The Annals of Probability
1997, Vol. 25, No. 1, 115–132

STOCHASTIC PARTICLE APPROXIMATIONS FOR GENERALIZED BOLTZMANN MODELS AND CONVERGENCE ESTIMATES

By Carl Graham and Sylvie Méléard

École Polytechnique and Université Paris 6

We specify the Markov process corresponding to a generalized mollified Boltzmann equation with general motion between collisions and nonlinear bounded jump (collision) operator, and give the nonlinear martingale problem it solves. We consider various linear interacting particle systems in order to approximate this nonlinear process. We prove propagation of chaos, in variation norm on path space with a precise rate of convergence, using coupling and interaction graph techniques and a representation of the nonlinear process on a Boltzmann tree. No regularity nor uniqueness assumption is needed. We then consider a nonlinear equation with both Vlasov and Boltzmann terms and give a weak pathwise propagation of chaos result using a compactness–uniqueness method which necessitates some regularity. These results imply functional laws of large numbers and extend to multitype models. We give algorithms simulating or approximating the particle systems.

1. Framework and main results. The Boltzmann equation describes the evolution of the limit density $f(t, x, v)$ of molecules of a rarefied gas at time t, position x and speed v; f is positive and $\int f(t, x, v)\,dx\,dv = 1$. It is given by

$$\begin{aligned} &\partial_t f(t, x, v) + v\cdot\nabla_x f(t, x, v) \\ &\quad = \int_{S^2} dn \int_{\mathbb{R}^3} dw\, q(v, w, n)\big(f(t, x, v^*) f(t, x, w^*) \\ &\qquad\qquad - f(t, x, v) f(t, x, w)\big), \end{aligned} \tag{1.1}$$

where x and v are in $\mathbb{R}^3$, and $v^* = v + ((w - v)\cdot n)n$ and $w^* = w + ((v - w)\cdot n)\, n$ represent the post-collisional velocities of two particles of speeds v and w having collided in a position in which their centers are on a line of direction given by the unit vector n (called the impact parameter). For more details, in particular on the form of the cross-section q, see Cercignani [2].

This equation has an intrinsic probabilistic interpretation consistent with its derivation from the underlying particle dynamics. Integrating test functions ϕ with respect to (1.1), transposing the operators so as to have them

---

Received February 1995; revised February 1996.

AMS 1991 *subject classifications*. 60K35, 60F17, 47H15, 65C05, 76P05, 82C40, 82C80.

Key words and phrases. Boltzmann equation, nonlinear diffusion with jumps, random graphs and trees, coupling, propagation of chaos, Monte Carlo algorithms.

carry on ϕ and setting $f_t(dx, dv) = f(t, x, v)\, dx\, dv$, we rewrite (1.1) as

$$\begin{aligned}\partial_t\langle \phi, f_t\rangle - \langle v\cdot\nabla_x\phi(x,v), f_t\rangle &= \int (\phi(x,v^*) - \phi(x,v))\, q(v,w,n)\, f(t,x,v) \\ &\quad \times f(t,x,w)\, dx\, dv\, dw\, dn \\ &= \left\langle \int (\phi(x,v^*) - \phi(x,v))\, q(v,w,n)\, dn \right. \\ &\qquad \left. \times f(t,x,w)\, dw,\, f_t(dx,dv)\right\rangle. \end{aligned} \tag{1.2}$$

This is the evolution equation for the flow of marginals of a nonlinear Markov process. The collision term depends on the law of the process itself, but locally through the density at x of the law. The mapping $f_t(dx, dw) \mapsto f(t, x, w)\, dw$ is not continuous, and we cannot bound $f(t, x, w)\, dw$ nor the jump rate because of possible accumulation of particle density at x. This leads to an ill-defined process. This difficulty is of the same nature as the one encountered by the analysts. There is an extensive physical and mathematical literature, but even existence results are rare and under restrictive assumptions, the best ones being in DiPerna and Lions [4].

Hence analysts, numericists and physicists consider a simplified tractable model: a mollified bounded-rate equation in which the interaction is delocalized. In (1.2), $f(t, x, w)$ is approximated by $\int I(x, y) f(t, y, w)\, dy$ and the whole measure $f_t(dx, dw) = f(t, x, w)\, dx\, dw$ intervenes instead of the local measure $f(t, x, w)\, dw$. The regularizing kernel $I(x, y)$ approximating the Dirac mass may be given by a grid method as in Perthame [15] and Pulvirenti, Wagner and Rossi [17], or be made more regular if needed. We obtain the mollified Boltzmann equation with delocalized cross section $q(x, v, y, w, n) = I(x, y) q(v, w, n)$:

$$\begin{aligned}\partial_t\langle \phi, f_t\rangle - \langle v\cdot\nabla_x\phi(x,v), f_t\rangle &= \left\langle \int (\phi(x,v^*) - \phi(x,v))\, q(x,v,y,w,n)\, dn, \right. \\ &\qquad \left. f_t(dx,dv)\, f_t(dy,dw)\right\rangle. \end{aligned} \tag{1.3}$$

Heuristically, the particle repeatedly samples for collision partners using the common law of the particles, but the equation does not say what happens to these collision partners. Hence the different particle interpretations derived by Bird, Nanbu and others. The physical binary symmetric collision derivation may lead us to write the collision term as

$$\begin{aligned}&\left\langle \int \tfrac{1}{2}(\phi(x,v^*) - \phi(x,v) + \phi(y,w^*) - \phi(y,w)) \right. \\ &\qquad \left. \times q(x,v,y,w,n)\, dn,\, f_t(dx,dv)\, f_t(dy,dw)\right\rangle. \end{aligned} \tag{1.4}$$

The form of the cross section gives causes of divergence of the jump rate due to large speeds and grazing collisions. Except in the spatially homogeneous hard-sphere case (cf. [18]), mathematicians and physicists usually make the cut-off assumption $\sup_{x, y, v, w} \int q(x, v, y, w, n)\, dn < \infty$.

It is important to approximate solutions to equation (1.3) and obtain error estimates. Because of its complexity and the probabilistic interpretation of the collision term as the jump operator of a Markov process, Monte Carlo algorithms based on stochastic interacting particles are widely used. Rigorous results on the famous Bird and Nanbu algorithms appear for instance in Wagner [21, 22], Pulvirenti, Wagner and Rossi [17] and Babovsky and Illner [1]. Several numerical methods are also discussed in Nanbu [13], Illner and Neunzert [9], Neunzert, Gropengeisser and Struckmeier [14], Chauvin [3] and Perthame [15].

We get new and better results using radically different techniques. We prove propagation of chaos in variation norm on the path space, with precise estimates of the proper order for fluctuations. This strong result gives asymptotics for sojourn times, hitting times and other functionals of the whole path and immediately implies the convergence of the flow of time marginals and a functional law of large numbers. We need no regularity assumption on the noninteracting motion no on the collision kernel which is very general (not parameterized) but bounded, and we do not use a uniqueness result.

Let us describe more precisely the contents of the paper. We wish to take into account more complex physical models with boundary conditions, energy exchanges, external fields, different species and so on. We thus define a Markov process with a general linear Markov evolution instead of the free flow and with any bounded jump measure with linear dependence on the law of the process, and its corresponding nonlinear martingale problem. This linear dependence on the law allows us to define a family of approximating binary mean-field interacting n-particle systems (for which the jump measure is a sum of terms in which only pairs of particles intervene) that gives a unified treatment of the Bird and Nanbu models. We prove Theorem 3.1.

THEOREM 1.1. *The law on path space on* $[0, T]$ *of a subsystem of size* k *of the interacting system converges in variation norm to the* k*-fold product of the law of a Boltzmann process constructed on a Boltzmann tree when* n *goes to infinity, with a precise* $O(1/n)$ *rate of convergence. This implies convergence in probability of the empirical measure to the Boltzmann law.*

The proof uses a pathwise representation of the particle systems with binary interaction graphs and a probabilistic argument called coupling. Graham and Méléard [8] use such ideas for pure jump processes, which we adapt here to general markovian evolution (without interaction) between the jumps. We obtain robust results which extend to multitype Boltzmann equations and interacting particle systems representing a mixture of different species as in realistic physical models. These representations give an effective simulation of the n-particle system using an acceptance–rejection or fictitious collision method. The empirical measure of this system approximates the Boltzmann law.

In the last part of the paper we consider more complex physical models mixing strong (Boltzmann) and weak (Vlasov) interaction, as in the Fokker–Planck equation which takes into account grazing collisions and in the Boltzmann–Maxwell equation for charged particles. Both the diffusion and the jump operators depend on the law, possibly nonlinearly but with some regularity assumptions. We introduce appropriate interacting particle systems. In this case, interaction graph methods do not work since there is interaction between collisions. We obtain Theorem 4.5.

THEOREM 1.2. *There is propagation of chaos for weak convergence on path space with the Skorokhod topology, which is equivalent to convergence in probability of the empirical measure.*

The proof uses the equivalence for exchangeable systems of propagation of chaos for weak convergence and of convergence in law of the empirical measures, and deduces the latter using compactness, martingale problems and uniqueness of the limiting martingale problem with nonlinear integro-differential operator. The complexity of the Skorokhod topology renders both the new result and Theorem 4.6 difficult. Theorem 4.6 states a nontrivial general result from which we deduce that Theorem 4.5 implies convergence for the flow of marginals and a functional law of large numbers.

2. A generalized Boltzmann equation and its approximating interacting particle systems. We generalize (1.3) in order to consider more complex models with complicated spatial evolution between collisions including boundary conditions. We consider the equation

$$\partial_t\langle \phi, \tilde{P}_t\rangle = \left\langle L\phi + K\phi\left(\cdot, \tilde{P}_t\right), \tilde{P}_t\right\rangle. \tag{2.1}$$

ASSUMPTION 2.1. The free evolution is given by a very general markovian operator L on $\mathbb{R}^d$ acting on a sufficiently large domain $\mathrm{Dom}(L)$ of $L^\infty(\mathbb{R}^d)$.

ASSUMPTION 2.2. The collision operator is a bounded operator K on $\mathbb{R}^d$ such that the following hold.

(i) For z in $\mathbb{R}^d$ and a probability measure p on $\mathbb{R}^d$, the jump amplitudes h leading from the precollisional state z to the postcollisional state $z^* = z + h$ are given by a positive measure $m(z, p)$, and for bounded ϕ, $K\phi(z, p) = \int(\phi(z+h) - \phi(z))m(z, p, dh)$.

(ii) m depends linearly on p, and we set $m(z, p, dh) = \langle \mu(z, a, dh), p(da)\rangle$.

(iii) m, or equivalently μ, is uniformly bounded.

Since m depends linearly on p as in Section 1, (2.1) can be written

$$\begin{aligned}\partial_t\langle \phi, \tilde{P}_t\rangle - \langle L\phi, \tilde{P}_t\rangle &\\ = \langle K\phi, \tilde{P}_t \otimes \tilde{P}_t\rangle &\\ = \left\langle \int (\phi(z+h) - \phi(z))\mu(z, a, dh), \tilde{P}_t(dz)\tilde{P}_t(da)\right\rangle.\end{aligned} \tag{2.2}$$

Dom(L) is usually $C_b^2(\mathbb{R}^d)$, or $C_b^1(\mathbb{R}^d)$ for a first-order operator as in Section 1, but is more complex in case of boundary conditions discussed in Cercignani [2], Neunzert, Gropengeisser and Struckmeier [14] and Babovsky and Illner [1]. An essential fact in all that follows is that adding mass at zero to m does not change K. In Section 1, $d = 6$, $z = (x, v)$ with x and v in $\mathbb{R}^3$, $L\phi(x, v) = v \cdot \nabla_x \phi(x, v)$ and $m(x, v, p)$ is the image of $\int_{y \in \mathbb{R}^3} q(x, v, y, w, n) p(dy\, dw)\, dn$ under $(w, n) \mapsto ((w - v) \cdot n)n$.

Equation (2.1) or (2.2) is satisfied by the flow of marginals of any solution to the following nonlinear martingale problem.

DEFINITION 2.3. Let Z be the canonical process on the Skorokhod space $D(\mathbb{R}_+, \mathbb{R}^d)$. A solution to the (Boltzmann) nonlinear martingale problem is a $\tilde{P} \in P(D(\mathbb{R}_+, \mathbb{R}^d))$ such that

$$\phi(Z_t) - \phi(Z_0) - \int_0^t L\phi(Z_s) + K\phi\big(Z_s, \tilde{P}_s\big)\, ds = M_t^\phi \tag{2.3}$$

is a $\tilde{P}$-martingale for any ϕ in Dom(L). $\tilde{P}_s = \tilde{P} \circ Z_s^{-1} = \text{law}(Z_s)$ is the nonlinearity.

REMARK 2.4. The martingale problem (2.3) gives much more information than (2.1) or (2.2) and enables us to consider multidimensional time marginals and quantities based on the whole process as extrema or hitting times which are of interest for gas dynamics.

In (2.1) the nonlinearity is a usual mean-field one as in the Vlasov models, and the obvious choice of a mean-field approximating system gives the Nanbu system defined below in (2.5). This is far from the physical interpretation of the system and does not preserve momentum. For any symmetrical jump kernel $\hat{\mu}(z, a, dh, dk)$, giving the joint postcollisional states of two particles, such that $\mu(z, a, dh)$ is the marginal $\hat{\mu}(z, a, dh \times \mathbb{R}^d)$ up to mass at zero, (2.2) can be written

$$\begin{aligned}\partial_t\langle \phi, \tilde{P}_t\rangle - \langle L\phi, \tilde{P}_t\rangle &\\ = \Big\langle \int \tfrac{1}{2}(\phi(z+h) - \phi(z) + \phi(a+k) - \phi(a)) &\\ \times \hat{\mu}(z, a, dh, dk), \tilde{P}_t(dz)\tilde{P}_t(da)\Big\rangle.\end{aligned} \tag{2.4}$$

The joint measure $\hat{\mu}$ cannot be deduced from the Boltzmann equation and involves consideration of an underlying physical process or some kind of choice. It leads to the Bird model defined in (2.6).

We define a family of mean-field interacting particle systems which approximate the solution to the nonlinear martingale problem, corresponding to different choices of the joint jump measure $\hat{\mu}$ in (2.4), for which existence and uniqueness is simple. Let $\mathbf{z}^n = (z_1, \dots, z_n)$ be the generic point in $(\mathbb{R}^d)^n$, $Z^{n1}, \dots, Z^{nn}$ be the canonical processes on $D(\mathbb{R}_+, (\mathbb{R}^d)^n)$, L_i be the extension of L on $(\mathbb{R}^d)^n$ acting only on the variable z_i. Define the empirical measures $\bar{z}^{ni} = (1/(n-1))\sum_{j \neq i,\ j=1}^n \delta_{z_j}$ and $\bar{Z}^{ni} = (1/(n-1))\sum_{j \neq i,\ j=1}^n \delta_{Z^{nj}}$ and the mapping $\mathbf{e}_i$: $h \in \mathbb{R}^d \mapsto \mathbf{e}_i \cdot h = (0, \dots, 0, h, 0, \dots, 0) \in (\mathbb{R}^d)^n$ with h at the ith place. We consider functions Φ on $(\mathbb{R}^d)^n$ that are in $\mathrm{Dom}(L_i)$ for all $1 \le i \le n$.

The Nanbu, or simple mean-field, model. The generator of the particle system is

$$\begin{aligned} &\sum_{i=1}^n L_i \Phi(\mathbf{z}^n) + \sum_{i=1}^n \int \big(\Phi(\mathbf{z}^n + \mathbf{e}_i \cdot h) - \Phi(\mathbf{z}^n)\big) m\big(z_i, \bar{z}^{ni}, dh\big) \\ &\quad = \sum_{i=1}^n L_i \Phi(\mathbf{z}^n) + \frac{1}{n-1} \sum_{1 \le i \ne j \le n} \int \big(\Phi(\mathbf{z}^n + \mathbf{e}_i \cdot h) \\ &\qquad\qquad - \Phi(\mathbf{z}^n)\big) \mu\big(z_i, z_j, dh\big). \end{aligned} \tag{2.5}$$

The Bird and other binary mean-field models. The generator is

$$\begin{aligned} \sum_{i=1}^n L_i \Phi(\mathbf{z}^n) + \frac{1}{n-1} \sum_{1 \le i \ne j \le n} \int \tfrac{1}{2}\big(\Phi\big(\mathbf{z}^n + \mathbf{e}_i \cdot h + \mathbf{e}_j \cdot k\big) - \Phi(\mathbf{z}^n)\big) \\ \times \hat{\mu}\big(z_i, z_j, dh, dk\big). \end{aligned} \tag{2.6}$$

In numerical practice one often allows self-interactions and replaces $\bar{z}^{ni}$ by $\bar{z}^n = (1/n)\sum_{j=1}^n \delta_{z_j}$ in (2.5) and $n-1$ by n and takes the sum over $1 \le i$, $j \le n$ in (2.5) and (2.6).

REMARK 2.5. In the specific Bird model, $\hat{\mu}(x, v, y, w)$ is the image of $q(x, v, y, w, n)\, dn$ by the mapping $n \mapsto (((w - v) \cdot n)n, ((v - w) \cdot n)n)$. This physical choice preserves momentum.

REMARK 2.6. Taking $\hat{\mu}(z, a, dh, dg) = \mu(z, a, dh) \otimes \delta_0(dg) + \delta_0(dh) \otimes \mu(a, z, dg)$ in (2.6) gives the Nanbu model (2.5).

The action of (2.5) and (2.6) on functions of a variable z_i coincide with any $\hat{\mu}$. For ϕ in $\mathrm{Dom}(L)$,

$$\begin{aligned} &\phi\big(Z_t^{ni}\big) - \phi\big(Z_0^{ni}\big) - \int_0^t L\phi\big(Z_s^{ni}\big) \\ &\quad + \int \big(\phi\big(Z_s^{ni} + h\big) - \phi\big(Z_s^{ni}\big)\big) m\big(Z_s^{ni}, \bar{Z}_s^{ni}, dh\big)\, ds = M_t^{\phi ni} \end{aligned} \tag{2.7}$$

is a martingale. If $A^{\phi ni}$ is the Doob–Meyer process of $L(Z^{ni})$ acting on ϕ, the Doob–Meyer bracket is

$$\langle M^{\phi ni} \rangle_t = A_t^{\phi ni} + \int_0^t \int \big(\phi\big(Z_s^{ni} + h\big) - \phi\big(Z_s^{ni}\big)\big)^2 m\big(Z_s^{ni}, \bar{Z}_s^{ni}, dh\big)\, ds. \tag{2.8}$$

For (2.5), $\langle M^{\phi n i}, M^{\phi n j}\rangle = 0$ for $i \neq j$ since there are no simultaneous jumps. Generally for (2.6)

$$\begin{aligned}
&\langle M^{\phi n i}, M^{\phi n j}\rangle_t \\
(2.9)\qquad &= \frac{1}{n-1}\int_0^t\int \big(\phi\big(Z_s^{ni}+h\big)-\phi\big(Z_s^{ni}\big)\big)\big(\phi\big(Z_s^{nj}+k\big)-\phi\big(Z_s^{nj}\big)\big) \\
&\qquad\qquad \times \hat{\mu}\big(Z_s^{ni}, Z_s^{nj}, dh, dk\big)\,ds.
\end{aligned}$$

Since at most two particles jump at the same time, (2.7), (2.8) and (2.9) characterize the martingale problem corresponding to (2.6). The factor $1/(n-1)$ in (2.9) is the key to the decoupling of the particles when n goes to infinity in the martingale problem methods and is the sign of weak interaction. All this involves the two-body empirical measure $\overline{Z}^{(2)n} = (1/n(n-1))\sum_{1\le i\neq j\le n}\delta_{Z^{ni},Z^{nj}}$.

3. Propagation of chaos in variation norm with estimates. Propagation of chaos is a probabilistic limit result in which the law of a fixed number k of particles in an interacting system converges to the k-fold product of a limit law as the number of particles goes to infinity. It is usually stated and proved in weak topology, but we shall prove it for the strong variation norm on path space. $|\ |_T$ is the variation norm on the space of signed bounded measures on $D([0,T],\mathbb{R}^d)$, and $\overline{Z}^n = (1/n)\sum_{i=1}^n \delta_{Z^{ni}}$.

THEOREM 3.1. *Assume* $(Z_0^{ni})_{1\le i\le n}$ *i.i.d. of law* P_0, *Assumptions* 2.1 *and* 2.2, $\sup_{x,a}|\hat{\mu}(z,a)| \le \Lambda$.

(i) *We then have propagation of chaos*: *for given* T *and* k

$$\left|\text{law}(Z^{n1},\dots,Z^{nk}) - \text{law}(Z^{n1})^{\otimes k}\right|_T \le 2k(k-1)\frac{\Lambda T + \Lambda^2 T^2}{n-1}$$

and there exists a law $\tilde{P}$ *defined uniquely using a Boltzmann tree such that*

$$\left|\text{law}(Z^{ni}) - \tilde{P}\right|_T \le 6\frac{\exp(\Lambda T)-1}{n+1},$$

which solves the nonlinear martingale problem (2.3).

(ii) $\text{law}(Z^{n1},\dots,Z^{nk})$ *converges weakly to* $\tilde{P}^{\otimes k}$ *for the Skorokhod topology on* $D(\mathbb{R}_+,(\mathbb{R}^d)^k)$.

(iii) *The empirical measure* $\overline{Z}^n$ *converges in probability to* $\tilde{P}$ *in* $P(D(\mathbb{R}_+,\mathbb{R}^d))$ *for the weak convergence for the Skorokhod metric on* $D(\mathbb{R}_+,\mathbb{R}^d)$, *with* $O(1/\sqrt{n})$ *estimates on* $[0,T]$.

PROOF. Graham and Méléard [8] prove such a result for pure jump processes. We give a compact proof adapted to our model with general markovian motion given by L. We separate the main steps.

3.1. *A sample path representation using interaction graphs.* We give for $k \le n$ and distinct $i_1, \dots, i_k$ in $\{1, \dots, n\}$ a pathwise representation of $(Z^{ni_1}, \dots, Z^{ni_k})$ on $[0, T]$. At time T the evolution of the state of a process will have been directly affected by other processes with which it has collided, and recursively these processes will have evolved according to collisions with other processes which thus influence indirectly the first process. We now describe this construction on $[0, T]$ in reverse time so that we can build the processes with the least amount of superfluous knowledge. This will define the past history of a particle. If the histories of two particles become disjoint when n goes to infinity, a coupling argument will show that they become independent.

This is described by interaction graphs, which are marked random subsets of $[0, T] \times \{1, \dots, n\}$. Considering (2.6), we introduce i.i.d. Poisson processes $(N_{ij})_{1 \le i < j \le n}$ of rate $\Lambda/(n-1)$, where $|\hat{\mu}| \le \Lambda$, and set $N_{ij} = N_{ji}$ for notational convenience. For $i \ne j$, N_{ij} is a random clock giving the epochs at which Z^{ni} and Z^{nj} are possibly authorized to jump simultaneously and interact. We imagine time as being vertical and directed upwards and the indices of particles as being on a horizontal level. We work our way in reverse time from time T to 0 to build a graph rooted on the given subset $\{i_1, \dots, i_k\}$. Every time we encounter a jump of a Poisson process N_{ij} for an i already in the graph, we include the index j in the graph at that time, and recursively so, and mark the graph with variables for jumps and free flow. The branching is binary and deterministic, given the Poisson processes; see Figure 1.

Once an index is selected, the whole vertical line from the time of selection down to 0 is included in the graph, and we proceed recursively from the selected indices. We do not have a tree since a particle may influence another several times. We can thus build down, from time T, an interaction graph $G^n_{i_1 \cdots i_k}$ rooted on $\{i_1, \dots, i_k\}$.

Once we reach time 0, we construct a pathwise representation of the process $(Z^{ni_1}, \dots, Z^{ni_k})$ in direct time on $[0, T]$. The interaction graph represents all the information necessary to construct the process, and we only need to consider at time t the indices appearing then in the graph. We use the independent variables Z_0^{ni} of law P_0 at time 0. Then we follow the indices in

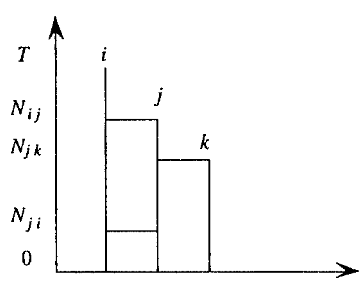

FIG. 1. *Interaction graph rooted at i.*

the graph. In between jumps of the N_{ij}, we use the pathwise representation of the process generated by L. At the jump times of N_{ij} we compute $\hat{\mu}$ at the position reached by (Z^{ni}, Z^{nj}); with probability $1 - |\hat{\mu}|/\Lambda$, we do nothing and with probability $|\hat{\mu}|/\Lambda$ we choose the joint amplitude of jumps according to the law $\hat{\mu}/|\hat{\mu}|$. All this must be done independently. This gives the correct evolution for the Markov process since $\Lambda(\hat{\mu}/\Lambda + (1 - \hat{\mu}/\Lambda)\delta_0)$ and $\hat{\mu}$ coincide up to mass at zero.

3.2. *The coupling.* The chaos property asserts that a fixed finite number of particles become independent as the total number grows. We construct a set of independent particles, couple it to an interacting system and show that the two do not differ much. We detail this for two indices i and j. We take two independent copies distinguished using indices i and j of the Poisson processes, jump variables, Markov paths corresponding to L and initial values. We construct two independent interaction graphs $G_i^{i,n}$ and $G_j^{j,n}$ and two independent processes Z_i^{ni} and Z_j^{nj}.

We build an interaction graph G_{ij}^n in such a way that the subgraph G_i^n stemming from i is as close as possible to $G_i^{i,n}$ and G_j^n to $G_j^{j,n}$. We must choose for any pair $\{k, l\}$ of indices between N_{kl}^i and N_{kl}^j to get N_{kl}. We set $N_{ik} = N_{ik}^i$ and $N_{jl} = N_{jl}^j$ if $k \neq j$ and $l \neq i$. There is a problem between N_{ij}^i and N_{ij}^j which we resolve by a fair toss-up. We still have to set priorities between the N_{kl}^i and N_{kl}^j for distinct i, j, k and l. The special properties of the Poisson process allow us to do this pathwise in reverse time. If $G_i^{i,n}$ is the first (in reverse time) to intersect $\{k, l\}$, we set $N_{kl} = N_{kl}^i$; if $G_j^{j,n}$ is the first, we set $N_{kl} = N_{kl}^j$ (there cannot be a tie because of independence). We choose the initial values and jump variables among the two independent copies accordingly.

Thus the subgraphs G_i^n and G_j^n will grow like $G_i^{i,n}$ and $G_j^{j,n}$ until we encounter a jump time of an N_{kl}^i and N_{kl}^j with $\{k, l\}$ intersecting both $G_i^{i,n}$ and $G_j^{j,n}$, when the priority rule will force us either to use it or discard it for both G_i^n and G_j^n. This introduces a difference between either G_i^n and $G_i^{i,n}$ or G_j^n and $G_j^{j,n}$. We denote this event by $A_{ij}^n = \{G_i^n \neq G_i^{i,n}\} \cup \{G_j^n \neq G_j^{j,n}\} = \{G_i^{i,n} \cap G_i^{j,n} \neq \varnothing\}$, and on the contrary event, the process (Z^{ni}, Z^{nj}) built on G_{ij}^n is equal to (Z_i^{ni}, Z_j^{nj}).

Naturally $\operatorname{law}(Z^{ni}) = \operatorname{law}(Z_i^{ni})$ and $\operatorname{law}(Z^{nj}) = \operatorname{law}(Z_j^{nj})$, and thus $\|\operatorname{law}(Z^{ni}, Z^{nj}) - \operatorname{law}(Z^{ni}) \otimes \operatorname{law}(Z^{nj})\|_T \leq 2P(A_{ij}^n)$, which leads us to estimate $P(A_{ij}^n)$. Similarly

$$\begin{aligned}\left\| \operatorname{law}(Z^{ni_1}, \dots, Z^{ni_k}) - \operatorname{law}(Z^{ni_1}) \otimes \cdots \otimes \operatorname{law}(Z^{ni_k}) \right\|_T &\leq 2P\left(\bigcup_{1 \leq p < q \leq k} A_{i_p i_q}^n \right) \\ &\leq k(k-1)P\left(A_{ij}^n\right).\end{aligned}$$

3.3. *Estimates on interaction chains and convergence.* A_{ij}^n happens only if we encounter a jump of one of a pair of conflicting Poisson processes before reaching time 0. There must then exist an interaction chain between i and j, an event we now describe. The simplest case is a jump of N_{ij}^i or N_{ij}^j on $[0, T]$, called a direct interaction.

Going from T to 0, there are indices $i_1, \dots, i_m$ and $j_1, \dots, j_p$, such that in $G_i^{i,n}$, i branched on i_1, then i_1 on i_2, and so on to i_m, and similarly, $G_j^{j,n}$ branched successively on $j_1, j_2, \dots, j_p$, and either i_m branched on j_p in $G_i^{i,n}$ or j_p on i_m in $G_j^{j,n}$. This last step may be taken from either side, hence at twice the branching rate, and is a direct interaction between i_m and j_p. We choose an interaction chain with the least number of indices involved, in which $i, j, i_1, \dots, i_m, j_1, \dots, j_p$ are distinct, and define the interaction chain length as $m + p + 1$. See Figure 2.

If $Q_T^n(q)$ is a bound on the probability of occurrence of a chain reaction of length q, then $Q_T^n = \sum_{q \ge 1} Q_T^n(q)$ is a bound on $P(A_{ij}^n)$. We evaluate $Q_T^n(q)$ by induction; $q = 1$ corresponds to direct interaction: either N_{ij}^i or N_{ij}^j must jump in $[0, T]$, an event of probability $Q_T^n(1) = 1 - \exp(-2\Lambda T/(n-1)) \le 2\Lambda T/(n-1)$.

Assume we have a bound $Q_T^n(q-1)$ for $q \ge 2$. Then for a chain reaction to happen, there must first (backwards from T) be the birth of a new branch at either i or j after a wait of t, and this new branch must be joined in a time $T - t$ by an interaction chain of length $q - 1$ to the one of i and j which did not branch first. The maximal rate for this new branch is 2Λ and thus

$$Q_T^n(q) = \int_0^T Q_{T-t}^n(q-1) 2\Lambda \exp(-2\Lambda t)\, dt \tag{3.1}$$

and $Q^n(q) = Q^n(q-1) * e_{2\Lambda} = Q^n(1) * e_{2\Lambda}^{*(q-1)}$, where e_θ denotes the exponential density of parameter θ, $e_\theta^{*k}(t)$ is the gamma function $\theta^k(t^{k-1}/(k-1)!)e^{-\theta t}$ and we obtain

$$\begin{aligned} Q_T^n &= Q_T^n(1) + 2\Lambda \int_0^T Q_{T-t}^n(1) \sum_{q \ge 2} \frac{(2\Lambda t)^{q-2}}{(q-2)!} \exp(-2\Lambda t)\, dt \\ &= Q_T^n(1) + 2\Lambda \int_0^T \left(1 - \exp\left(-2\frac{\Lambda(T-t)}{n-1}\right)\right) dt \\ &= (n-2)\left(\exp\left(-\frac{2\Lambda T}{n-1}\right) - 1 + \frac{2\Lambda T}{n-2}\right) \\ &\le 2\frac{\Lambda T + \Lambda^2 T^2}{n-1} \end{aligned} \tag{3.2}$$

which gives us the first bound in Theorem 3.1.

3.4. *The limit Boltzmann tree.* A similar coupling argument between the interaction graph issued from one index for given n and a limit Boltzmann tree, where the links are taken among an infinite supply of independent similar links, shows that the law of one of the interacting processes converges as n goes to infinity to the law $\tilde{P}$ of a process constructed similarly on the tree.

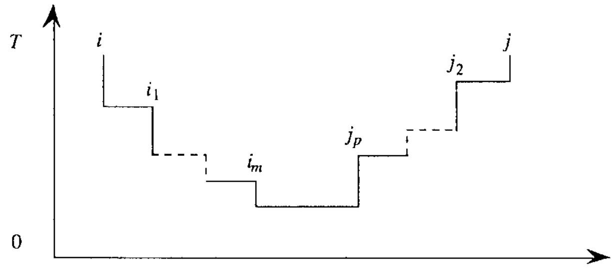

FIG. 2. *Interaction chain between i and j.*

For given n and index i, we build an interaction tree T_i^n inspired from G_i^n. Each branch of the tree has a distinct label, which is a finite sequence in $\{1, \dots, n\}$ giving the successive filiation from the root. We need independent Poisson processes of rate $\Lambda/(n-1)$, jump variables. Markov paths corresponding to L and initial values, corresponding to all branches, which we index by the labels. We then construct a process $\tilde{Z}^{ni}$ using T_i^n, similar to the way we constructed Z^{ni} using G_i^n. This is the Boltzmann process, constructed without self-interactions.

For the second bound in Theorem 3.1 we have to couple an interaction graph G_i^n to the tree T_i^n, and the process in the n-particle system Z^{ni} to $\tilde{Z}^{ni}$. To obtain the graph from the tree, we simply retain the last index in the labelling of the branches as we build the graph in reverse time, and we need priority rules to decide between conflicting Poisson processes. We do as for the graph coupling, and a conflict happens if two branches in the tree are linked by an interaction chain, an event we call an interaction loop. We obtain the second bound of Theorem 3.1 using a recursive evaluation of the number of such pairs of branches and the bound (3.2). We shall not further detail this (see Section 5 in [8]).

3.5. *A law of large numbers on the empirical measure.*

$$
\begin{aligned}
E\left(\langle \phi, \bar{Z}^n - \tilde{P}\rangle^2\right) &= E\left(\left(\frac{1}{n}\sum_{i=1}^{n} \phi(Z^{ni}) - \langle \phi, \tilde{P}\rangle\right)^2\right) \\
&= \frac{1}{n} E\left(\left(\phi(Z^{n1}) - \langle \phi, \tilde{P}\rangle\right)^2\right) \\
&\quad + \frac{n-1}{n} E\left(\left(\phi(Z^{n1}) - \langle \phi, \tilde{P}\rangle\right)\left(\phi(Z^{n2}) - \langle \phi, \tilde{P}\rangle\right)\right) \\
&= E\left(\phi(Z^{n1})\phi(Z^{n2})\right) - 2\langle \phi, \tilde{P}\rangle E\left(\phi(Z^{n1})\right) \\
&\quad + \langle \phi, \tilde{P}\rangle^2 + O(1/n) \\
&= O(1/n)
\end{aligned}
\tag{3.3}
$$

uniformly on T and $\|\phi\|_\infty$ for ϕ on $D([0,T],\mathbb{R}^d)$. The Skorokhod space is Polish and there is a countable convergence determining family $(\phi_i)_{i\geq 1}$ with $\|\phi_i\|_\infty = 1$: m^n converges weakly to m in $P(D(\mathbb{R}_+,\mathbb{R}^d))$ if and only if $\langle \phi_i, m^n\rangle$ converges to $\langle \phi_i, m\rangle$ for all $i \geq 1$. Then $d(m,m')^2 = \sum_{i\geq 1}(1/i^2)\langle \phi_i, m-m'\rangle^2$ defines a weak convergence metric, and $E(d(\overline{Z}^n,\tilde{P})^2) = O(1/n)$, implying convergence in probability with $O(1/\sqrt{n})$ rate. This convergence result extends easily to $\mathbb{R}_+$.

3.6. *The nonlinear martingale problem.* The convergence of the law of one process Z^{ni} being in variation norm, there is no need of regularity to prove that the Boltzmann process solves the nonlinear martingale problem. We consider the martingale problem (2.7) satisfied by one particle in the interacting system together with (2.8) and (2.9) and the convergence of the empirical measures. We use a characterization of martingales suited for taking the limit to show that the limit law solves the limit nonlinear martingale problem. Details on these convergence techniques are given in the next section. This ends the proof. □

REMARK 3.2. (i) It is immediate that

$$\sup_{0\leq t\leq T}\left|\operatorname{law}\left(Z_t^{n1},\dots,Z_t^{nk}\right) - \operatorname{law}\left(Z_t^{n1}\right)^{\otimes k}\right| \leq \left|\operatorname{law}(Z^{n1},\dots,Z^{nk}) - \operatorname{law}(Z^{n1})^{\otimes k}\right|_T,$$

$$\sup_{0\leq t\leq T}\left|\operatorname{law}\left(Z_t^{ni}\right) - \tilde{P}_t\right| \leq \left|\operatorname{law}(Z^{ni}) - \tilde{P}\right|_T.$$

(ii) If the marginal of rank k of the law of the n-particle system at time t has a density $f_{k,t}^n$ on $(\mathbb{R}^d)^k$ and if the limit law at time t has density f_t (this is the case if L contains a hypoelliptic operator or if the initial values have a density and L does not tend to concentrate the laws), then

$$\left|\operatorname{law}\left(Z_t^{n1},\dots,Z_t^{nk}\right) - \operatorname{law}\left(Z_t^{n1}\right)^{\otimes k}\right| = \left\| f_{k,t}^n - \left(f_{1,t}^n\right)^{\otimes k}\right\|_1,$$

$$\left|\operatorname{law}\left(Z_t^{ni}\right) - \tilde{P}_t\right| = \| f_{1,t}^n - f_t\|_1$$

and we can use the estimates in Theorem 3.1 to get results comparable to those in Pulvirenti, Wagner and Rossi [17].

REMARK 3.3. Certain chaotic initial conditions may work as in Theorem 1.4 in Graham and Méléard [7].

REMARK 3.4. This rate of convergence allows to get some results in which the regularizing kernel in (1.3) converges to the Dirac measure at the same time as the number of particles increases sufficiently fast. In the grid methods this corresponds to the cell size going to zero. See, for instance, results in this direction in Babovsky and Illner [1].

REMARK 3.5 (Algorithms for the Boltzmann equation). The empirical measure $\overline{Z}^n$ approximates the law of the Boltzmann process and depends on the choice of the joint jump measure $\hat{\mu}$ in (2.4). The ideas behind the proof of Theorem 3.1 can be used to simulate the particle systems. There are n particles and the total rate for the $n(n-1)/2$ pairs of possible interactions is $n\Lambda/2$ as seen in (2.6). A Poisson process of rate $n\Lambda/2$ gives the sequence of collision times. At each of these we choose uniformly the pair of particles which interact, update the states of these particles under L, compute $|\hat{\mu}|$ at these states, discard the jump with probability $1-|\hat{\mu}|/\Lambda$ and with probability $|\hat{\mu}|/\Lambda$ chose the joint jump amplitude according to $\hat{\mu}/|\hat{\mu}|$. All this is done independently. We evaluate at each step only the cross section of the interacting pair and not those of the $n(n-1)/2$ pairs. This simulation is exact if we solve the free motion explicitly and if we simulate the exponential variables exactly (instead of discretizing time).

4. Weak convergence by martingale problem methods in case of Vlasov terms. In physical models mixing strong (Boltzmann) interaction with weak (Vlasov) interaction as in Cercignani ([2], Sections II-4, II-9 and III-2), the diffusion operator also is nonlinear and interaction graph techniques fail. We use martingale problem techniques involving uniqueness results and give results for weak convergence (for the Skorokhod topology on the path space) instead of the variation norm. The complexity of the Skorokhod topology (discontinuity of the projections...) renders the proof and the deduction of results on the flow of time marginals intricate and original. The situation is more sensitive to the actual form of $L(z,p)$ and the uniqueness result is delicate. We need more regularity on $K(z,p)$, but may take a nonlinear dependence of $m(z,p)$ on the law p.

We metrize weak convergence plus convergence of the first moment on $P(\mathbb{R}^d)$ using the Kantorovitch–Rubinstein metric $\rho(p,q)=\sup\{\langle f,p\rangle - \langle f,q\rangle: |f(x)-f(y)|\le|x-y|\}$ which we extend to bounded jump measures m and n by $\rho(m,n)=\sup\{\langle f,m\rangle-\langle f,n\rangle: |f(x)-f(y)|\le|x-y|,\ f(0)=0\}$.

ASSUMPTION 4.1. σ and b are Lipschitz in (z,p) for the metric ρ, $a=\sigma\sigma^*$, and for ϕ in $C_b^2(\mathbb{R}^d)$,

$$L\phi(z,p)=\sum_{i=1}^d b_i(z,p)\partial_i\phi(z)+\tfrac{1}{2}\sum_{i,j=1}^d a_{ij}(z,p)\partial_{ij}^2\phi(z).$$

ASSUMPTION 4.2. (i) $\sup_{z,p}|m(z,p)|<\infty$, and $m(z,p)$ is Lipschitz for the metric ρ.

(ii) Define $e(z,m\,p)=\int h\,m(z,p,dh)$, $c_{ij}(z,p)=\int h_i h_j\, m(z,p,dh)$. Either P_0 has a second moment and $|e(z,p)|^2+\operatorname{tr}(c(z,p))\le K(1+|z|^2)$ uniformly in p, or P_0 has a first moment, $\sigma=0$ and $|e(z,p)|\le K(1+|z|)$ uniformly in p.

(iii) For bounded ϕ, $K\phi(z,p)=\int(\phi(z+h)-\phi(z))m(z,p,dh)$.

We say that $\tilde{P} \in P(\mathbb{R}_+, \mathbb{R}^d))$ solves the Boltzmann–Vlasov nonlinear martingale problem if

$$\phi(Z_t) - \phi(Z_0) - \int_0^t L\phi\big(Z_s, \tilde{P}_s\big) + K\phi\big(Z_s, \tilde{P}_s\big)\, ds = M_t^\phi \tag{4.1}$$

is a $\tilde{P}$-martingale for any ϕ in $C_b^2(\mathbb{R}^d)$.

THEOREM 4.3. *Under Assumptions* 4.1 *and* 4.2, *there is existence and uniqueness for the nonlinear martingale problem* (4.1) *starting at* P_0.

This is proved in Theorem 2.2 in Graham [5]. We wish to apply it to the model in Section 1. We need a Lipschitz cross section q and use a Lipschitz regularizing kernel I.

PROPOSITION 4.4. *Assume* $\int q(x, v, y, w, n)\, dn$ *is uniformly bounded,* $q(x, v, y, w, n)$ *is Lipschitz in* (x, v) *and* $\big| |(v-w)\cdot n| q(x, v, y, w, n) - |(v - \overline{w})\cdot n| q(x, v, \overline{y}, \overline{w}, n)\big| \le K_{x,v,n} |(y, w) - (\overline{y}, \overline{w})|$ *for a* $K_{x,v,n}$ *such that* $\int K_{x,v,n}\, dn$ *is bounded uniformly in* x *and* v. *Then the image measure of* $\int_{y\in\mathbb{R}^3} q(x, v, y, w, n) p(dy\, dw)\, dn$ *by the mapping* $(w, n) \mapsto ((w-v)\cdot n)n$ *is Lipschitz in* (x, v, p).

PROOF. Set $F_{xv}^f(y, w) = \int f(((w-v)\cdot n)n) q(x, v, y, w, n)\, dn$. The ρ-distance between the images by $(w, n) \mapsto ((w-v)\cdot n)n$ of $\int_{y\in\mathbb{R}^3} q(x, v, y, w, n) p(dy\, dw)\, dn$ and $\int_{y\in\mathbb{R}^3} q(x, v, y, w, n)\overline{p}(dy\, dw)\, dn$ is $\sup\{\langle F_{xv}^f, p - \overline{p}\rangle: |f(v) - f(w)| \le |v - w|, f(0) = 0\}$. This is less than $K\rho(p, \overline{p})$ as soon as F_{xv}^f is K-Lipschitz, which is true under our hypotheses since

$$\begin{aligned}
&\big|F_{xv}^f(y, w) - F_{xv}^f(\overline{y}, \overline{w})\big| \\
&\quad \le \int \big| f(((w-v)\cdot n)n)\, q(x, v, y, w, n) \\
&\qquad\quad - f(((\overline{w}-v)\cdot n)n)\, q(x, v, \overline{y}, \overline{w}, n)\big|\, dn \\
&\quad \le \int \big| f(((w-v)\cdot n)n) - f(((\overline{w}-v)\cdot n)n)\big| q(x, v, y, w, n) \\
&\qquad + \big| f(((\overline{w}-v)\cdot n)n)\big|\, \big| q(x, v, y, w, n) - q(x, v, \overline{y}, \overline{w}, n)\big|\, dn \\
&\quad \le |w - \overline{w}| \int q(x, v, y, w, n)\, dn \\
&\qquad + \int |(v - \overline{w})\cdot n|\, \big| q(x, v, y, w, n) - q(x, v, \overline{y}, \overline{w}, n)\big|\, dn. \qquad \square
\end{aligned}$$

We now introduce interacting particle systems. The Vlasov terms have a traditional mean-field interpretation. If m does not depend linearly on p, the only possible interpretation is the Nanbu, or simple mean-field, model. We define on $D(\mathbb{R}_+(\mathbb{R}^d)^n)$ a Markov process with generator

$$\sum_{i=1}^n L_i \Phi(\mathbf{z}^n, \overline{z}^{ni}) + \sum_{i=1}^n \int \big(\Phi(\mathbf{z}^n + \mathbf{e}_i \cdot h) - \Phi(\mathbf{z}^n)\big) m\big(z_i, \overline{z}^{ni}, dh\big). \tag{4.2}$$

If $m(z, p, dh) = \langle \mu(z, a, dh), p(da)\rangle$, we may interpret the collision term as in the Bird model (2.6). The systems are characterized by the martingale problem (2.7) [with $L\phi(Z_s^{ni}, \overline{Z}_s^{ni})$], (2.8) and (2.9). We have $A_t^{\phi n, i} = \int_0^t (\nabla\phi(Z_s^{ni})^* a(Z_s^{ni}, \overline{Z}_s^{ni})\nabla\phi(Z_s^{ni}))\, ds$.

THEOREM 4.5. *Let $(Z_0^{ni})_{1 \le i \le n}$ be exchangeable and P_0-chaotic. Let the hypotheses of Theorem* 4.3 *hold, and $\tilde{P}$ be the unique solution to the nonlinear martingale problem* (4.1) *starting at P_0. Then $(Z^{ni})_{1 \le i \le n}$ is $\tilde{P}$-chaotic*; $(\text{law}(Z^{n1}, \dots, Z^{nk}))_{n \ge 1}$ *converges weakly to $\tilde{P}^{\otimes k}$ in $P(D(\mathbb{R}_+, (\mathbb{R}^d)^k))$ for any k. Moreover, $\overline{Z}^n$ converges in probability to $\tilde{P}$.*

PROOF. Let $P^n \in P(D(\mathbb{R}_+, (\mathbb{R}^d)^n))$ be the law of the n-particle system; $\text{law}(Z^{n1}, \dots, Z^{nk}) = P^n \circ (Z^{n1}, \dots, Z^{nk})^{-1}$ and $\text{law}(\overline{Z}^n) = P^n \circ (\overline{Z}^n)^{-1}$. Since P^n is exchangeable, $P^n \circ (Z^{n1}, \dots, Z^{nk})^{-1}$ converges weakly to $\tilde{P}^{\otimes k}$ in $P(D(\mathbb{R}_+, (\mathbb{R}^d)^k))$ for all k if and only if $P^n \circ (\overline{Z}^n)^{-1}$ converges weakly to the Dirac mass at $\tilde{P}$ in $P(P(D(\mathbb{R}_+, \mathbb{R}^d)))$ (cf. Proposition 2.2 in [19]). Thus to prove propagation of chaos it is enough to prove tightness of the empirical measures and that the accumulation points solve a nonlinear martingale problem with unique solution.

Tightness of $(P^n \circ (\overline{Z}^n)^{-1})_{n \ge 1}$ in $P(P(D(\mathbb{R}_+, \mathbb{R}^d)))$ is equivalent to tightness of $(P^n \circ (Z^{n1})^{-1})_{n \ge 1}$ in $P(D(\mathbb{R}_+, \mathbb{R}^d))$ by Proposition 2.2 in [19]. Under Assumptions 4.1 and 4.2, Lemma 3.2.2 in Joffe and Métivier [10] proves uniform square integrability on every $[0, T]$ by using the Gronwall lemma to propagate the initial second moment assumption. Tightness of $(P^n \circ (Z^{n1})^{-1})_{n \ge 1}$ follows in Proposition 3.2.3 in [10] from this and the Aldous and Rebolledo criteria. If $\sigma = 0$ we only need the first moments. Thus $(P^n \circ (\overline{Z}^n)^{-1})_{n \ge 1}$ is relatively compact, and we now characterize its accumulation points.

Let π in $P(P(D(\mathbb{R}_+, \mathbb{R}^d)))$ be an accumulation point of $(\pi^n)_{n \ge 1} = (P^n \circ (\overline{Z}^n)^{-1})_{n \ge 1}$, limit of $(\pi^{n_k})_{k \ge 1}$. For $\phi \in C_b^2$, $0 \le s_1, \dots, s_q \le s \le t$, $g_1, \dots, g_q \in C_b$, Q in $P(D(\mathbb{R}_+, \mathbb{R}^d))$,

$$F_{sts_1 \cdots s_q}(Q) = \left\langle \left(\phi(Z_t) - \phi(Z_s) - \int_s^t L\phi(Z_u, Q_u) + K\phi(Z_u, Q_u)\, du \right) \times g_1(Z_{s_1}) \cdots g_q(Z_{s_q}), Q \right\rangle. \tag{4.3}$$

$F_{sts_1 \cdots s_q}$ is not continuous since the projections $Z \mapsto Z_t$ are not continuous for the Skorokhod metric. Let $J = \{u \in \mathbb{R}_+ : \pi(Q: Q(Z: |\Delta Z_u| > 0) > 0) > 0\}$. Clearly using monotone convergence of probability measures $J = \bigcup_{k \ge 1} J_k$, where $J_k = \{u \in [0, k]: \pi(Q: Q(Z: |\Delta Z_u| > 1/k) > 1/k) > 1/k\}$.

Since π is a probability measure, there are less than k laws Q such that for some u in $\mathbb{R}_+$, $\pi(Q: Q(Z: |\Delta Z_u| > 1/k) > 1/k) > 1/k$. Similarly there are less than k paths Z such that for some u, $Q(Z: |\Delta Z_u| > 1/k) > 1/k$. Finally since any path Z is right-continuous with left-hand limits, there are finitely many u on $[0, k]$ such that $|\Delta Z_u| > 1/k$. Hence J is countable, and for all $s, t, s_1, \dots, s_q$ in its complement $F_{sts_1 \cdots s_q}$ is π-a.s. continuous.

Then $F^2_{sts_1\cdots s_q}$ is π-a.s. continuous and $\pi^{n_k}\circ(F^2_{sts_1\cdots s_q})^{-1}$ converges weakly to $\pi\circ(F^2_{sts_1\cdots s_q})^{-1}$, and the Fatou lemma (using, for instance, the Skorokhod representation theorem) gives $\langle F^2_{sts_1\cdots s_q},\pi\rangle \le \liminf_k\langle F^2_{sts_1\cdots s_q},\pi^{n_k}\rangle$. By developing the square and using exchangeability,

$$\begin{aligned}\langle F^2_{sts_1\cdots s_q},\pi^n\rangle &= E^n\Big(F^2_{sts_1\cdots s_q}(\overline{Z}^n)\Big)\\ &= \frac{1}{n}E^n\Big(\Big(\big(M_t^{\phi n1}-M_s^{\phi n1}\big)g_1\big(Z^{n1}_{s_1}\big)\cdots g_q\big(Z^{n1}_{s_q}\big)\Big)^2\Big)\\ &\quad+\frac{n-1}{n}E^n\Big(\big(M_t^{\phi n1}-M_s^{\phi n1}\big)\big(M_t^{\phi n2}-M_s^{\phi n2}\big)\\ &\qquad\times g_1\big(Z^{n1}_{s_1}\big)\cdots g_q\big(Z^{n1}_{s_q}\big)g_1\big(Z^{n2}_{s_1}\big)\cdots g_q\big(Z^{n2}_{s_q}\big)\Big),\end{aligned}$$

which goes to zero using the $1/(n-1)$ in (2.9) and the uniform L^2 bounds, and $\langle F^2_{sts_1\cdots s_q},\pi\rangle = 0$.

Then $F_{sts_1\cdots s_q}(Q)=0$, π-a.s., for all choices of $s, t, s_1,\dots,s_q$ outside of the countable set J and of $g_1,\dots,g_q$ in C_b. Convergence of the initial values is immediate, and π-a.s., Q solves the nonlinear martingale problem (4.1) starting at P_0. This problem has a unique solution $\tilde{P}$ as stated in Theorem 4.3, and π is the Dirac mass at $\tilde{P}$. Thus $(P^n\circ(\overline{Z}^n)^{-1})_{n\ge 1}$ converges to the Dirac mass at $\tilde{P}$. □

This result on path space implies a functional law of large numbers. Such an implication is simple for processes with continuous paths, but not here because of topological difficulties related to the jumps. We give a general topological result in which we adapt an intermediate result in the proof of Lemma 2.8 in [12].

THEOREM 4.6. *Let a random sequence $(Q^n)_{n\ge 0}$ converge in probability in $P(D([0,T],\mathbb{R}^d))$ to Q such that ω-a.s., $Q(\omega)(Z\colon |\Delta Z_u|>0)=0$ for any u in $[0,T]$. Then the flow $(Q^n_t)_{t\ge 0}$ converges in probability to $(Q_t)_{t\ge 0}$ in $D([0,T], P(\mathbb{R}^d))$ with the uniform norm.*

PROOF. Convergence in probability is equivalent to the fact that from any subsequence one can extract a further subsequence which converges a.s. Take a subsequence; let $(n_k)_{k\ge 1}$ be a subsubsequence such that Q^{n_k} converges a.s. to Q for the weak topology on $P(D([0,T],\mathbb{R}^d))$ and let ω be such that $Q^{n_k}(\omega)$ converges to $Q(\omega)$. By the Skorokhod representation theorem, there exists a probability space $(W,\mathcal{W},\pi)$ and $D([0,T],\mathbb{R}^d)$-valued random variables Y^k and Y such that $\pi\circ(Y^k)^{-1}=Q^{n_k}(\omega)$, $\pi\circ(Y)^{-1}=Q(\omega)$, and Y^k converges to Y. Hence the theorem is true if

$$\lim_{k\to\infty}\sup_{0\le t\le T}E^\pi\big(|Y^k_t-Y_t|\wedge 1\big)=0. \tag{4.4}$$

For $x\in D([0,T],\mathbb{R}^d)$, $S\subset[0,T]$ and $\delta>0$ we set $w_x(S)=\sup\{|x_t-x_s|\colon s,t\in S\}$ and $w'_x(\delta)=\inf_{\{t_i\}}\max_{1\le i\le r} w_x([t_{i-1},t_i[)$ where the infimum is over all finite sets $\{t_i\}$ such that $0 = t_0 < t_1 < \cdots < t_r = T$ and $t_i - t_{i-1} > \delta$ for all i in $\{1, \ldots, r\}$. It is easy to see that for $0 \le a \le b \le T$,

$$w_x(]a, b[) \le 2 w'_x(b - a) + \sup_{s \in]a, b[} |\Delta x_s|. \tag{4.5}$$

Let $\alpha > 0$ and x, y in $D([0, T], \mathbb{R}^d)$ be such that $d(x, y) < \alpha$, where d denotes the Skorokhod metric. There is a time-change λ such that $\sup_{0 \le t \le T} |\lambda(t) - t| < \alpha$ and $\sup_{0 \le t \le T} |y_t - x_{\lambda(t)}| < \alpha$, and thus $d(x, y) < \alpha$ implies $|x_t - y_t| < \alpha + w_x(]t - \alpha, t + \alpha[)$. Using (4.5), this shows that

$$\begin{aligned} &E^\pi\big(|Y_t^k - Y_t| \wedge 1\big) \\ &\quad\le \pi\big(d(Y^k, Y) \ge \alpha\big) + E^\pi\big(|Y_t^k - Y_t| \wedge 1 : d(Y^k, Y) < \alpha\big) \\ &\quad\le \pi\big(d(Y^k, Y) \ge \alpha\big) + \alpha + E^\pi\big(w_Y(]t - \alpha, t + \alpha[) \wedge 1\big) \\ &\quad\le \pi\big(d(Y^k, Y) \ge \alpha\big) + \alpha + E^\pi\Bigg(\bigg(2 w'_Y(2\alpha) + \sup_{s \in]t-\alpha, t+\alpha[} |\Delta Y_s|\bigg) \wedge 1\Bigg). \end{aligned}$$

Y^k converges to Y in $D([0, T], \mathbb{R}^d)$ and $\lim_{\delta \to 0} w'_x(\delta) = 0$ for any x in $D([0, T], \mathbb{R}^d)$. Using the dominated convergence theorem and $\pi \circ (Y)^{-1} = Q(\omega)$, (4.4) is true as soon as

$$\begin{aligned} &\lim_{\alpha \to 0} \sup_{0 \le t \le T} E^\pi\bigg(\sup_{s \in]t-\alpha, t+\alpha[} |\Delta Y_s| \wedge 1\bigg) \\ &\qquad = \lim_{\alpha \to 0} \sup_{0 \le t \le T} E^{Q(\omega)}\bigg(\sup_{s \in]t-\alpha, t+\alpha[} |\Delta Z_s| \wedge 1\bigg) = 0. \end{aligned} \tag{4.6}$$

Since for $\varepsilon > 0$, $\{Z: \ |\Delta Z_u| \ge \varepsilon\}$ is the decreasing limit of $\{Z: \sup_{s \in]u-1/k, u+1/k[} |\Delta Z_s| \ge \varepsilon\}$, $Q(\sup_{s \in]u-\alpha_u, u+\alpha_u[} |\Delta Z_s| \ge \varepsilon) \le \varepsilon$ and hence $E^Q(\sup_{s \in]u-\alpha_u, u+\alpha_u[} |\Delta Z_s| \wedge 1) \le 2\varepsilon$ for small enough $\alpha_u > 0$. Let $\alpha > 0$ be half of the minimal length of the overlaps of a finite covering of $[0, T]$ by the $]u - \alpha_u, u + \alpha_u[$. For any t in $[0, T]$, $]t - \alpha, t + \alpha[$ is included in an $]u - \alpha_u, u + \alpha_u[$, hence $\sup_{0 \le t \le T} E^Q(\sup_{s \in]t-\alpha, t+\alpha[} |\Delta Z_s| \wedge 1) \le 2\varepsilon$ and (4.6) holds. □

COROLLARY 4.7. *Under the assumptions of Theorem* 4.3, *we have convergence of the flows* $(\overline{Z}_t^n)_{t \ge 0}$ *to* $(\tilde{P}_t)_{t \ge 0}$ *in law and probability in* $D(\mathbb{R}_+, P(\mathbb{R}^d))$ *with the topology of uniform convergence on compact sets and* $(\mathrm{law}(Z_t^{n1}, \ldots, Z_t^{kn}))_{t \ge 0}$ *to* $(\tilde{P}_t^{\otimes k})_{t \ge 0}$ *uniformly on compact sets.*

PROOF. We use Theorems 4.5 and 4.6. If $|m| \le \Lambda$,

$$\tilde{P}\big(\sup_{s \in]u-\alpha, u+\alpha[} |\Delta Z_s| > 0\big) \le 2\Lambda\alpha. \qquad \square$$

REFERENCES

[1] BABOVSKY, H. and ILLNER, R. (1994). A convergence proof for Nanbu's simulation method for the full Boltzmann equation. *SIAM J. Numer. Anal.* **26** 45–65.

[2] CERCIGNANI, C. (1988). *The Boltzmann Equation and Its Applications*. Springer, New York.

[3] Chauvin, B. (1993). Branching processes, trees and the Boltzmann equation. *Proceedings du Congrès Probabilités Numériques*, INRIA.

[4] DiPerna, R. J. and Lions, P. L. (1989). On the Cauchy problem for the Boltzmann equation: global existence and weak stability. *Ann. Math.* **130** 321–366.

[5] Graham, C. (1992). Nonlinear diffusion with jumps. *Ann. Inst. H. Poincaré* **28** 393–402.

[6] Graham, C. (1992). McKean–Vlasov Ito–Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Process. Appl.* **40** 69–82.

[7] Graham, C. and Méléard, S. (1993). Propagation of chaos for a fully connected loss network with alternate routing. *Stochastic Process. Appl.* **44** 159–180.

[8] Graham, C. and Méléard, S. (1994). Chaos hypothesis for a system interacting through shared resources. *Probab. Theory Related Fields* **100** 157–173.

[9] Illner, R. and Neunzert, H. (1987). On simulation methods for the Boltzmann equation. *Transport Theory Statist. Phys.* **16** 141–154.

[10] Joffe, A. and Métivier, M. (1986). Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Adv. in Appl. Probab.* **18** 20–65.

[11] Lanford, III, O. E. (1975). Time evolution of large classical systems. *Lecture Notes in Phys.* **38** 1–111. Springer, Berlin.

[12] Léonard, C. (1995). Large deviations for long range interacting particle systems with jumps. *Ann. Inst. H. Poincaré.*

[13] Nanbu, K. (1983). Interrelations between various direct simulation methods for solving the Boltzmann equation. *J. Phys. Soc. Japan* **52** 3382–3388.

[14] Neunzert, H., Gropengeisser, F. and Struckmeier, J. (1991). Computational methods for the Boltzmann equation. In *Applied and Industrial Mathematics* (R. Spigler, ed.) 111–140. Kluwer, Dordrecht.

[15] Perthame, B. (1994). Introduction to the theory of random particle methods for Boltzmann equation. In *Progresses on Kinetic Theory*. World Scientific, Singapore.

[16] Perthame, B. and Pulvirenti, M. (1996). On some large systems of random particles which approximate scalar conservation laws. *Asympt. Anal.* To appear.

[17] Pulvirenti, M., Wagner, W. and Zavelani Rossi, M. B. (1993). Convergence of particle schemes for the Boltzmann equation. Preprint 49, Institut fur Angewandte Analysis und Stochastik, Berlin.

[18] Sznitman, A. S. (1984). Equations de type de Boltzmann, spatialement homogènes. *Z. Wahrsch. Verw. Gebeite* **66** 559–592.

[19] Sznitman, A. S. (1991). Propagation of chaos. *Ecole d'Été de Probabilités de Saint-Flour 1989* (*Lecture Notes in Math.* **1464** 165–251). Springer, Berlin.

[20] Uchiyama, K. (1988). Derivation of the Boltzmann equation from particle dynamics. *Hiroshima Math. J.* **18** 245–297.

[21] Wagner, W. (1992). A convergence proof for Bird's direct simulation method for the Boltzmann equation. *J. Statist. Phys.* **66** 1011–1044.

[22] Wagner, W. (1994). A functional law of large numbers for Boltzmann type stochastic particle systems. Preprint 93, Institut fur Angewandte Analysis und Stochastik, Berlin.

CMAP, École Polytechnique
F-91128 Palaiseau
France
E-mail: mata@cmapx.polytechnique.fr

Laboratoire de Probabilités
Universite Paris 6
F-75231 Paris
France