state $z_k \in \underline{z}$. We call (q, e', σ) the state of the particle with label l in $\underline{\xi}$; then $z_k = (q, e, \sigma)$, if $(q, e, \sigma) \notin \underline{\xi}$. Otherwise $z_k = (q', e, \sigma)$, where q' is the lexicographically closest point to q for which $(q', e, \sigma) \notin \underline{\xi}$. By iteration we specify the whole $\underline{z}$, thus completing the definition of $\underline{\xi}(\cdot)$ and of $\underline{\xi}^0(\cdot)$.

We shall need the following extension of Proposition 6.3.

Proposition 6.5. Given a skeleton π with m interaction times, then for any $\lambda > 0$, $d > 0$, and k there is a c for which the following holds. Fix any increasing sequence of times $\underline{s} = (s_1,..., s_m)$ in $[0, t]$ and call t_i^0 and t_i^f the time when particle i is born, respectively dies, in $\underline{\xi}(\cdot)$, $\underline{\xi}^0(\cdot)$. Call $q_i(s)$, $q_i^0(s)$ its position in $\underline{\xi}(s)$, $\underline{\xi}^0(s)$. Then for any u there is a c so that

$$\mathbb{P}(\sup_i \sup_{t_i^0 + \varepsilon^{-d} \leqslant s \leqslant t_i^f} |q_i(s) - q_i^0(s)| \, (s - t_i^0)^{-\lambda} \leqslant 1) \geqslant 1 - c\varepsilon^u$$

Proof. Proposition 6.5 is actually a corollary of Proposition 6.3; in fact, the streaming updatings of $q_i(s)$ and $q_i^0(s)$ are the same by construction; therefore, the difference $|q_i(s) - q_i^0(s)|$ can only change at the stirring updatings, just as in Proposition 6.3; hence the proof of the present proposition. We omit the details. ∎

Going back to (6.14), we write the first term on its right-hand side as

$$\mathbb{E}(1_{\{T = T_1\}} 1_{\{(\underline{x}, I, \underline{\sigma}^*) \in \pi\}} D(\{\underline{x}(s), I(s), s \leqslant T_1\})\qquad (6.15a)$$

Call $\mathscr{C}^{(i)}$, $i = 1,..., m$, the clusters at the ith interaction time as specified by the skeleton π, and let $\chi_{(i)}$ be the characteristic function that the state of the particles at the ith interaction time have the cluster structure $\mathscr{C}^{(i)}$. Denote by $\underline{s} = (s_1,..., s_m)$ an increasing sequence of times in $[0, t]$ and write

$$I_{\underline{s}}(s) = \begin{cases} 0 & \text{if } \; s \notin \underline{s} \\ 1 & \text{otherwise} \end{cases}$$

Then we have

$$(6.15a) = \sum_{\underline{s}} \text{Est}(\pi, \underline{s}), \qquad \text{Est}(\pi, \underline{s}) = \mathbb{E}\left(D(\underline{\xi}, I_{\underline{s}}) \prod_{i=1}^{m} \chi_{(i)} \right) \quad (6.15b)$$

Since we have fixed π and $\underline{s}$, $\underline{\xi}(\cdot)$ is completely specified by $\underline{\sigma}^*(\cdot)$; hence the expectation in (6.15b) is over the Bernoulli law $\mathbb{P}$ on Σ^*. Denote by k the largest particles' label which appears in π. Then using Proposition 6.5, for any $u > 0$ there is a c so that

$$\mathrm{Est}(\pi, \underline{s}) = \mathbb{E}\left(D(\underline{\xi}, I_{\underline{s}}) \prod_{i=1}^{m} \chi_{(i)}(\{ \sup_{t_k^0 + \varepsilon^{-d} \leqslant s \leqslant t_k^f} |q_k(s) - q_k^0(s)| \right.$$

$$\left. \times (s - t_k^0)^{-\lambda} \leqslant 1 \}) \right) + c\varepsilon^u$$

$$\leqslant \mathbb{E}\left(D^{(k)}(\underline{\xi}, I_{\underline{s}}) \prod_{i=1}^{m} \chi_{(i)}^{(0,k)} \right) \phi(t_k^f) + c\varepsilon^u \tag{6.16a}$$

where $D^{(k)}$, unlike D, does not have the factor coming from the death of particle k; such a factor [see (6.11)] is bounded by $\phi(t_k^f)$:

$$\phi(s) = c\varepsilon^{-\zeta} \begin{cases} \max\{\varepsilon^{-\beta}(t-s)^{-1}, \varepsilon\} & \text{if } s \text{ is a collision time} \\ \max\{\varepsilon^{-\beta}(t-s)^{-1}, \varepsilon\}\, \varepsilon^{\beta/2} & \text{if } s \text{ is a stirring time and } t - s \geqslant \varepsilon^{-\beta} \\ (t-s)^{-1/2} & \text{if } s \text{ is a stirring time and } t - s < \varepsilon^{-\beta} \end{cases} \tag{6.16b}$$

Finally, $\chi_{(i)}^{(0,k)} = \chi_{(i)}$ if the particle k is not alive at the ith interaction time. If this is not the case, then $\chi_{(i)}^{(0,k)}$ is the characteristic function of the cluster structure obtained from $\mathscr{C}^{(i)}$ by requiring that all l and l' different from k are in the same cluster if and only if they were in the same cluster in $\mathscr{C}^{(i)}$. Therefore $\chi_{(i)}^{(0,k)}$ does not impose constraints on x_k, but, on the other hand, there are conditions on x_k^0, the state of the independent particle with label k: for $s_i - t_k^0 > \varepsilon^{-d}$, s_i being either a collision time or $s_i = t_k^f$, then it is requested that $q_k^0(s_i)$ is within distance $|s_i - t_k^0|^\lambda$ from the cluster to which particle k belongs in $\mathscr{C}^{(i)}$, unless this cluster has single occupancy.

The second inequality in (6.16a) is obtained by changing the $\chi_{(i)}$ into the $\chi_{(i)}^{(0,k)}$ by exploiting the characteristic function that $|q_k(s) - q_k^0(s)| \leqslant (s - t_k^0)^\lambda$. Once this is done we bound by 1 the characteristic function on $q_k - q_k^0$. Since $D \leqslant D^{(k)}\phi(t_k^f)$, we then get (6.16a).

Call, for notational simplicity,

$$t_0 \equiv t_k^0 < t_1 < \cdots < t_l \equiv t_k^f \tag{6.17}$$

where the t_i, $1 \leqslant i \leqslant l-1$, are the collision interaction times when the particle k is in a cluster with more than one particle. At these intermediate times no particle dies nor is born in the cluster to which k belongs; in fact, the particles with highest labels are those which die first, while, by definition, k is the last particle born. Denote then by $\mathscr{C}_i^{(k)}$ the cluster structure obtained from $\mathscr{C}^i$ by deleting particle k and let $\chi_{(i)}^{(k)}$ be the corresponding characteristic function. Then we have the following result.

Lemma 6.6. With the above notation

$$\mathrm{Est}(\pi, \underline{s}) \leqslant \mathrm{Est}_{(k)}(\pi, \underline{s})\, \phi(t_k^f)\, c\varepsilon^{-[d+(1+\nu)2\lambda]l} \prod_{i=1}^{l} (t_i - t_{i-1})^{-1} + c\varepsilon^{u} \tag{6.18a}$$

$$\mathrm{Est}_{(k)}(\pi, \underline{s}) = \mathbb{E}\left(D^{(k)}(\underline{\xi}, I_{\underline{s}}) \prod_{i=1}^{m} \chi_{(i)}^{(k)} \right) \tag{6.18b}$$

Proof. According to the definition of the ξ process, the trajectories $\xi_i(s)$, $i<k$ and $s \leqslant t$, are measurable, i.e., are completely determined by the trajectories $\sigma_i^*(s)$, with i and s as above. Therefore the increments of the variable $\xi_k^0(s)$, $s \geqslant t_0$, are independent of the $\xi_i(s)$. The constraints imposed by the characteristic functions $\chi_i^{(0,k)}$ imply that at the times $t_i > t_0 + \varepsilon^{-d}$, the position $q^0(t_i)$ of the variable $\xi_k^0(t_i) - \xi_k(t_0)$ [$\xi_k(t_0) = \xi_k^0(t_0)$ by definition] is in a square of side $2(t_i - t_0)^{\lambda}$ centered around a point which is determined by the $\xi_j(s)$, $j<k$, $s \leqslant t_i$. Therefore, given the positions $\xi_j(s)$, $j<k$, $s \leqslant t$, we have to estimate the probability that a symmetric random walk is at time t_i in a square of side $2(t_i - t_0)^{\lambda}$ centered around a given fixed point, $i = 1,\ldots, l$. Recall that we are now considering $t_i \geqslant t_0 + \varepsilon^{-d}$, so that we can use Proposition 6.5. We first estimate the condition on the last $t_{\bar{l}}$. We condition on the positions of the symmetric random walk at all $s \leqslant t_{\bar{l}-1}$, so we obtain a bound of the form $(t_{\bar{l}} - t_0)^{2\lambda}/(t_{\bar{l}} - t_{\bar{l}-1})$ uniformly on the conditioning. By iterating, we then get the bound

$$c \prod_{i:\, t_i > t_0 + \varepsilon^{-d}} \frac{(t_i - t_0)^{2\lambda}}{t_i - t_{i-1}} \leqslant c \prod_{i:\, t_i > t_0 + \varepsilon^{-d}} \frac{\varepsilon^{-(1+\nu)2\lambda}}{t_i - t_{i-1}}$$

$$\leqslant c\varepsilon^{-(1+\nu)2\lambda l} \varepsilon^{-dl} \prod_{i=1}^{l} \frac{1}{t_i - t_{i-1}} \tag{6.19}$$

From this the lemma follows. ∎

$\mathrm{Est}_{(k)}(\pi, \underline{s})$ looks very much like $\mathrm{Est}(\pi, \underline{s})$, since all references to the trajectory of the particle k have disappeared. There is, however, a difference which has to be taken into account. Assume that the particle k at the moment of its death is in a cluster with only one other particle, say particle h, and particle h also dies. In the new cluster structure $\{\mathscr{C}_i^{(k)}\}$, obtained by erasing particle k, there is therefore a particle, namely particle h, which dies alone (i.e., in a cluster with single occupancy).

The bound for $\mathrm{Est}_{(k)}(\pi, \underline{s})$ is just like that for $\mathrm{Est}(\pi, \underline{s})$ except in the case we have just mentioned and when $h = k-1$. Let us therefore assume that we are in such a case. Call $t'_0 < t'_1 < \cdots < t'_{l'} < t_l$ the times referring to particle $k-1$: t'_0 is the time when it is born, and the times t'_i, $i \leqslant l'$, are the collision times when the particle is in a cluster with multiple occupancy; if there are no such times, then $t'_{l'} = t'_0$; t_l is the time when it dies. Then, after extracting from $D^{(k)}$ the factor coming from the death of particle $k-1$, we get

$$\mathrm{Est}_k(\pi, \underline{s}) \leqslant \mathbb{E}\left(D^{(k-1)}(\underline{\xi}, I_{\underline{s}}) \prod_{i=1}^{m} \chi_{(i)}^{(k)} b(q_{k-1}(t_l), t_l)\right) \tag{6.20a}$$

where $D^{(k-1)}$ does not have the factors referring to the deaths of particles k and $k-1$, while for all q,

$$b(q, t_l) = \begin{cases} \rho_{(q, t-t_l)}/(t-t_l)^{1/2} & \text{if } \; t_l \text{ is a stirring time} \\ \rho_{(q, t-t_l)} & \text{otherwise} \end{cases} \tag{6.20b}$$

Lemma 6.7. Under the above conditions for any u there is a constant c so that

$$\mathrm{Est}_k(\pi, \underline{s}) \leqslant \mathbb{E}\left(D^{(k-1)}(\underline{\xi}, I_{\underline{s}}) \prod_{i=1}^{m} \chi_{(i)}^{(k)}\right) c\phi(t'_{l'})\, e^{-(1+\nu)2\lambda} + c\varepsilon^u$$

Proof. If $t - t_l \geqslant \varepsilon^{-1-\beta}$, then $\rho(\cdot, t-t_l)$ is bounded by $c\varepsilon^{1-\zeta}$, so that we can replace b by $\phi(t'_{l'})$. We shall therefore assume from now on that $t - t_l < \varepsilon^{-1-\beta}$, $t_l - t'_0 \leqslant \varepsilon^{-d}$; then $t'_{l'} = t'_0$ (consecutive collision times are separated by $\varepsilon^{-\nu}$ and we are tacitly assuming that $d < \nu$). We therefore have

$$\frac{1}{t-t_l} \leqslant \frac{1}{t-(t'_0+\varepsilon^{-d})} \leqslant \frac{2}{t-t'_0}$$

because $t - t'_0 \geqslant \varepsilon^{-\nu} > 2\varepsilon^{-d}$. Hence if $t_l - t'_0 \leqslant \varepsilon^{-d}$, then $b(q, t_l) \leqslant c\phi(t'_{l'})$, recall that $t'_{l'}$ is by definition a collision time, so that the factor $(t-t_l)^{-1/2}$ present in (6.20b) can be bounded by 1.

Let us now consider the case $t_l - t'_0 \geqslant \varepsilon^{-d}$. Using Proposition 6.5, we have that for any positive u there is a c so that

$$\begin{aligned} \mathrm{Est}_k(\pi) &\leqslant \mathbb{E}\left(D^{(k-1)}(\underline{\xi}, I_{\underline{s}}) \prod_{i=1}^{m} \chi_{(i)}^{(k)} \sum_{|q - q^0_{k-1}(t_l)| \leqslant \varepsilon^{-(1+\nu)\lambda}} b(q, t_l)\right) + c\varepsilon^u \\ &\leqslant \mathbb{E}(\alpha(\xi^0_{k-1}(t'_{l'}))\, \beta(\xi^0_{k-1}(t'_{l'}))) + c\varepsilon^u \end{aligned} \tag{6.21a}$$

where

$$\begin{aligned} \alpha(\xi) &= \mathbb{E}\left(D^{(k-1)}(\underline{\xi}, I_{\underline{s}}) \prod_{i=1}^{m} \chi_{(i)}^{(k)} \mid \{\xi^0_{k-1}(t'_{l'}) = \xi\}\right) \\ \beta(\xi) &= \mathbb{E}\left(\sum_{|q - q^0_{k-1}(t_l)| \leqslant \varepsilon^{-(1+\nu)\lambda}} b(q, t_l) \mid \{\xi^0_{k-1}(t'_{l'}) = \xi\}\right) \end{aligned} \tag{6.21b}$$

We have used that $\{\xi_i(\cdot),\ i<k-1;\ \xi_{k-1}(s),\ s\leqslant t'_{l'}\}$ is independent of $\{q^0_{k-1}(s),\ s\geqslant t'_{l'}\}$, given $\xi^0_{k-1}(t'_{l'})$, as follows from the definition of the processes ξ and ξ^0; cf. also the proof of Lemma 6.6. We now have, using the translational invariance of the process $q^0_{k-1}(s)$,

$$\begin{aligned}
\sup_{\xi} \beta(\xi) &= \sup_{\xi} \mathbb{E}\Bigg(\sum_{|q-q^0_{k-1}(t_l)| \leqslant \varepsilon^{-(1+\nu)\lambda}} b(q, t_l) \mid \{\xi^0_{k-1}(t'_{l'}) = \xi\} \Bigg) \\
&= \sup_{\xi} \mathbb{E}\Bigg(\sum_{|q| \leqslant \varepsilon^{-(1+\nu)\lambda}} b(q^0_{k-1}(t_l) + q, t_l) \mid \{\xi^0_{k-1}(t'_{l'}) = \xi\} \Bigg) \\
&= \sup_{\xi} \mathbb{E}\Bigg(\sum_{|q| \leqslant \varepsilon^{-(1+\nu)\lambda}} b(q^0_{k-1}(t_l), t_l) \mid \{\xi^0_{k-1}(t'_{l'}) = \xi + q\} \Bigg) \\
&\leqslant \varepsilon^{-(1+\nu)2\lambda} \sup_{\xi} \mathbb{E}(b(q^0_{k-1}(t_l), t_l)) \mid \{\xi_{k-1}(t'_{l'}) = \xi\} \\
&\leqslant c\varepsilon^{-(1+\nu)2\lambda} \phi(t'_{l'})
\end{aligned}$$

where, if $\xi = (q', e, \sigma)$, $\xi + q = (q' + q, e, \sigma)$. In the last inequality we have used Lemma 6.4. The proof of the lemma is therefore completed. ∎

If $t'_{l'} = t'_0$, we simply bound $\chi^{(k)}_{(i)}$ by $\chi^{(k-1)}_{(i)}$, which denotes the characteristic function relative to the cluster structure $\mathscr{C}^{(k-1)}_i$, obtained from $\mathscr{C}_i$ by disregarding both particles k and $k-1$. If $t'_{l'} \neq t'_0$, then we are just in the same situation as when estimating $\mathrm{Est}(\pi, \underline{s})$. So we can proceed iteratively till the contribution of all the particles, those born during the branchings and those alive from the beginning, have been taken into account. By recalling that λ and d may be chosen abitrarily small, we may draw the following conclusions stated in Proposition 6.8 below, after some notation.

Notation. The sup norm over $q \in \Lambda_{\varepsilon,s}$ of $\rho_{(q,s)}$ is denoted by ρ_s. Furthermore, for each particle label $i \in [1, k]$, we denote by $t^{(i)}_j$, $0 \leqslant j \leqslant l^{(i)}$, the relevant times relative to the particle i: $t^{(i)}_0$ is the time when particle i is born ($=0$ if $i \leqslant n$); the $t^{(i)}_j$, $0 < j < l^{(i)}$, are the collision times when the particle i is in a cluster where there is at least a particle with label $h < i$. The last time in the sequence is the time when the particle i dies, if the cluster to which particle i belongs when it dies has at least a particle with label $h < i$. Otherwise the last time is the last collision time when it was in a cluster with at least a particle with lower label.

We have so far proven the following result.

Proposition 6.8. Fix π in $\{T = T_1\}$ and assume that the initial configuration η satisfies the conditions stated in the paragraph, Assumptions on the Initial Configuration η, before Lemma 6.4. Then, using the notation introduced above, for any positive δ and u there is a c so that

$$\mathrm{Est}(\pi, \underline{s}) \leqslant c\left[\prod_{i\in\mathscr{S}} \rho_{s_i}\right] \prod_{i=1}^{k} \mathscr{T}_i + c\varepsilon^u \tag{6.22a}$$

where $\mathscr{S}$ is the subset of all the collision times s_i in $\underline{s}$ when all the clusters have single occupancy [cf. Definition: The Function D, after (6.13)]

$$\mathscr{T}_i = \phi(t^{(i)}_{l(i)}) \prod_{j=0}^{l(i)-1} \frac{\varepsilon^{-\delta}}{t^{(i)}_{j+1} - t^{(i)}_j} \tag{6.22b}$$

We have to bound now the sum over $\underline{s}$ of (6.22a). Recalling that

$$\rho_s \leqslant c\varepsilon^{-\zeta} \max\{\varepsilon^{-\beta}/s, \varepsilon\} \tag{6.23}$$

and that each s_i in $\underline{s}$ is such that $s_i \leqslant t \leqslant \varepsilon^{-1-\nu+a}$, we have that

$$\sum_{\{s_i, i\in\mathscr{S}\}} \prod_{i\in\mathscr{S}} \rho_{s-i} \leqslant c\varepsilon^{(a-\zeta)|\mathscr{S}|} \tag{6.24}$$

We have used the inequality

$$\sum_{\substack{s\in\varepsilon^{-\nu}\mathbb{N}\\ s\leqslant\varepsilon^{-1-\nu}}} \frac{1}{s} \leqslant c\varepsilon^{\nu} \log \varepsilon^{-1} \tag{6.25}$$

which covers the cases $s \leqslant \varepsilon^{-1-\beta}$; this will turn out to be the smallest contribution. The contribution due to $s \geqslant \varepsilon^{-1-\beta}$ gives in fact

$$\varepsilon^{-\zeta}\varepsilon\varepsilon^{\nu}\varepsilon^{-1-\nu+a} \leqslant c\varepsilon^{-\zeta}\varepsilon^{a}$$

so that each $i \in \mathscr{S}$ contributes a factor $c\varepsilon^{a-\zeta}$.

We need now to estimate the contribution of the terms $\mathscr{T}_i$. We proceed iteratively, starting from $\mathscr{T}_k$. Call again $t_0, \ldots, t_l$ the times involved in $\mathscr{T}_k$.

Definition: The Free Times. We shall say that the time t_j is "free" if it does not appear in any other of the $\mathscr{T}_i$.

We shall extensively use the following inequality: for any $\delta > 0$ there is a c so that

$$\sum_{\substack{t_j\in\varepsilon^{-\nu}\mathbb{N}\\ t_{j-1}<t_j<t_{j+1}}} \frac{1}{t_j - t_{j-1}} \frac{1}{t_{j+1} - t_j} \leqslant c\varepsilon^{\nu-\delta} \frac{1}{t_{j+1} - t_{j-1}} \tag{6.26}$$

By (6.26) we can perform the sum of $\mathscr{T}_k$ over the set $\underline{t}_k$ of all the free times among $t_1, \ldots, t_{l-1}$; let their number be m. We obtain a bound consisting of the product of the factor $c\varepsilon^{(\nu-\delta)m}$ times the same $\mathcal{T}_k$, but with the free times dropped. Let us relabel the remaining times as t_j, $0 \leqslant j < l'$. For $0 < j \leqslant l'$ we use the inequality (recall that $t_j \in \varepsilon^{-\nu}\mathbb{N}$ for $j < l'$)

$$\frac{1}{t_j - t_{j-1}} \frac{1}{t_{j+1} - t_j} \leqslant c\varepsilon^{\nu} \frac{1}{t_{j+1} - t_{j-1}} \tag{6.27}$$

We have therefore proven that

$$\sum_{t_k} \mathcal{T}_k \leqslant c\varepsilon^{(\nu-\delta)(l-1)} \frac{1}{t_l - t_0} \phi(t_l) \tag{6.28}$$

Recall that π determines whether t_0 and t_l are or are not free. Set then

$$\tilde{\psi}_k = \begin{cases} \sum_{t_l} \phi(t_l)/(t_l - t_0) & \text{if} \quad t_l \text{ is free} \\ \phi(t_l)/(t_l - t_0) & \text{otherwise} \end{cases} \tag{6.29a}$$

$$\psi_k = \begin{cases} \sum_{t_0} \tilde{\psi}_k & \text{if} \quad t_0 \text{ is free} \\ \tilde{\psi}_k & \text{otherwise} \end{cases} \tag{6.29b}$$

We then have the following result.

Lemma 6.9. For any $\delta > 0$ there is a c so that

$$\psi_k \leqslant \bar{\psi}_k \equiv c\varepsilon^{-\delta-\zeta} \begin{cases} \varepsilon^{\beta/2} & \text{if} \quad k > n \\ \varepsilon^{\beta/2} \max\{\varepsilon^{-\beta}/t, \varepsilon\} & \text{if} \quad k \leqslant n \end{cases} \tag{6.30}$$

Proof. We start from $\tilde{\psi}_k$ and we distinguish various possibilities.

1. *Assume t_l is a stirring time.* Then if t_l is free, $\tilde{\psi}_k$ is bounded by the sum of the following three terms, each corresponding to a range of variation of t_l:

$$c\varepsilon^{-\zeta} \begin{cases} \varepsilon^{-\beta/2-\delta}(t-t_0)^{-1} & \text{for} \quad \varepsilon^{-1-\beta} > t - t_l > \varepsilon^{-\beta} \\ \varepsilon^{-\beta/2}(t-t_0)^{-1} & \text{for} \quad t - t_l < \varepsilon^{-\beta} \\ \varepsilon^{\beta/2-\delta}\varepsilon & \text{for} \quad t - t_l > \varepsilon^{-1-\beta} \end{cases} \tag{6.31a}$$

while if t_l is not free [using (6.27) for the first bound below]

$$\tilde{\psi}_k \leqslant c\varepsilon^{-\zeta} \begin{cases} \varepsilon^{-\beta/2}(t-t_0)^{-1} & \text{for} \quad \varepsilon^{-1-\beta} > t - t_l > \varepsilon^{-\beta} \\ (t-t_0)^{-1}(t-t_l)^{-1/2} & \text{for} \quad t - t_l < \varepsilon^{-\beta} \\ \varepsilon^{\beta/2}\varepsilon & \text{for} \quad t - t_l > \varepsilon^{-1-\beta} \end{cases} \tag{6.31b}$$

To get (6.30), we argue as follows. Let first $k > n$. For the first case in (6.31a), summing over t_0 if free,

$$\varepsilon^{-\beta/2-\delta} \sum_{t_0 \in \varepsilon^{-\nu}\mathbb{N}} (t-t_0)^{-1} \leqslant c\varepsilon^{-\beta/2-\delta}\varepsilon^{\nu-\delta}$$

By our choice of β, much smaller than ν, we have that $\nu - \beta/2 > \beta/2$. Redefining the value of δ, we get the first case in (6.30). If t_0 is not free, since it is a collision time, we have $t - t_0 \geqslant \varepsilon^{-\nu}$, hence the same bound as before.

The second case in (6.31a) is just the same as the previous one, so we get again the right bound. The third one is already the right bound if t_0 is not free, while if t_0 is free, we get

$$\sum_{t_0 \in \varepsilon^{-\nu}\mathbb{N}} \varepsilon^{\beta/2}\varepsilon\varepsilon^{-\delta} \leqslant \varepsilon^{\beta/2}\varepsilon\varepsilon^{-\delta} t\varepsilon^{\nu} \leqslant c\varepsilon^{\beta/2-\delta}$$

If t_l is not free, we have to look at (6.31b), but the bounds in (6.31b) are smaller than the corresponding ones in (6.31a); hence (6.30) for $k > n$ is proved when t_l is a stirring time.

When $k \leqslant n$, $t_0 = 0$ and we obtain directly the bound in (6.30).

2. *Assume t_l is a collision time.* If t_l is free, $\tilde{\psi}_k(t_0)$ is bounded by the sum of the following two terms, which correspond to different ranges of variation of t_l:

$$c\varepsilon^{-\zeta} \begin{cases} \varepsilon^{-\beta+\nu-\delta}(t-t_0)^{-1} & \text{for} \quad t - t_l < \varepsilon^{-1-\beta} \\ \varepsilon^{\nu-\delta}\varepsilon & \text{for} \quad t - t_l \geqslant \varepsilon^{-1-\beta} \end{cases} \tag{6.32a}$$

If t_l is not free,

$$\tilde{\psi}_k \leqslant c\varepsilon^{-\zeta} \begin{cases} \varepsilon^{-\beta+\nu}(t-t_0)^{-1} & \text{for} \quad t - t_l < \varepsilon^{-1-\beta} \\ \varepsilon^{\nu}\varepsilon & \text{for} \quad t - t_l \geqslant \varepsilon^{-1-\beta} \end{cases} \tag{6.32b}$$

For $k > n$ and t_0 free, we get from the first bound in (6.32a)

$$\sum_{t_0 \in \varepsilon^{-\nu}\mathbb{N}} \varepsilon^{\nu-\delta}(t-t_0)^{-1} \leqslant c\varepsilon^{2\nu-2\delta}$$

which agrees with (6.30). The second bound in (6.32a) gives

$$\sum_{t_0 \in \varepsilon^{-\nu}\mathbb{N}} \varepsilon^{\nu-\delta}\varepsilon \leqslant \varepsilon^{\nu-\delta}\varepsilon t\varepsilon^{\nu} \leqslant c\varepsilon^{\nu-\delta}$$

which is also compatible with (6.30). If t_0 is not free, (6.32a) *a fortiori* gives the desired bound. The bounds in (6.32b) are better than the corresponding ones in (6.32a), so that this completes the proof of (6.30) for $k > n$. For $k \leqslant n$, $t_0 = 0$ and (6.30) readily follows from (6.32). This concludes the proof of the lemma. ∎

By (6.28)–(6.30) we obtain a bound on the sum of $\mathcal{T}_k$ over the free times appearing in $\mathcal{T}_k$; the bound is uniform on all the other times appearing in $\mathcal{T}_k$. We therefore have from (6.22a), using (6.24) and denoting by $\underline{s}^{(k)}$ the subset of $\underline{s}$ obtained by dropping the times s_i, $i \in \mathcal{S}$, and those which are free in $\mathcal{T}_k$,

$$\mathrm{Est}(\pi, \underline{s}) \leqslant c\varepsilon^{(a-\zeta)|\mathcal{S}|} \bar{\psi}_k \sum_{\underline{s}^{(k)}} \prod_{i=1}^{k-1} \mathcal{T}_i + c\varepsilon^u \tag{6.33}$$

In (6.33) it is understood that the sum over each s_i in $\underline{s}^{(k)}$ is extended from $s_j + 1$ to $s_h - 1$ if j and h are the labels closest to i among those left in $\underline{s}^{(k)}$.

We are now in the same situation as when estimating the contribution of $\mathcal{T}_k$; the only difference occurs when we are estimating a $\mathcal{T}_i$ in which there is only one time t_0, that is, the time when the particle i was born. This case occurs if the first time the particle i is involved in a nontrivial cluster is when it dies, and it dies together with particles which all have higher label. We have $\mathcal{T}_i = \varepsilon^{-\delta}\phi(t_0)$, and using the same arguments as in the proof of Lemma 6.9, we get

$$\psi_i = \begin{cases} \sum_{t_0} \phi(t_0) & \text{if} \quad t_0 \text{ is free} \\ \phi(t_0) & \text{otherwise} \end{cases}$$

$$\leqslant \psi_i^* \equiv c\varepsilon^{-\delta-\zeta} \begin{cases} \varepsilon^{\nu-\beta} & \text{if} \quad i > n,\ t < \varepsilon^{-1-\beta} \\ \varepsilon^a & \text{if} \quad i > n,\ t > \varepsilon^{-1-\beta} \\ \max\{\varepsilon, \varepsilon^{-\beta}/t\} & \text{if} \quad i \leqslant n \end{cases} \tag{6.34}$$

We notice that any interaction time t_j either appears in some of the $\mathcal{T}_l$ or $j \in \mathcal{S}$; we can then conclude that

$$\sum_{\underline{s}} \mathrm{Est}(\pi, \underline{s}) \leqslant c\varepsilon^{-\delta}\varepsilon^{(a-\zeta)|\mathcal{S}|} \prod_{i=1}^{k} \times \begin{cases} \varepsilon^{\nu(l_i-1)}\bar{\psi}_i & \text{if} \quad l_i \geqslant 1 \\ \psi_i^* & \text{otherwise} \end{cases} \tag{6.35}$$

where $l_i + 1$ is the number of times appearing in $\mathcal{T}_i$. From (6.35) we finally get

$$\sum_{\underline{s}} \mathrm{Est}(\pi, \underline{s}) \leqslant c\varepsilon^{-\delta}[\varepsilon^{-\zeta} \max\{\varepsilon, \varepsilon^{-\beta}/t\}\, \varepsilon^{\beta/4}]^n \tag{6.36}$$

because (1) the contribution of each of the factors in (6.35) is bounded (recall that a is chosen much larger than ζ) and (2) for each $i \leqslant n$ such that $l_i = 0$ there is at least another $j > i$, j being the label of one of the particles which die with i, which contributes a factor $\bar{\psi}_i$, containing a factor $\varepsilon^{\beta/2}$.

Going back to the second term of the right-hand side of (6.14), we use the induction hypothesis and (4.18), so that for any k and any $\delta > 0$ there is a c such that

$$|v_k(\underline{x}, \tau)| \leqslant c[\varepsilon^{b-a-\zeta} \max\{\varepsilon, c\varepsilon^{-\beta}/T_\varepsilon(r)\}]^k$$

for any τ which is an HPP time in $[T'_\varepsilon(r), T'_\varepsilon(r+1))$, where $r \leqslant h-1$ and

$$T'_\varepsilon(r) = \begin{cases} T_\varepsilon(r) & \text{if} \quad r \leqslant h-1 \\ t & \text{if} \quad r = h \end{cases}$$

We then get

$$\begin{aligned} &\mathbb{E}(1_{\{T=T_2\}} |v_{|\underline{x}(T_2)|}(\underline{x}(T_2), t-T_2)| \, D(\{(\underline{x}(t'), I(t')), t' \leqslant T_2\})) \\ &\quad \leqslant c \sum_{r=1}^{h-1} \sum_{\tau \in (T'_\varepsilon(r), T'_\varepsilon(r+1)]} \mathbb{E}(1_{\{T=T_2=t-\tau\}} D((\underline{x}(\cdot), I(\cdot)) \\ &\quad \times [\varepsilon^{b-a-\zeta} \max\{\varepsilon, c\varepsilon^{-\beta}/T_\varepsilon(r)\}]^{|\underline{x}(T_2)|}) \end{aligned} \tag{6.37}$$

We fix r and $\tau \in (T'_\varepsilon(r), T'_\varepsilon(r+1)]$ ad a skeleton π with at least N_{h-r} collision times. Call s_i the ith interaction time and consider all possible increasing strings $\underline{s}$ of such times, so that there are $< N_1$ collision times in $\underline{s}$ in $[0, t - T_\varepsilon(h-1)]$, there are $N'_2 < N_2$ in the interval $[t - T_\varepsilon(h-1), t - T_\varepsilon(h-2)]$, and so on, except for the last one: there are exactly N_r collision times in $[t - T_\varepsilon(h-r+1), \tau]$. For each of these choices we construct the $\underline{\xi}$ and $\underline{\xi}^0$ processes and then sum in (6.37) over $\underline{s}$ for fixed π and τ. We call the corresponding term in (6.37) $F(\pi, \tau)$. We shall now adapt the previous analysis to the case of $F(\pi, \tau)$. Let k be the highest particle label, call t_0 the time when k was born, and let $t_1, \ldots, t_l$ be the successive times when particle k is involved in nontrivial clusters, just as before. The difference with the previous cases is that at t_l particle k may survive and still be alive till time $t - \tau$. If this is so, particle k is one of the particles which contributes one of the factors

$$[\varepsilon^{b-a-\zeta} \max\{\varepsilon, c\varepsilon^{-\beta}/T_\varepsilon(r)\}]$$

appearing in (6.37). To take this into account, we need to change the definition of $\phi(s)$ given in (6.16b) and introduce the quantity

$$\phi(k, s) = \begin{cases} \phi(s) & \text{if} \quad \text{particle } k \text{ dies} \\ \varepsilon^{b-a-\zeta} \max\{\varepsilon, c\varepsilon^{-\beta}/T_\varepsilon(r)\} & \text{otherwise} \end{cases} \tag{6.38}$$

With this change Proposition 6.8 remains valid. Of course, when particle k dies, the estimate of $\mathscr{T}_k$ is just the same. If it does not die, the analysis of the contribution of the intermediate times $t_1,..., t_{l-1}$ is again the same and the estimate (6.28) extends to the present case. By our conventions, t_l is a collision time, since no particle dies in the cluster to which k belongs at time t_l. If t_l and t_0 are not both free, we get

$$\psi(k) \leqslant \bar{\psi}(k) \equiv c\varepsilon^{\nu}[\varepsilon^{b-a-\zeta} \max\{\varepsilon, c\varepsilon^{-\beta}/T_\varepsilon(r)\}] \tag{6.39a}$$

If they are both free, we need to distinguish whether particle k was born after or before $t - T_\varepsilon(r+1)$. We have

$$\psi(k) \leqslant \bar{\psi}(k) = c\varepsilon^{2\nu-\delta}[\varepsilon^{b-a-\zeta} \max\{\varepsilon, c\varepsilon^{-\beta}/T_\varepsilon(r)\}] \times \begin{cases} t & \text{if} \quad t_0 \geqslant t - T_\varepsilon(r+1) \\ T_\varepsilon(r+1) & \text{otherwise} \end{cases} \tag{6.39b}$$

The analysis of $\mathscr{T}_j$, $j<k$, is completely analogous. We now notice that all the $\bar{\psi}(j)$ are bounded by a constant times $\varepsilon^{a-\zeta}$, except for the first case in (6.39b), where $\bar{\psi}(j)$ is bounded by a constant times ε^{-1}. We have

$$F(\pi, \tau) \leqslant c\varepsilon^{-(N_1 + \cdots + N_{r-1}) + (a-\zeta)N_r} \tag{6.40}$$

Since the sum over τ has at most $c\varepsilon^{-1}$ terms, we choose N_r so that

$$\varepsilon^{-1}\varepsilon^{(a-\zeta)N_1} \leqslant \varepsilon^{2n(1+\nu)} \tag{6.41a}$$

$$\varepsilon^{-(N_1 + \cdots + N_{r-1} - 1) + (a-\zeta)N_r} \leqslant \varepsilon^{2n(1+\nu)}, \qquad r > 1 \tag{6.41b}$$

and this concludes the estimates of $v_n(\underline{x}, t)$.

APPENDIX. PROOF OF LEMMA 6.4

As noticed before stating the lemma, (6.11) and (6.13) have already been proved in Section 5, so that we only need to prove (6.12). We fix s as in (6.12); then

$$\rho(q + c_\sigma, e, \sigma, s+1 \mid \eta) = \frac{1}{4} \sum_{\sigma'=1}^{4} \rho(q, e, \sigma', s \mid \eta)$$

Call $x' = (q, e, \sigma')$ and $x = (q, e, \sigma)$; we then need to show that

$$\Delta\rho \equiv |\rho(x', s \mid \eta) - \rho(x, s \mid \eta)| \leqslant \frac{1}{\sqrt{s}} \phi(s) \tag{A.1a}$$

where

$$\phi(s)=c\begin{cases}\varepsilon^{-\beta}/s & \text{for}\quad s<\varepsilon^{-1-\beta}\\ \varepsilon & \text{otherwise}\end{cases}\tag{A.1b}$$

Let $P_{s,s'}(x\to y)$, $s>s'$, be the probability defined in (4.14). Then by (5.1) and (5.2)

$$\Delta\rho\leqslant\Delta_0\rho+\Delta_1\rho\tag{A.2a}$$

$$\Delta_0\rho=\sum_y |P_{s,0}(x\to y)-P_{s,0}(x'\to y)|\,\eta(y)\tag{A.2b}$$

$$\Delta_1\rho=\sum_{\substack{s'\in\varepsilon^{-\nu}\mathbb{N}\\ s'<s}}\sum_y |P_{s,s'}(x\to y)-P_{s,s'}(x'\to y)|\,\rho^2_{(q,s')}\tag{A.2c}$$

In (A.2c), q denotes the position in the state y.

We start from $\Delta_1\rho$. We call s'' the largest collision time smaller than $s/2$. Then, using (6.11),

$$\Delta_1\rho\leqslant\phi(s/2)^2\sum_{\substack{s'\in\varepsilon^{-\nu}\mathbb{N}\\ s'\geqslant s''}}\sum_y |P_{s,s'}(x\to y)-P_{s,s'}(x'\to y)|$$
$$+\sum_{\substack{s'\in\varepsilon^{-\nu}\mathbb{N}\\ s'<s''}}\sum_{z,y}|P_{s,s''}(x\to z)-P_{s,s''}(x'\to z)|\,P_{s'',s'}(z\to y)\,\rho^2_{(q,s')}\tag{A.3}$$

We postpone to the end of this Appendix the proof that for all $t'<s$

$$\sum_z |P_{s,t'}(x\to z)-P_{s,t'}(x'\to z)|\leqslant\frac{c}{(s-t')^{1/2}}\tag{A.4}$$

By (A.4) the first term in (A.3) is bounded by

$$c\varepsilon^{\nu}\sqrt{s}\,\phi(s/2)^2\leqslant\frac{c}{\sqrt{s}}\,\phi(s)$$

(recall that $\nu>\beta$ and that $s\leqslant\varepsilon^{1-\nu+a}$). By (6.13) we can bound the second term in (A.3) by

$$\frac{c}{\sqrt{s}}\sum_{s'\in\varepsilon^{-\nu}\mathbb{N}}\phi(s')\,\phi(s)\leqslant\frac{c'}{\sqrt{s}}\,\phi(s)$$

since the sum over $\phi(s')$ is bounded (because $s'\in\varepsilon^{-\nu}\mathbb{N}$).

For $\Delta_0 \rho$, by (6.13), we get

$$\Delta_0 \rho \leqslant \sum_z |P_{s,s''}(x \to z) - P_{s,s''}(x' \to z)| \, P_{s'',0}(z \to y) \, \eta(y)$$

$$\leqslant \frac{c}{\sqrt{s}} \phi(s'')$$

To complete the proof of (6.12), we need to show the validity of (A.4). We consider two walks starting from x and x'. We couple them independently till the first time when one of the spatial coordinates, say the x coordinate (i.e., along the horizontal axis), of the two walks is the same. After this time the rotations of the two motions are obtained as follows. We choose with probability 1/2 the x or the y coordinates. If the x coordinate is chosen, then with equal probability we choose the value $\sigma = 1, 3$, the same for both particles. If the y coordinate is chosen, then we choose $\sigma = 2, 4$ with equal probability, but independently for the two particles. Since the e-velocities of the two particles are equal, the x coordinate in this coupling remains the same in the two motions. The above rule is valid till the time when also the y coordinates become the same. After that the particles move in the same way. The probability that the two walks are different at time τ is bounded by $c/\sqrt{\tau}$, if they start from two given sites of the same even or odd sublattice of $\mathbb{Z}^2$: if $q = (q_1, q_2) \in \mathbb{Z}^2$, then it is in the even (odd) sublattice if $q_1 + q_2$ is even (odd). Notice that if a particle is in the even sublattice, then at the successive time it is in the odd one, and vice versa. In the case considered in Lemma 6.4 the two particles are in the same sublattice, so that we have proven (A.4) and completed the proof of Lemma 6.4. ∎

ACKNOWLEDGMENTS

We thank Herbert Spohn and Maria Eulalia Vares for useful discussions. This work has been supported by a CNR-MMAIT grant.

REFERENCES

1. B. M. Boghosian and D. Levermore, A cellular automaton for Burgers equation, *Complex Systems* **1**:17–30 (1987).
2. S. Caprino, A. De Masi, E. Presutti, and M. Pulvirenti, A derivation of the Broadwell equation, *Commun. Math. Phys.* **135**:443–465 (1991).
3. R. E. Caflisch and G. C. Papanicolaou, The fluid-dynamical limit of a nonlinear model Boltzmann equation, *Commun. Pure Appl. Math.* **33**:589–616 (1979).
4. A. De Masi, R. Esposito, J. L. Lebowitz, and E. Presutti, Hydrodynamics of stochastic HPP cellular automata, *Commun. Math. Phys.* **125**:127–145 (1989).

5. A. De Masi and E. Presutti, Mathematical methods for hydrodynamic limits, Vol. 1501 (Springer-Verlag, 1991).
6. J. Hardy, O. de Pazzis, and Y. Pomeau, Molecular dynamics of classical lattice gas: Transport properties and time correlation functions, *Phys. Rev. A* **13**:1949 (1976).
7. R. Illner and M. Pulvirenti, Global validity of the Boltzmann equation for a two- and three-dimensional rare gas in a vacuum: Erratum and improved results, *Commun. Math. Phys.* **121**:143–166 (1989).
8. O. E. Lanford III, Time evolution of large classical systems, in *Lecture Notes in Physics*, Vol. 38, J. Moser, ed. (1975), pp. 1–111.
9. R. Lang and X. X. Nguyen, Smoluchowski's theory of coagulation in colloids holds rigorously in the Boltzmann–Grad limit, *Z. Wahrsch. Verw. Geb.* **54**:227–280 (1980).
10. T. M. Liggett, *Interacting Particle Systems* (Springer-Verlag, Berlin, 1985).
11. H. Spohn, *Large Scale Dynamics of Interacting Particles* (Springer-Verlag, to appear).
12. A. Sznitman, *Topics in Propagation of Chaos* (Ecole d'Eté de Probabilitées de Saint Flour, 1989).
13. K. Uchiyama, On the Boltzmann–Grad limit for the Broadwell model of the Boltzmann equation, *J. Stat. Phys.* **52**:331–355 (1988).