IMS Collections
Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz
Vol. 4 (2008) 63–83

DOI: 10.1214/074921708000000309

From Particles with Random Potential to a Nonlinear Vlasov–Fokker–Planck Equation

Jin Feng[1,*]

University of Kansas

Abstract: We consider large time and infinite particle limit for a system of particles living in random potentials. The randomness enters the potential through an external ergodic Markov process, modeling oscillating environment with good statistical averaging properties.

From each individual particle's point of view, both law of large number and central limit theorem type of averaging are possible. Problems of this type have been well studied and are known as random evolutions. Instead of one particle, we focus on the collective behavior of infinite particles. We separately rescale potential functions (type one) which annihilates the equilibrium measure of the ergodic environment process, and the potential functions which may not annihilate such measure (type two). Appropriately rescaled to the macroscopic limit, type two potentials give a transport term while type one potentials give a nonlinear diffusion term. The resulting equation is a version of nonlinear Vlasov–Fokker–Planck equation. We will also prove the uniqueness of solution for such equation.

1. Introduction

We rigorously derive a version of nonlinear Vlasov–Fokker–Planck equation

$$\partial_t \rho(t,x,v) + v\cdot\nabla_x \rho(t,x,v) - \Big(\rho_X * \nabla\bar\Phi_1 + \nabla\bar\Psi_1\Big)(t,x)\cdot\nabla_v\rho(t,x,v) = a(\rho_X;x)\cdot D^2_{vv}\rho(t,x,v) \tag{1.1}$$

as multi-scaled limit of infinite interacting systems with random potential and prove its uniqueness (Theorem 6.9).

In the above equation, $x, v \in R^d$, D^2_{vv} is the Hessian matrix where the derivatives are only taken with respect to v; a is some square matrix specified in (1.15) using potential functions defining a dynamic at the microscopic level; and by $M\cdot N$ for two $d\times d$ matrices $M=(m_{ij})_{d\times d}$ and $N=(n_{ij})_{d\times d}$, we mean

$$M\cdot N \equiv \sum_{i,j=1}^d m_{ij}n_{ji}. \tag{1.2}$$

---

*The author would like to thank Tom Kurtz for his sound advice during the author's PhD training, and for sharing his scientific curiosities many years after the author's graudation. This research has been supported in part by a New Faculty Research Grant at the University of Kansas, and US Army grant W911NF-08-1-0064.

[1]Department of Mathematics, University of Kansas, Lawrence, KS 66045, e-mail: jfeng@math.ku.edu

AMS 2000 subject classifications: Primary 60F17, 60K35. Secondary 82C40

Keywords and phrases: random potential, nonlinear Vlasov–Fokker–Planck equation, perturbed test function method

The whole equation (1.1) is understood in the weak (Schwartz distributional) sense, where in particular $\rho(t,x,v)$ is understood as a probability measure in x, v-variables, and ρ_X is the x-marginal probability measure of ρ (i.e. $\rho_X(dx) = \rho(dx, R^d)$). Macroscopic functions $\bar{\Phi}_1, \bar{\Psi}_1$ and $a(\rho_X, x)$ are obtained from microscopically defined functions Φ_i, Ψ_i, $i = 1,2$ by (1.7) and (1.15).

At least formally, the following conservation of mass can be directly verified

$$\frac{d}{dt}\int \rho(t,x,v)dxdv = 0.$$

Indeed, a constructive existence theory in this article will show that ρ is a probability measure. In the special case of $\bar{\Psi}_1 = 0$, and imposing a natural assumption (1.3) on Φ_1 (hence on $\bar{\Phi}_1$), we also have conservation of momentum

$$\frac{d}{dt}\int v\rho(t,x,v)dxdv = 0.$$

There is even an entropy-entropy production formula

$$\frac{d}{dt}\int_{x,v}(\rho\log\rho)dxdv = -\int_{x,v} a(\rho_X;x)\cdot I(\rho;x,v)dxdv$$

with Fisher information matrix

$$I(\rho;x,v) = \Big(\frac{\partial_{v_i}\rho(x,v)\partial_{v_j}\rho(x,v)}{\rho(x,v)}\Big)_{d\times d}.$$

However, in general, there is no macroscopic version of energy function we can define so that (1.1) conserves it.

(1.1) belongs to a type of equation known as kinetic equations. They play important roles in the kinetic theory of gas, and in scaling limit issues which lead to derivation of (system of) nonlinear PDEs describing evolution of spatial density(ies) of momentum or(and) energy. See for instance Cercignani [2] on discussions of Boltzmann equation and its variants (such as kinetic Fokker–Planck equation (9.19) in Chapter II). For such equations in general, uniqueness and the derivation from microscopic particle models are extremely hard to show, with many issues still open. (1.1) is a toy model. In the hierarchy of scaling limits, it is much farther way from the Boltzmann equation or its variants, but closer to plain transport equations. In this article, we identify a microscopic situation where lots of simplification can be made, yet basic mechanical features (such as conservation of momentum and mass) are retained. We will introduce a random environment with some independence property. The model and scaling we choose are so that we will have very strong uniform ergodic property, avoiding the need to justify a what so called local equilibrium hypothesis in statistical mechanics. These simplifications allow us to justify the passage of limit from microscopic to macroscopic level relatively painlessly, using generalization of classical averaging/homogenization techniques (in the spirit of Kurtz [5], Bensoussan, Lions and Papanicolaou [1]). Because the resulting measure-valued PDE (1.1) has smooth coefficients, we can also prove its uniqueness using more than one approaches - we will adapt a recent method developed by Kurtz in [6]. Of course, losing conservation of energy is a big downside of the model (1.1). But in situations where only conservation of mass and momentum are crucial, such model can still be useful.

1.1. The microscopic model

Let $x = (x_1, \ldots, x_N)$ denote the position of N-particles, $N = 1, 2, \ldots$. Each $x_i \in R^d$. Similarly, we also use v_i to denote the velocity of the i-th particle. We assume that each particle has unit mass.

We use function $\Phi_n(z; y_i, y_j) : (z, y_i, y_j) \in R^d \times S \times S \mapsto R$ to model pair-wise interaction potential between two particles at locations x_i, x_j $(z = x_i - x_j)$ when "environmental" states of the particles are respectively $y_i, y_j \in S$. S is a compact metric space modeling environment. We assume that

$$\Phi(z; y_1, y_2) = \Phi(z; y_2, y_1), \quad \Phi(z; y_1, y_2) = \Phi(-z; y_1, y_2). \tag{1.3}$$

The random environment is determined by a large external Markov process $Y(t) = (Y_1(t), \ldots, Y_N(t)) \in S^N$ with generator $\mathbf{B}$. To free us from possible complications caused by boundary conditions in the x-variable, we consider the whole space R^d and introduce an external potential $\Psi_n(x; y)$, acting on each individual particle at location x when the environment is in state y. We will look at a particular scaling which corresponds to large time (controlled by parameter n) behavior of the system:

$$\begin{aligned} \dot{x}_i(t) &= v_i(t) \\ \dot{v}_i(t) &= n\Big\{ -\frac{1}{N}\sum_{j=1}^{N} \Big(\nabla\Phi_n\Big)\Big(x_i(t) - x_j(t); Y_i(n^2 t), Y_j(n^2 t)\Big) \\ &\qquad - \nabla\Psi_n\Big(x_i(t); Y_i(n^2 t)\Big)\Big\}. \end{aligned} \tag{1.4}$$

For simplicity, we write

$$\nabla\Phi_n(x; y_1, y_1) = \nabla_x \Phi_n(x; y_1, y_2), \quad \nabla\Psi_n(x, y) = \nabla_x \Psi_n(x, y). \tag{1.5}$$

Under appropriate smoothness and growth conditions on Ψ_n, Φ_n and regularities on Y, the above equation has a unique solution.

1.2. Structural assumptions

Let Y_is be independent Markov process with weak infinitesimal generator B in $C_b(S)$ in the following sense. For bounded measurable function f, let

$$S(t)f(y) = E[f(Y_i(t))|Y_i(0) = y].$$

We assume that $S(t) : C_b(S) \mapsto C_b(S)$, and define domain $D(B)$ of B to be functions in $C_b(S)$ such that $\limsup_{t\to 0} \sup_{y\in S} t^{-1}|S(t)f(y) - f(y)| < +\infty$ and that the limit $Bf(y) = \lim_{t\to 0+} t^{-1}(S(t)f(y) - f(y))$ exists as a function in $C_b(S)$.

Similarly, we define generator $\mathbf{B}$ for $Y = (Y_1, \ldots, Y_N)$. Then

$$\mathbf{B}\varphi(y_1, \ldots, y_N) = \sum_{k=1}^{m} B\varphi(y_1, \ldots, y_{i-1}, \cdot, y_{i+1}, \ldots, y_N)(y_i).$$

whenever $\varphi(y_1, \ldots, y_{i-1}, \cdot, y_{i+1}, \ldots, y_N) \in D(B)$, and

$$B\varphi(y_1, \ldots, y_{i-1}, \cdot, y_{i+1}, \ldots, y_N)(y_i) \in C_b(S^N).$$

We assume that Y_i has the following ergodic properties.

Condition 1.1. *Let Y be the Markov process on a compact metric space S with weak infinitesimal generator B.*

1. *Y has a unique stationary probability measure $\pi_0(dy)$ in the sense that*

$$\lim_{t\to+\infty} S(t)f(x) = \lim_{t\to+\infty} E[f(Y(t))|Y(0)=x] = \int f(y)\pi_0(dy), \quad \forall x \in S.$$

2. *for each bounded measurable function h, there exists constant $C_h > 0$ such that*

$$|S(t)h(x) - \pi_0 h(x)| \le C_h(1+t^2)^{-1}.$$

Typical examples satisfying the above requirements are random walks (continuous time) with a communication condition (in the sense that there is positive probability to reach each point from any other point), Brownian motions on compact manifold, etc, etc. By the second part of Condition 1.1, the measure

$$\nu(y,dz) = \int_0^\infty (P(Y(t)\in dz|Y(0)=y) - \pi_0(dz))dt$$

is well defined. Let $\varphi = \varphi(y_1,\dots,y_m) \in C_b(S^m)$, define

$$(P_{\nu\otimes\dots\otimes\nu}\varphi)(y_1,\dots,y_m) = \int_S \dots \int_S \varphi(z_1,\dots,z_m)\nu(y_1,dz_1)\dots\nu(y_m,dz_m),$$

and constant function

$$P_{\pi_0\otimes\dots\otimes\pi_0}\varphi(y_1,\dots,y_m) = \int_S \dots \int_S \varphi(z_1,\dots,z_m)\pi_0(dz_1)\dots\pi_0(dz_m).$$

Since $\mathbf{B}P_{\pi_0\otimes\dots\otimes\pi_0}\varphi = 0$,

$$\mathbf{B}P_{\nu\otimes\dots\otimes\nu}\varphi(y_1,\dots,y_m) + \varphi(y_1,\dots,y_m) = P_{\pi_0\otimes\dots\otimes\pi_0}\varphi. \tag{1.6}$$

For a smooth function f on R^d, we denote

$$|D^k f(x)| = \sum_{i_1+\dots+i_d=k} \Big|\frac{\partial^k}{\partial x_1^{i_1}\dots\partial x_d^{i_d}} f(x)\Big|, \quad x\in R^d,$$

and write

$$\bar\Phi_1(x) = \int_S\int_S \Phi_1(x;y_1,y_2)\pi_0(dy_1)\pi_0(dy_2), \quad \bar\Psi_1(x) = \int_{y\in S}\Psi_1(x,y)\pi_0(dy). \tag{1.7}$$

Condition 1.2.

1. *Φ_n and Ψ_n satisfy*

$$\begin{aligned}\Phi_n(x;y_1,y_2) &= n^{-1}\Phi_1(x;y_1,y_2) + \Phi_2(x;y_1,y_2),\\ \Psi_n(x;y) &= n^{-1}\Psi_1(x;y) + \Psi_2(x;y);\end{aligned} \tag{1.8}$$

2. *$\nabla\Phi_1(z;y_1,y_2), \nabla\Phi_2(z;y_1,y_2) \in C_b(R^d\times S\times S)$ and*

$$\sup_{z\in R^d, y_1,y_2\in S} (|D^1\Phi_2(z,y_1,y_2)| + |D^2\Phi_2(z,y_1,y_2)|) < +\infty,$$

where the $D^k, k=1,2$ applies to the z-variable (recall (1.5) for the notation $\nabla\Phi_2, \nabla\Psi_2$);

3.

$$\text{(1.9)} \quad \int_{S\times S} \nabla\Phi_2(x;y_1,y_2)\pi_0(dy_1)\pi_0(dy_2)=0, \qquad \int_S \nabla\Psi_2(x,y)\pi_0(dy)=0;$$

4. $\Phi_1,\Psi_1\geq 0$, $\nabla\Psi_2(z;y)\in C_b(R^d\times S)$; *and for each* $x\in R^d$, $\nabla\Psi_1(x;\cdot)\in C_b(S)$, $\bar\Psi_1\in C(R^d)$, *and*

$$\lim_{M\to+\infty}\inf_{|z|>M}\bar\Psi_1(z)=+\infty.$$

5. The number of particles N *and the multiple time scale parameter* n *satisfy*

$$\text{(1.10)} \qquad \lim_{n\to+\infty} nN^{-1}=0.$$

1.3. Main result

The symmetry among particle labels suggests that we can identify the stochastic system $(x_i(\cdot),v_i(\cdot),Y_i(n^2\cdot))$ with its empirical measure without any lose of generality:

$$\text{(1.11)} \qquad \gamma_n(t,dx,dv,dy)\equiv\frac{1}{N}\sum_{i=1}^N \delta_{\{x_i(t),v_i(t),Y_i(n^2t)\}}(dx,dv,dy).$$

Let

$$\text{(1.12)} \qquad \rho_n(t,dx,dv)=\gamma_{n,X,V}(t,dx,dv)=\gamma_n(t,dx,dv,S),$$

and

$$\text{(1.13)} \qquad \rho_{n,X}(t,dx)\equiv\rho_n(t,dx,R^d).$$

In Theorem 4.1, we show that $\{\rho_n(\cdot):n=1,2,\ldots\}$ is a tight sequence as stochastic processes with trajectories in the space $C_{\mathcal{P}(R^d\times R^d)}[0,+\infty)$. In Theorem 5.1, we conclude that any limit point of the above sequence is a solution to the stochastic Vlasov–Fokker–Planck equation (1.1). Finally, with additional smoothness hypothesis on $\Phi_i,\Psi_i, i=1,2$, in Theorem 6.9, we show that solution to (1.1) is unique, hence ρ_n converges to the solution of (1.1).

We specify the coefficients in (1.1) next: let

$$\text{(1.14)} \qquad \bar{\bar\Phi}_2(x;y)=\int_{\bar y\in S}\int_{z\in S}\Phi_2(x;y,z)\nu(\bar y,dz)\pi_0(d\bar y);$$

and define the $d\times d$ square matrix $a(\rho,x)=\big(a_{ij}(\rho,x)\big)_{i,j=1,\ldots,d}$ by

$$\begin{aligned}\text{(1.15)} \qquad a_{ij}(\rho,x)&=a_{ij}(\rho_X,x)\\ &=\mathcal{E}(P_\nu(\rho_X*\nabla^{(i)}\bar{\bar\Phi}_2(x;\cdot)+\nabla^{(i)}\Psi_2(x;\cdot))),\\ &\qquad P_\nu(\rho_X*\nabla^{(j)}\bar{\bar\Phi}_2(x;\cdot)+\nabla^{(j)}\Psi_2(x;\cdot))))\end{aligned}$$

where $\mathcal{E}$ is the Dirichlet form associated with B:

$$\mathcal{E}(f,g)=-\frac{1}{2}\int(fBg+gBf)d\pi_0, \quad f,g\in D(B).$$

By symmetry property of $\mathcal{E}$, $a(\rho;x)$ is a non-negative definite matrix.

1.4. Notations

If E_0 is a metric space, we use $M_\pm(E_0)$ to denote space of signed Borel measures on E_0. For a function $g : \gamma \in M_\pm(E_0) \mapsto R$, we define its first and second variational derivatives as functions

$$\frac{\delta g}{\delta \gamma} : E_0 \mapsto R, \quad \frac{\delta^2 g}{\delta \gamma^2} : E_0 \times E_0 \mapsto R$$

satisfying the Taylor's expansion

$$g(\gamma + t\gamma') - g(\gamma) = t \left\langle \frac{\delta g}{\delta \gamma}, \gamma' \right\rangle + \frac{1}{2} t^2 \left\langle \frac{\delta^2 g}{\delta \gamma^2}, \gamma' \otimes \gamma' \right\rangle + o(t^2) \tag{1.16}$$

for each $\gamma' \in M_\pm(E_0)$ with compact support. $\gamma' \otimes \gamma'$ means the product measure on $E_0 \times E_0$. As an example, let $E_0 = R^d \times R^d \times S$, for $\varphi_k \in B(R^d \times R^d \times S)$, $\psi \in C^2(R^m)$, $\gamma \in M_\pm(E_0)$ and

$$f(\gamma) = \psi(\langle \varphi_1, \gamma \rangle, \ldots, \langle \varphi_m, \gamma \rangle), \tag{1.17}$$

$$\frac{\delta f}{\delta \gamma}(x, v; y) = \sum_{k=1}^{m} \partial_k \psi(\langle \varphi_1, \gamma \rangle, \ldots, \langle \varphi_m, \gamma \rangle) \varphi_k(x, v, y),$$

and

$$\frac{\delta^2 f}{\delta \gamma^2}(x, v, y; \bar{x}, \bar{v}, \bar{y}) = \sum_{k,l=1}^{m} \partial_{kl}^2 \psi(\langle \varphi_1, \gamma \rangle, \ldots, \langle \varphi_m, \gamma \rangle) \varphi_k(x, v, y) \varphi_l(\bar{x}, \bar{v}, \bar{y}).$$

Throughout, we denote

$$E'_n = \left\{ \gamma(dx, dv, dy) = N^{-1} \sum_{i=1}^{N} \delta_{\{x_i, v_i, y_i\}}(dx, dv, dy) : x_i, v_i \in R^d, y_i \in S \right\}, \tag{1.18}$$

and

$$E_n = \{\rho(dx, dv) = \gamma(dx, dv; S) : \gamma \in E'_n\}. \tag{1.19}$$

We also denote $E = \mathcal{P}(R^d \times R^d)$ and $E' = \mathcal{P}(R^d \times R^d \times S)$. Typical element in E is denoted ρ and typical element in E' is γ. Topologies on these spaces we use are always the weak convergence of probability measure topology. Convergence of sequences is denoted $\rho_n \Rightarrow \rho$ or $\gamma_n \Rightarrow \gamma$. Each $\rho \in E_n$ is compactly supported (on a finite number of points), it has moments up to all orders.

For an operator B, $D(B)$ denote the domain of B. If $f = f(x, v, y)$: $R^d \times R^d \times S \mapsto R$ and $f(x, v, \cdot) \in D(B)$, we write $B_y f(x, v, y) = (Bf(x, v, \cdot))(y)$.

2. Martingale problem

Let A_n be generator (through martingale problem) for the Markov process $\gamma_n(t)$ in (1.11). First, we identify A_n for a relatively simpler class of test functions (1.17). Later, we will find that another type of test functions is also needed, but the calculations follow similarly.

We denote $\mathcal{F}_t^n = \sigma\{Y_i(r) : n = 1, 2, \ldots, N; 0 \le r \le n^2 t\}$. Let $\varphi = \varphi(x, v; y)$, $\nabla_x\varphi$, $\nabla_v\varphi \in C(R^d \times R^d \times S)$, and $\varphi(x, v, \cdot) \in D(B)$ and $B_y\varphi(x, v, y)$ is bounded continuous. By Lemma 4.3.4 of Ethier and Kurtz [4],

$$\begin{aligned}\varphi(x_i(t), v_i(t), Y_i(n^2t)) - \int_0^t \Big\{ &v_i(r)\nabla_x\varphi(x_i(r), v_i(r), Y_i(n^2r)) \\ &- n\Big(\int_{R^d\times R^d\times S} \nabla\Phi_n(x_i(r) - x; Y_i(n^2r), y)\gamma_n(r; dx, dv, dy) \\ &+ \nabla\Psi_n(x_i(r); Y_i(n^2r))\Big) \cdot \nabla_v\varphi(x_i(r), v_i(r), Y_i(n^2r)) \\ &- n^2 B\varphi(x_i(r), v_i(r), Y_i(n^2r))\Big\} dr \equiv M_i^{\varphi}(t);\end{aligned}$$

is a martingale, with co-quadratic variation

$$[M_i^{\varphi_1}, M_j^{\varphi_2}](t) = \delta_{ij} \int_0^t n^2 (\varphi_1, \varphi_2)_B(x_i(r), v_i(r), Y_i(n^2r)) dr,$$

where

$$\begin{aligned}(\varphi_1, \varphi_2)_B(x, v, y) = B_y(\varphi_1\varphi_2)(x, v, y) - \varphi_1(x, v, y)B_y\varphi_2(x, v, y) \\ -\varphi_2(x, v, y)B_y\varphi_1(x, v, y).\end{aligned}$$

First, we consider test functions f of the form (1.17) with φ satisfying the requirements given in the last paragraph. It follows that probability measure valued process $\gamma_n(\cdot)$ solves the martingale problem (i.e.

$$f(\gamma_n(t)) - f(\gamma_n(0)) - \int_0^t A_n f(\gamma_n(s)) ds \tag{2.1}$$

is a martingale) given by generator

(2.2) $A_n f(\gamma)$

$$\begin{aligned}&= \Big\langle \gamma, \Big(v \cdot \nabla_x - n\big(\int_{R^d\times R^d\times S} \nabla\Phi_n(x - \bar{x}; y, \bar{y})\gamma(d\bar{x}, d\bar{v}, d\bar{y}) \\ &\qquad + \nabla\Psi_n(x, y)\big) \cdot \nabla_v + n^2 B_y\Big) \frac{\delta f}{\delta\gamma} \Big\rangle \\ &\qquad + \frac{n^2}{2N} \Big\langle \gamma, \sum_{k,l=1}^m \partial_{kl}^2 \psi(\langle\varphi_1, \gamma\rangle, \ldots, \langle\varphi_m, \gamma\rangle)(\varphi_k, \varphi_l)_B \Big\rangle \\ &= \Big\langle \gamma, v \cdot \nabla_x \frac{\delta f}{\delta\gamma} - \Big(\int \nabla\Phi_1(x - \bar{x}; y, \bar{y})\gamma(d\bar{x}, d\bar{v}, d\bar{y}) + \nabla\Psi_1(x; y)\Big) \cdot \nabla_v \frac{\delta f}{\delta\gamma} \Big\rangle \\ &\quad - n\Big\langle \gamma, \Big(\int_{R^d\times R^d\times S} \nabla\Phi_2(x - \bar{x}; y, \bar{y})\gamma(d\bar{x}, d\bar{v}, d\bar{y}) + \nabla\Psi_2(x, y)\Big) \cdot \nabla_v \frac{\delta f}{\delta\gamma} \Big\rangle \\ &\quad + n^2 \Big\langle \gamma, B_y \frac{\delta f}{\delta\gamma} \Big\rangle + \frac{n^2}{2N} \Big(\Big\langle \gamma, B \frac{\delta^2 f}{\delta\gamma^2}(x, v, \cdot; x, v, \cdot)(y) \Big\rangle \\ &\quad - 2\Big\langle \gamma(dx, dv, dy) \otimes \delta_{\{x,v,y\}}(d\bar{x}, d\bar{v}, d\bar{y}), B_y \frac{\delta^2 f}{\delta\gamma^2}(x, v, y; \bar{x}, \bar{v}, \bar{y}) \Big\rangle \Big).\end{aligned}$$

Having only test functions f of the form (1.17) is not good enough. Later, we need to consider test functions of the form such as g in (3.5) and h in (3.6). These are special cases of

$$f(\gamma) = \phi_1(\langle\rho,\varphi_1\rangle,\dots,\langle\rho,\varphi_m\rangle)\int\int\int\int \phi_2(x_1,x_2,x_3,x_4;y_1,y_2,y_3,y_4)\phi_3(x_1,v_1)$$
$$\gamma(dx_1,dv_1,dy_1)\gamma(dx_2,dv_2,dy_2)\gamma(dx_3,dv_3,dy_3)\gamma(dx_4,dv_4,dy_4),$$

where $\varphi_k \in C_c^\infty(R^d\times R^d)$, $\phi_1 \in C^2(R^m)$ and ϕ_2,ϕ_3 are bounded continuous functions. For such case, at least when $\gamma \in E_n'$, the first two variational derivatives $\delta f/\delta\gamma$ and $\delta^2 f/\delta\gamma^2$ are well defined smooth functions. Then Ito's formula allow us to conclude that (2.1) is still a martingale.

3. Generator convergence for a class of perturbed test functions

We recall earlier convention that ρ denote the X,V-marginal of γ (i.e. $\rho(dx,dv) = \gamma(dx,dv,S)$). For each

(3.1)
$$f \in D_0 = \{f(\rho) = \psi(\langle\varphi_1,\rho\rangle,\dots,\langle\varphi_m,\rho\rangle):\quad \varphi_k\in C_c^\infty(R^d\times R^d),\psi\in C^2(R^m)\},$$

we define

$$(3.2)\qquad Af(\rho) = \Big\langle \rho, v\cdot\nabla_x\frac{\delta f}{\delta\rho}(x,v) - (\rho_X * \nabla\bar\Phi_1 + \nabla\bar\Psi_1)(x,y)\cdot\nabla_v\frac{\delta f}{\delta\rho} + a(\rho,x)\cdot D^2_{vv}\frac{\delta f}{\delta\rho}(x,v)\Big\rangle$$

with $\bar\Phi_1,\bar\Psi_1$ given by (1.7) and square matrix $a(\rho,x)$ given by (1.15).

This section proves the following

Lemma 3.1. *For each $f\in D_0\subset C_b(E)$, there exists $g,h\in C_b(E')$ (given by (3.5) and (3.6) below). Let*

$$(3.3)\qquad f_n(\gamma) = f(\rho) + n^{-1}g(\gamma) + n^{-2}h(\gamma).$$

Then

$$\lim_{n\to+\infty} A_n f_n(\gamma_n) = Af(\rho),$$

whenever $\rho_n(dx,dv) = \gamma_n(dx,dv,S) \Rightarrow \rho(dx,dv)$ and $\sup_n\int|v|d\gamma_n < +\infty$. Moreover, there exists $C_0,C_1>0$,

$$(3.4)\qquad |A_nf_n(\gamma)| \le C_0 + n^{-1}C_1\int|v|d\gamma,\quad \gamma\in E_n'.$$

We prove the above lemma. Let

$$(3.5)\qquad g(\gamma) = -\int\int\Big(P_{\nu\otimes\nu}\nabla\Phi_2(x-\bar x;y,\bar y) + P_\nu\nabla\Psi_2(x,y)\Big)\cdot\nabla_v\frac{\delta f}{\delta\rho}(x,v)$$
$$\gamma(dx,dv,dy)\gamma(d\bar x,d\bar v,d\bar y),$$

where $P_{\nu\otimes\nu}\nabla\Phi_2(x;y,\bar{y}) = \int_{z\in S}\int_{\bar{z}\in S}\nabla\Phi_2(x;z,\bar{z})\nu(y,dz)\nu(\bar{y},d\bar{z})$; and

$$\begin{aligned}
(3.6)\quad & h(\gamma) \\
&= \int P_\nu(-\nabla\Psi_1)(x,y)\cdot\nabla_v\frac{\delta f}{\delta\rho}(x,v)\gamma(dx,dv,dy) \\
&\quad + \int\int P_{\nu\otimes\nu}\nabla\Phi_1(x-\bar{x};y,\bar{y})\cdot\nabla_v\frac{\delta f}{\delta\rho}(x,v)\gamma(dx,dv,dy)\gamma(d\bar{x},d\bar{v},d\bar{y}) \\
&\quad + \int\int\int P_{\nu\otimes\nu\otimes\nu}a_1(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}})\cdot D^2_{vv}\frac{\delta f}{\delta\rho}(x,v) \\
&\qquad\qquad \gamma(dx,dv,dy)\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}}) \\
&\quad + \int\int\int\int P_{\nu\otimes\nu\otimes\nu\otimes\nu}a_2(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}};\hat{x},\hat{y})\cdot\nabla_{\bar{\bar{v}}}\nabla_v\frac{\delta^2 f}{\delta\rho^2}(x,v;\bar{\bar{x}},\bar{\bar{v}}) \\
&\qquad\qquad \gamma(dx,dv,dy)\gamma(d\hat{x},d\hat{v},d\hat{y})\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}}).
\end{aligned}$$

See (3.15) and (3.16) for the definition of matrices a_1, a_2. In the above (and below), we denote matrix

$$\nabla_{\bar{\bar{v}}}\nabla_v\frac{\delta^2 f}{\delta\rho^2}(x,v;\bar{\bar{x}},\bar{\bar{v}}) = \Big(\frac{\partial^2}{\partial v_i\partial\bar{\bar{v}}_j}\frac{\delta^2 f}{\delta\rho^2}(x,v;\bar{\bar{x}},\bar{\bar{v}})\Big)_{d\times d}.$$

Therefore,

$$\begin{aligned}
& P_{\nu\otimes\nu\otimes\nu\otimes\nu}a_2(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}};\hat{x},\hat{y})\cdot\nabla_{\bar{\bar{v}}}\nabla_v\frac{\delta^2 f}{\delta\rho^2}(x,v;\bar{\bar{x}},\bar{\bar{v}}) \\
&= \sum_{k,l=1}^m \partial^2_{kl}\psi(\langle\rho,\varphi_1\rangle,\ldots,\langle\rho,\varphi_m\rangle) \\
&\qquad \nabla_{\bar{\bar{v}}}\varphi_k(\bar{\bar{x}},\bar{\bar{v}})\cdot P_{\nu\otimes\nu\otimes\nu\otimes\nu}a_2(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}};\hat{x},\hat{y})\cdot\nabla_v\varphi_l(x,v).
\end{aligned}$$

Since $\delta f/\delta\rho\in C_c(R^d\times R^d)$, $\delta^2 f/\delta\rho^2\in C_c(R^d\times R^d\times R^d\times R^d)$, both $g,h\in C_b(E'_n)$. Since

$$\begin{aligned}
(3.7)\quad & A_n f_n(\gamma) \\
&= A_n f(\rho) + n^{-1}A_n g(\gamma) + n^{-2}A_n h(\gamma) \\
&= \Big\langle\gamma, v\cdot\nabla_x\frac{\delta f}{\delta\rho} - \Big(\int\nabla\Phi_1(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y}) + \nabla\Psi_1\Big)\cdot\nabla_v\frac{\delta f}{\delta\rho}\Big\rangle \\
&\quad + \Big\langle\gamma, -\Big(\int_{R^d\times R^d\times S}\nabla\Phi_2(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y}) + \nabla\Psi_2\Big)\cdot\nabla_v\frac{\delta g}{\delta\gamma}\Big\rangle \\
&\qquad\qquad + \Big\langle\gamma, B_y\frac{\delta h}{\delta\gamma}\Big\rangle \\
&\quad + n\Big\langle\gamma, -\Big(\int_{R^d\times R^d\times S}\nabla\Phi_2(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y}) + \nabla\Psi_2\Big)\cdot\nabla_v\frac{\delta f}{\delta\rho} \\
&\qquad\qquad + B_y\frac{\delta g}{\delta\gamma}\Big\rangle + o(1),
\end{aligned}$$

where

$$
\begin{aligned}
& o(1) \\
&= n^{-1}\Big\{\Big\langle \gamma, v\cdot\nabla_x \frac{\delta g}{\delta\gamma} - \Big(\int \nabla\Phi_1(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y}) + \nabla\Psi_1\Big)\cdot\nabla_v \frac{\delta g}{\delta\gamma}\Big\rangle \\
&\quad + \frac{n^2}{2N}\Big(\Big\langle \gamma, B\frac{\delta^2 g}{\delta\gamma^2}(x,v,\cdot;x,v,\cdot)(y)\Big\rangle \\
&\quad - 2\Big\langle \gamma(dx,dv,dy)\otimes\delta_{\{x,v,y\}}(d\bar{x},d\bar{v},d\bar{y}), B_y\frac{\delta^2 g}{\delta\gamma^2}(x,v,y;\bar{x},\bar{v},\bar{y})\Big\rangle\Big)\Big\} \\
&\quad + \frac{1}{n^2}\Big\{\Big\langle \gamma, v\cdot\nabla_x \frac{\delta h}{\delta\gamma} - \Big(\int \nabla\Phi_1(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y}) + \nabla\Psi_1)\Big)\cdot\nabla_v \frac{\delta h}{\delta\gamma}\Big\rangle \\
&\quad - n\Big\langle \gamma, \Big(\int_{R^d\times R^d\times S} \nabla\Phi_2(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y}) + \nabla\Psi_2\Big)\cdot\nabla_v \frac{\delta h}{\delta\gamma}\Big\rangle \\
&\quad + \frac{n^2}{2N}\Big(\Big\langle \gamma, B\frac{\delta^2 h}{\delta\gamma^2}(x,v,\cdot;x,v,\cdot)(y)\Big\rangle \\
&\quad - 2\Big\langle \gamma(dx,dv,dy)\otimes\delta_{\{x,v,y\}}(d\bar{x},d\bar{v},d\bar{y}), B_y\frac{\delta^2 h}{\delta\gamma^2}(x,v,y;\bar{x},\bar{v},\bar{y})\Big\rangle\Big)\Big\}.
\end{aligned}
$$

If g is taken to be (3.5), then following Taylor expansion (1.16), we identify

$$
\begin{aligned}
\frac{\delta g}{\delta\gamma}(x,v,y) &= -\int P_{\nu\otimes\nu}\nabla\Phi_2(x-\bar{x};y,\bar{y})\cdot\nabla_v\frac{\delta f}{\delta\rho}(x,v)\gamma(d\bar{x},d\bar{v},d\bar{y}) \\
&\quad - \int (P_{\nu\otimes\nu}\nabla\Phi_2(\bar{\bar{x}}-x;\bar{\bar{y}},y)\nabla_{\bar{\bar{v}}}\frac{\delta f}{\delta\rho}(\bar{\bar{x}},\bar{\bar{v}})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}}) \\
&\quad - P_\nu\nabla\Psi_2(x,y)\cdot\nabla_v\frac{\delta f}{\delta\rho}(x,v) \\
&\quad - \int\int \Big(P_{\nu\otimes\nu}\nabla\Phi_2(\bar{\bar{x}}-\bar{x};\bar{\bar{y}},\bar{y}) + P_\nu\nabla\Psi_2(\bar{\bar{x}},\bar{\bar{y}})\Big) \\
&\qquad \cdot\nabla_{\bar{\bar{v}}}\frac{\delta^2 f}{\delta\rho^2}(\bar{\bar{x}},\bar{\bar{v}};x,v)\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}})\gamma(d\bar{x},d\bar{v},d\bar{y}).
\end{aligned}
\tag{3.8}
$$

Similarly, we can also compute $\delta^2 g/\delta\gamma^2$ explicitly. From the conditions on Φ_2, Ψ_2 (Condition 1.2), it follows that

$$
\nabla_x\frac{\delta g}{\delta\gamma}(x,v;y), \nabla_v\frac{\delta g}{\delta\gamma}(x,v;y), B\frac{\delta^2 g}{\delta\gamma^2}(x,v,\cdot;x,v,\cdot)(y), B_y\frac{\delta^2 g}{\delta\gamma^2}(x,v,y;\bar{x},\bar{v},\bar{y}) \tag{3.9}
$$

are all bounded over their respective domain. Analogous estimates holds when g is replaced by the h in (3.6). Note that because $\delta f/\delta\rho, \delta^2 f/\delta\rho^2$ have compact support,

$$
\nabla_v\frac{\delta g}{\delta\gamma}(x,v;y) \in C_c(R^d\times R^d\times S). \tag{3.10}
$$

To emphasize the dependence of $o(1)$ on n and on γ, we write it $o(1;n,\gamma)$. By (3.9) and similar bounded estimates for h and by (3.10), there exists $C_1, C_2, C_3 > 0$,

$$
|o(1;n,\gamma)| \le n^{-1}(C_2 + C_1\int |v|d\gamma) + \frac{n}{N}C_3, \quad \gamma\in E_n'. \tag{3.11}
$$

Therefore

$$
\lim_{n\to+\infty} \sup_{\gamma\in E_n', \int |v|d\gamma\le C} |o(1;n,\gamma)| = 0, \quad \forall C>0.
$$

Hence $o(1)$ is a higher order term.

First of all, the g is chosen so that

$$\Big\langle \gamma, -\Big(\int_{R^d\times R^d\times S} \nabla\Phi_2(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y}) + \nabla\Psi_2(x,y)\Big)\cdot\nabla_v \frac{\delta f}{\delta\gamma} + B_y \frac{\delta g}{\delta\gamma}\Big\rangle = 0. \tag{3.12}$$

This can be directly verified as $\delta g/\delta\gamma$ is given by (3.8) and (1.6) and (1.9) hold,

$$\begin{aligned}
&\int B_y \frac{\delta g}{\delta\gamma}(x,v,y)\gamma(dx,dv;dy) \qquad (3.13)\\
&= \int\int \Big\{ \Big(B_y(-P_{\nu\otimes\nu}\nabla\Phi_2(x-\bar{x};y,\bar{y})) \\
&\qquad + B_{\bar{y}}(-P_{\nu\otimes\nu}\nabla\Phi_2(x-\bar{x};y,\bar{y}))\Big)\cdot\nabla_v\frac{\delta f}{\delta\rho}(x,v) \\
&\qquad + B_y(-P_\nu\nabla\Psi_2(x,y))\cdot\frac{\delta f}{\delta\rho}(x,v)\Big\}\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(dx,dv,dy) \\
&= \int\int (\nabla\Phi_2(x-\bar{x};y,\bar{y}) + \nabla\Psi_2(x,y))\cdot\nabla_v\frac{\delta f}{\delta\rho}(x,v)\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(dx,dv,dy).
\end{aligned}$$

$A_n f_n$ now simplifies into

$$\begin{aligned}
&A_n f_n(\gamma) \qquad (3.14)\\
&= \Big\langle \gamma, v\cdot\nabla_x\frac{\delta f}{\delta\rho} - \Big(\int \nabla\Phi_1(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y}) + \nabla\Psi_1(x;y)\Big)\cdot\nabla_v\frac{\delta f}{\delta\rho}\Big\rangle \\
&\quad + \Big\langle \gamma, \int\int a_1(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}})\cdot D^2_{vv}\frac{\delta f}{\delta\rho}(x,v)\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}})\Big\rangle \\
&\quad + \langle \gamma, \int\int\int a_2(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}};\hat{x},\hat{y})\cdot\nabla_{\bar{\bar{v}}}\nabla_v\frac{\delta^2 f}{\delta\rho^2}(x,v;\bar{\bar{x}},\bar{\bar{v}}) \\
&\qquad \gamma(d\hat{x},d\hat{v},d\hat{y})\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}})\Big\rangle \\
&\quad + \Big\langle\gamma, B_y\frac{\delta h}{\delta\gamma}\Big\rangle + o(1).
\end{aligned}$$

where

$$D^2_{vv}\varphi(x,v) = \Big(\frac{\partial^2}{\partial v^{(i)}\partial v^{(j)}}\varphi(x,v)\Big)_{d\times d}, \quad v = (v^{(1)},\dots,v^{(d)}).$$

and matrices

$$a_1(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}}) = \Big(\nabla\Phi_2(x-\bar{x};y,\bar{y}) + \nabla\Psi_2(x,y)\Big) \times \Big(P_{\nu\otimes\nu}\nabla\Phi_2(x-\bar{\bar{x}};y,\bar{\bar{y}}) + P_\nu\nabla\Psi_2(x,y)\Big) \tag{3.15}$$

(matrix $\xi\times\eta = (\xi_i\eta_j)_{ij}$ for vectors $\xi = (\xi_1,\dots,\xi_d)$, $\eta = (\eta_1,\dots,\eta_d)\in R^d$), and

$$a_2(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}};\hat{x},\hat{y}) = \Big(\nabla\Phi_2(x-\bar{x};y,\bar{y}) + \nabla\Psi_2(x,y)\Big) \times \Big(P_{\nu\otimes\nu}\nabla\Phi_2(\bar{\bar{x}}-\hat{x};\bar{\bar{y}},\hat{y}) + P_\nu\nabla\Psi_2(\bar{\bar{x}},\bar{\bar{y}})\Big). \tag{3.16}$$

In the above simplification, we used

$$
\begin{aligned}
(3.17)\quad & \nabla_v \frac{\delta g}{\delta\gamma}(x,v,y) \\
&= \Big(- \int P_{\nu\otimes\nu} \nabla\Phi_2(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y}) - P_\nu \nabla\Psi_2(x,y)\Big)\cdot D_{vv}\frac{\delta f}{\delta\rho}(x,v) \\
&\quad - \nabla_v \int\int \Big(P_{\nu\otimes\nu}\nabla\Phi_2(\bar{\bar{x}}-\bar{x};\bar{\bar{y}},\bar{y}) + P_\nu\nabla\Psi_2(\bar{\bar{x}},\bar{y})\Big) \\
&\qquad \cdot \nabla_{\bar{\bar{v}}}\frac{\delta^2 f}{\delta\rho^2}(\bar{\bar{x}},\bar{\bar{v}};x,v)\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}})\gamma(d\bar{x},d\bar{v},d\bar{y}).
\end{aligned}
$$

The idea of choosing h is such that limit of $A_n f_n$ only depends on ρ (the X, V-marginal of γ). To achieve this, we need to "average out" all the $y, \bar{y}, \bar{\bar{y}}, \hat{y}$ dependence in (3.14). We verify this next.

First, we make three interesting observations (3.18), (3.19) and (3.20) in order to simplify the expression of $\langle \gamma, B_y \delta h/\delta\gamma\rangle$. Let $\varphi_1 = \varphi_1(y,\hat{y},\bar{y},\bar{\bar{y}}) \in C_b(S^4)\cap D(\mathbf{B})$, and $\varphi_2 = \varphi_2(\rho; x, v; \bar{\bar{x}}, \bar{\bar{v}}) \in C_b(E\times(R^d\times R^d)^2)$ where $\delta\varphi_2/\delta\rho$ is still a bounded function. By (1.6),

$$
\begin{aligned}
(3.18)\quad & \langle \gamma, B\frac{\delta}{\delta\gamma}\int\int\int\int P_{\nu\otimes\nu\otimes\nu\otimes\nu}\varphi_1(y,\hat{y},\bar{y},\bar{\bar{y}})\varphi_2(\rho;x,v;\bar{\bar{x}},\bar{\bar{v}}) \\
&\qquad \gamma(dx,dv,dy)\gamma(d\hat{x},d\hat{v},d\hat{y})\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}})\rangle \\
&= \int\int\int\int \Big(\mathbf{B}P_{\nu\otimes\nu\otimes\nu\otimes\nu}\varphi_1(y,\hat{y},\bar{y},\bar{\bar{y}})\Big)\varphi_2(\rho;x,v;\bar{\bar{x}},\bar{\bar{v}}) \\
&\qquad \gamma(dx,dv,dy)\gamma(d\hat{x},d\hat{v},d\hat{y})\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}}) \\
&= \int\int\int\int \Big(- \varphi_1(y,\hat{y},\bar{y},\bar{\bar{y}}) + P_{\pi_0\otimes\pi_0\otimes\pi_0\otimes\pi_0}\varphi_1\Big)\varphi_2(\rho;x,v;\bar{\bar{x}},\bar{\bar{v}}) \\
&\qquad \gamma(dx,dv,dy)\gamma(d\hat{x},d\hat{v},d\hat{y})\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}}).
\end{aligned}
$$

We note that

$$
\begin{aligned}
(3.19)\quad & P_{\pi_0\otimes\pi_0\otimes\pi_0\otimes\pi_0} a_2(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}};\hat{x},\hat{y}) \\
&= \Big\{ \Big(P_{\pi_0\otimes\pi_0}\mathbf{B}(-P_{\nu\otimes\nu}\nabla\Phi_2(x-\bar{x};y,\bar{y}) - P_\nu\nabla\Psi_2(x,y))\Big) \\
&\qquad \times\Big(P_{\pi_0\otimes\pi_0}(P_{\nu\otimes\nu}\nabla\Phi_2(\bar{\bar{x}}-\hat{x};\bar{\bar{y}},\hat{y}) + P_\nu\nabla\Psi_2(\bar{\bar{x}},\bar{\bar{y}}))\Big)\Big\} = 0,
\end{aligned}
$$

and similarly

$$
\begin{aligned}
(3.20)\quad & \int\int P_{\pi_0\otimes\pi_0\otimes\pi_0} a_1(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}})\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}}) \\
&= \int_{y\in S}\Big\{\Big(- BP_\nu(\rho_X * \nabla\bar{\bar{\Phi}}_2(x;y) + \nabla\Psi_2(x,y))\Big) \\
&\qquad \times\Big(P_\nu(\rho_X*\nabla\bar{\bar{\Phi}}_2(x;y)+\nabla\Psi_2(x,y))\Big)\Big\}\pi_0(dy),
\end{aligned}
$$

where $\bar{\bar{\Phi}}_2(x;y)$ is defined in (1.14).

In view of the form of h in (3.6), the above three identities enable us to simplify (3.14) into

$$
(3.21)\qquad A_n f_n(\gamma) = Af(\rho) + o(1).
$$

4. Energy estimates and tightness

In this section, we prove

Theorem 4.1. *Suppose that (4.4) hold. Then the processes* $\{\rho_n(\cdot) : n = 1, 2, \ldots\}$ *is tight with trajectories in space* $C_{\mathcal{P}(R^d \times R^d)}[0, +\infty)$.

Following the general method given by Theorem 9.1 in Chapter 3 of [4], there are two key steps in proving tightness for processes. First, we prove a compact containment property (Lemma 4.2) by studying the time evolution of an energy function which is effectively averaged at the macroscopic scale.

Let

$$\mathcal{E}(\rho) = \int_{x,v} \int_{\bar{x},\bar{v}} \Big(\frac{1}{2}|v|^2 + \frac{1}{2}\bar{\Phi}_1(x-\bar{x}) + \bar{\Psi}_1(x)\Big)\rho(d\bar{x}, d\bar{v})\rho(dx, dv), \tag{4.1}$$

and

$$V(\rho) = \log(1 + \mathcal{E}(\rho)). \tag{4.2}$$

Since $\bar{\Phi}_1, \bar{\Psi}_1 \geq 0$ and $\bar{\Psi}_1$ has compact level set in R^d (Condition 1.2.4 is crucial here), $\mathcal{E}$ and V have compact level sets in $\mathcal{P}(R^d \times R^d)$ under the weak convergence of probability measure topology.

We now define stopping time

$$\tau_{n,M} = \inf\{t \geq 0 : V(\rho_n(t)) > M\}. \tag{4.3}$$

Lemma 4.2. *Suppose that*

$$\sup_n E[V(\rho_n(0))] < +\infty. \tag{4.4}$$

Then

$$\lim_{M \to +\infty} \sup_n P(\tau_{n,M} \leq T) = 0, \qquad \forall T > 0. \tag{4.5}$$

Consequently, we have a compact containment property: for each $T > 0$, $\epsilon > 0$, *there exists compact set* $K = K(T, \epsilon) \subset \mathcal{P}(R^d \times R^d)$ *such that*

$$P(\exists t, 0 \leq t \leq T, \rho_n(t) \notin K) < \epsilon.$$

Furthermore, we have energy estimate

$$E[V(\rho_n(t))] \leq E[V(\rho_n(0))] + t(n^{-1}C_0 + \|\mathrm{tr}(a_1)\|_\infty) \tag{4.6}$$

where $C_0 > 0$ *is some deterministic constant depending on* Ψ_i, Φ_i *and* B.

We prove this result using a stochastic Lyapunov function technique. Let $\rho \in E_n$. First, we note that for $\varphi(x, v) : R^d \times R^d \mapsto R$, $\int \varphi d\rho = N^{-1} \sum_{i=1}^N \varphi(x_i, v_i) < +\infty$. In particular, $\int (\bar{\Psi}_1(x) + |v|^2) d\rho < +\infty$.

Proof. From the definition of the first two variational derivatives in (1.16), for $\rho \in E_n$ and $\gamma \in E'_n$ (see (1.19) and (1.18)),

$$\frac{\delta V}{\delta \rho}(x, v) = (1 + \mathcal{E}(\rho))^{-1} \left(\frac{1}{2}|v|^2 + \int_{\bar{x}} \bar{\Phi}_1(x - \bar{x})\rho_X(d\bar{x}) + \bar{\Psi}_1(x)\right),$$

and

$$\frac{\delta^2 V}{\delta\rho^2}(x,v;\bar{x},\bar{v}) = (1+\mathcal{E}(\rho))^{-1}\bar{\Phi}_1(x-\bar{x}) - \frac{\delta V}{\delta\rho}(x,v)\frac{\delta V}{\delta\rho}(\bar{x},\bar{v}).$$

We let $g_V = g_V(\gamma)$ and $h_V = h_V(\gamma)$ be respectively defined as in (3.5) and (3.6) with $\delta f/\delta\rho$ and $\delta^2 f/\delta\rho^2$ replaced by $\delta V/\delta\rho$ and $\delta^2 V/\delta\rho^2$. In view of the special form for $\rho \in E_n, \gamma \in E_n'$,

$$\begin{aligned} g_V(\gamma) &= -(1+\mathcal{E}(\rho))^{-1}\int\int \Big(P_{\nu\otimes\nu}\nabla\Phi_2(x-\bar{x};y,\bar{y}) + P_\nu\nabla\Psi_2(x;y)\Big)\cdot v \\ &\qquad\qquad \gamma(dx,dv,dy)\gamma(d\bar{x},d\bar{v},d\bar{y}) \\ &= -(1+\mathcal{E}(\rho))^{-1}N^{-2}\sum_{i=1}^N\sum_{j=1}^N \Big(P_{\nu\otimes\nu}\nabla\Phi_2(x_i-x_j;y_i,y_j) \\ &\qquad\qquad + P_\nu\nabla\Psi_2(x_i;y_i)\Big)\cdot v_i; \end{aligned}$$

and

$$\sup_n \sup_{\gamma\in E_n'} |g_V(\gamma)| < +\infty.$$

Similarly, for $\gamma \in E_n'$,

$$\begin{aligned} h_V(\gamma) &= (1+\mathcal{E}(\rho))^{-1}\Big\{\int\int v\cdot(P_\nu(-\nabla\Psi_1)(x,y) + P_{\nu\otimes\nu}\nabla\Phi_1(x-\bar{x};y,\bar{y})) \\ &\qquad \gamma(dx,dv,dy)\gamma(d\bar{x},d\bar{v},d\bar{y}) \\ &\quad + \int\int\int \mathrm{Tr}(P_{\nu\otimes\nu\otimes\nu}a_1(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}})) \\ &\qquad \gamma(dx,dv,dy)\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}})\Big\} \\ &\quad + (1+\mathcal{E}(\rho))^{-2}\int\int\int\int P_{\nu\otimes\nu\otimes\nu\otimes\nu}a_2(x,y;\bar{x},\bar{y};\bar{\bar{x}},\bar{\bar{y}};\hat{x},\hat{y})\cdot(v\times\bar{\bar{v}}) \\ &\qquad \gamma(dx,dv,dy)\gamma(d\hat{x},d\hat{v},d\hat{y})\gamma(d\bar{x},d\bar{v},d\bar{y})\gamma(d\bar{\bar{x}},d\bar{\bar{v}},d\bar{\bar{y}}); \end{aligned}$$

and it follows

$$\sup_n \sup_{\gamma\in E_n'} |h_V(\gamma)| < +\infty.$$

Furthermore, following similar computations as in (3.8) and (3.17), we can verify

$$\begin{aligned} \sup_n \sup_{\gamma\in E_n'} \Big(&\Big|\Big\langle\gamma, v\cdot\nabla_x\frac{\delta g_V}{\delta\gamma}\Big\rangle\Big| + \Big|\Big\langle\gamma, \nabla\Psi_1(x,y)\cdot\nabla_v\frac{\delta g_V}{\delta\gamma}\Big\rangle\Big| \\ &+ \Big|\Big\langle\gamma, B_y\frac{\delta^2 g_V}{\delta\gamma^2}(x,v,y;x,v,y)\Big\rangle\Big| \\ &+ \Big|\Big\langle\gamma(dx,dv,dy)\otimes\delta_{\{x,v,y\}}(d\bar{x},d\bar{y},d\bar{y}), B_y\frac{\delta^2 g_V}{\delta\gamma^2}(x,v,y;\bar{x},\bar{v},\bar{y})\Big\rangle\Big|\Big) < +\infty. \end{aligned}$$

Analogous estimates hold when g_V is replaced by h_V as well.

Let

$$V_n(\gamma) = V(\rho) + n^{-1}g_V(\gamma) + n^{-2}h_V(\gamma).$$

We extend the definition of A_n to V_n according to the second expression in (2.2) where everything is expressed in terms of $\delta V_n/\delta\gamma$ and $\delta^2 V_n/\delta\gamma^2$. Recall that for each fixed n, $\gamma_n(t)$ is a probability measure concentrated only on a finite number of points. By the usual finite dimensional Ito's formula,

$$V_n(\gamma_n(t\wedge\tau_{n,M})) - V_n(\gamma_n(0)) - \int_0^{t\wedge\tau_{n,M}} A_nV_n(\gamma_n(s))ds$$

is a martingale. We estimate A_nV_n next.

Following similar computations as in the previous section, we have for $\gamma\in E'_n$,

$$\begin{aligned}
(4.7)\ & A_nV_n(\gamma)\\
&= \Big\langle \rho, v\cdot\nabla_x\frac{\delta V}{\delta\rho}(x,v) - \Big(\rho * \nabla\bar{\Phi}_1(x) + \nabla\bar{\Psi}_1(x)\Big)\cdot\nabla_v\frac{\delta V}{\delta\rho}(x,v)\\
&\qquad + a(\rho,x)\cdot D^2_{vv}\frac{\delta V}{\delta\rho}(x,v)\Big\rangle\\
&\quad + \frac{1}{n}\Big\{\Big\langle\gamma, v\cdot\nabla_x\frac{\delta g_V}{\delta\gamma} - \Big(\int\nabla\Phi_1(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y})\\
&\qquad + \nabla\Psi_1(x;y)\Big)\cdot\nabla_v\frac{\delta g_V}{\delta\gamma}\Big\rangle + \frac{n^2}{2N}\Big(\Big\langle\gamma, B\frac{\delta^2 g_V}{\delta\gamma^2}(x,v,\cdot;x,v,\cdot)(y)\Big\rangle\\
&\qquad - 2\Big\langle\gamma(dx,dv,dy)\otimes\delta_{\{x,v,y\}}(d\bar{x},d\bar{v},d\bar{y}), B_y\frac{\delta^2 g_V}{\delta\gamma^2}(x,v,y;\bar{x},\bar{v},\bar{y})\Big\rangle\Big)\Big\},\\
&\quad + \frac{1}{n^2}\Big\{\Big(\Big\langle\gamma, v\cdot\nabla_x\frac{\delta h_V}{\delta\gamma} - \Big(\int\nabla\Phi_1(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y})\\
&\qquad + \nabla\Psi_1(x;y)\Big)\cdot\nabla_v\frac{\delta h_V}{\delta\gamma}\Big\rangle\Big)\\
&\qquad + \frac{n^2}{2N}\Big(\Big\langle\gamma, B\frac{\delta^2 h_V}{\delta\gamma^2}(x,v,\cdot;x,v,\cdot)(y)\Big\rangle\\
&\qquad - 2\Big\langle\gamma(dx,dv,dy)\otimes\delta_{\{x,v,y\}}(d\bar{x},d\bar{v},d\bar{y}), B_y\frac{\delta^2 h_V}{\delta\gamma^2}(x,v,y;\bar{x},\bar{v},\bar{y})\Big\rangle\Big)\Big\}\\
&\qquad + \frac{1}{n}\Big\{\Big\langle\gamma, \Big(\int_{R^d\times R^d\times S}\nabla\Phi_2(x-\bar{x};y,\bar{y})\gamma(d\bar{x},d\bar{v},d\bar{y})\\
&\qquad + \nabla\Psi_2(x,y)\Big)\cdot\nabla_v\frac{\delta h_V}{\delta\gamma}\Big\rangle\Big\}\\
&\le \|\mathrm{Tr}(a_1)\|_\infty + n^{-1}C_0
\end{aligned}$$

where $C_0 > 0$ is some deterministic constant depending on Ψ_i, Φ_i and B.

Consequently,

$$E[V_n(\gamma_n(t\wedge\tau_{n,M})] \le E[V_n(\gamma_n(0))] + t(n^{-1}C_0 + \|\mathrm{tr}(a_1)\|_\infty). \tag{4.8}$$

For each n fixed, $V(\rho_n(t))$ is a process which is continuous in time. Let $T, M > 0$,

$$\begin{aligned}
&MP(\tau_{n,M}\le T) \le E[V(\rho_n(\tau_{n,M}))\chi_{\tau_{n,M}\le T}] \le E[V(\rho_n(T\wedge\tau_{n,M}))\chi_{\tau_{n,M}\le T}]\\
&\le E[V_n(\gamma_n(T\wedge\tau_{n,M}))] + \sup_m\sup_{\gamma\in E'_m}(n^{-1}|g_V(\gamma)| + n^{-2}|h_V(\gamma)|)\\
&\le E[V_n(\gamma_n(0))] + T(n^{-1}C_0 + \|\mathrm{tr}(a_1)\|_\infty) + \sup_m\sup_{\gamma\in E'_m}(n^{-1}|g_V(\gamma)| + n^{-2}|h_V(\gamma)|).
\end{aligned}$$

Therefore (4.5) follows. Let $\epsilon > 0$, by selecting M large enough, then the compact containment property also follows. Finally, (4.6) follows from (4.8) by taking $M\to+\infty$ and by noting lower semicontinuity of V. □

We choose a special $M = M_n = 2\log(n)$, and denote

$$\tau_n = \tau_{n,M_n}.$$

Then from this definition and the definition of V,

$$\chi_{r\le\tau_n} n^{-1} e^{V(\rho_n(r))/2} \le n^{-1}e^{M_n/2} = 1, \quad r > 0. \tag{4.9}$$

Lemma 4.3. *For each $f \in D_0$ (see (3.1)), $\{f(\rho_n(\cdot \wedge \tau_n)) : n = 1, 2, \ldots\}$ is a sequence of relatively compact real valued processes.*

Proof. We apply Theorem 9.4 in Chapter 3 of Ethier and Kurtz [4].

Let $T > 0$. Let f_n be given as in (3.3) with g, h defined by (3.5), (3.6). First, we note

$$\sup_n E[\sup_{0\le t\le T} |f_n(\gamma_n(t\wedge\tau_n)) - f(\rho_n(t\wedge\tau_n))|] \le n^{-1}\|g\| + n^{-2}\|h\| \to 0, \text{ as } n \to +\infty.$$

This verifies (9.17) in the above mentioned Theorem 9.4 of [4].

Since

$$f_n(\gamma_n(t\wedge\tau_n)) - f_n(\gamma_n(0)) - \int_0^{t\wedge\tau_n} A_n f_n(\gamma_n(r))dr$$

is a martingale. By estimate (3.4) (Lemma 3.1) and noting (4.9), for $T > 0$,

$$\begin{aligned}
&\sup_{n=1,2,\ldots} E\Big[\sup_{0\le t\le T}\Big|\int_0^t \chi(r\le\tau_n)A_n f_n(\gamma_n(r))dr\Big|\Big] \\
&\le \sup_n E\Big[\sup_{0\le t\le T}\Big|\int_0^t \chi(r\le\tau_n)\Big(C_0 + n^{-1}C_1\sqrt{\int |v|^2 d\gamma_n(r)}\Big)dr\Big|\Big] \\
&\le \sup_n E\Big[\sup_{0\le t\le T}\Big|\int_0^t \chi(r\le\tau_n)(C_0 + \sqrt{2}n^{-1}C_1 \exp\{V(\rho_n(r))/2\})dr\Big|\Big] \\
&\le (C_0 + 2C_1)T < +\infty.
\end{aligned}$$

This verifies (9.18) in Theorem 9.4 in Chapter 3 of [4].

The conclusion of this lemma follows from the above two estimates. □

We conclude the tightness property in Theorem 4.1.

First, we claim that $\{\rho_n(\cdot\wedge\tau_n) : n = 1, 2, \ldots\}$ is tight in $C_{\mathcal{P}(R^d\times R^d)}[0, +\infty)$. Given Lemmas 4.2 and 4.3, this follows from Theorem 9.1 in Chapter 3 of Ethier and Kurtz [4].

Next, the conclusion of Theorem 4.1 follows from tightness of stopped processes $\{\rho_n(\cdot\wedge\tau_n) : n = 1, 2, \ldots\}$ and observation

$$P(\sup_{0\le t\le T} r(\rho_n(t\wedge\tau_n), \rho_n(t)) > \epsilon) \le P(\tau_n \le T) \to 0, \quad \text{as } n \to +\infty, \quad \forall \epsilon > 0,$$

where r can be any metric given topology of weak convergence of probability measures on $\mathcal{P}(R^d \times R^d)$.

5. Identifying the limit equation

Let f_n be given by (3.3) and $\tau_{n,M}$ be given by (4.3). By Ito's formula,

$$E\Big[\Big(f_n(\gamma_n(t\wedge\tau_{n,M})) - f_n(\gamma_n(s\wedge\tau_{n,M})) - \int_{s\wedge\tau_{n,M}}^{t\wedge\tau_{n,M}} A_n f_n(\gamma_n(r))dr\Big) \Pi_{i=1}^k h_i(\rho_n(t_i))\Big] = 0,$$

and for every $h_1, \ldots, h_k \in C_b(E)$, $0 \le t_1 \le \ldots, t_k \le s \le t$, and $k = 1, 2, \ldots$. Let $\rho_0(\cdot)$ be a limit processes of $\{\rho_n(\cdot) : n = 1, 2, \ldots\}$ (Theorem 4.1). Then by the convergence of $A_n f_n$ to Af in Lemma 3.1,

$$E\Big[\Big(f(\rho_0(t)) - f(\rho_0(s)) - \int_s^t Af(\rho_0(r))dr\Big)\Pi_{i=1}^k h_i(\rho_0(t_i))\Big] = 0,$$

implying ρ_0 is a solution to the martingale problem

$$f(\rho_0(t)) - f(\rho_0(0)) - \int_0^t Af(\rho_0(s))ds = M^f(t) \tag{5.1}$$

where M^f is a martingale with respect to the natural filtration induced by ρ_0.

We notice that A is really a first-order differential operator in infinite dimensions in the sense that $A(fg) = fAg + gAf$, if $f, g \in D(A)$. In particular,

$$A(f^2) - 2fAf \equiv 0, \quad f \in D(A).$$

Therefore, by Exercise 29 on page 93 of Ethier and Kurtz [4], the quadratic variation

$$[M^f, M^f](t) = \int_0^t (Af^2(\rho_0(s)) - 2f(\rho_0(s))Af(\rho_0(s)))ds \equiv 0,$$

implying $M^f \equiv 0$. In particular, taking $f(\rho) = \langle \varphi, \rho \rangle$ for $\varphi \in C_c^\infty(R^d \times R^d)$, this means that ρ is just a weak solution (Schwartz distributional sense) to (1.1). Therefore we arrive at

Theorem 5.1. *Suppose (4.4) holds. Then under Condition 1.2, stochastic process* $\{\rho_n : n = 1, 2, \ldots\}$ *is tight, and every convergent subsequence converges in probability to a weak (i.e. Schwartz distributional) solution of (1.1) in the path space.*

If we can show that weak solution to (1.1) is unique, then the above conclusion can be strengthened from convergence of subsequence to convergence. We pursue this with additional mild conditions on $\Phi_i, \Psi_i, i = 1, 2$ next.

6. Uniqueness

Uniqueness of McKean-Vlasov type PDE with mean-field type interactions (such as (1.1)) can be proved using probabilistic approach. See for instance, a review by Méléard [8]. Here we follow an approach recently proposed by Kurtz [6].

We consider a countably infinite system of stochastic differential equations

$$\begin{aligned} &\dot{x}_i(t) = v_i(t) \\ &dv_i(t) = \rho_X(t) * \nabla\bar{\Phi}_1(x_i(t))dt + \nabla\bar{\Psi}_1(x_i(t))dt + \sqrt{2}a^{1/2}(\rho(t); x_i(t))dW_i(t); \end{aligned} \tag{6.1}$$

where a is given by (1.15). Note that a is non-negative definite square matrix, therefore its square root is well defined. In the above equation,

$$\rho(t, dx, dv) = \lim_{n\to+\infty} \frac{1}{n}\sum_{j=1}^n \delta_{x_i(t), v_i(t)}(dx, dv), \tag{6.2}$$

and $(W_1, W_2, \ldots, W_k, \ldots)$ is a countably infinite sequence of i.i.d. standard Brownian motions. The following additional smoothness condition on Φ_2, Ψ_2 will be useful in this section.

Condition 6.1. *$\Psi_2(x;y), \Phi_2(x;y_1,y_2)$ have continuous derivatives in the x-variable upto the third order. Moreover,*

$$\sup_{x\in R^d;y_1,y_2\in S} |D^1\Phi_2(x;y_1,y_2)| + |D^2\Phi_2(x;y_1,y_2)| + |D^3\Phi_2(x;y_1,y_2)| < +\infty$$

and

$$\sup_{x\in R^d,y\in S} |D^1\Psi_2(x;y)| + |D^2\Psi_2(x;y)| + |D^3\Psi_2(x;y)| < +\infty$$

where all derivatives are taken with respect to x.

With the above condition, by Theorem 5.2.3 of Stroock and Varadhan [9], and by the defining structure of a in (1.15) (also noting (1.6)), $a^{1/2}$ is Lipschitz in the sense

$$|a^{1/2}(\rho;x) - a^{1/2}(\gamma,y)| \le C(|x-y| + d_W(\rho,\gamma)),$$

where d_W is the order-0 Wasserstein metric on $\mathcal{P}(R^d \times R^d)$:

$$(6.3)\quad d_W(\rho,\gamma) = \sup_{f\in C_b(R^d\times R^d),|f|\le 1,|f(x,v)-f(y,u)|\le|x-y|+|v-u|} \Big| \int f d\rho - \int f d\gamma \Big|.$$

Condition 6.2. *$\nabla\bar\Phi_1, \nabla\bar\Psi_1 : R^d \mapsto R$ are Lipschitz continuous.*

By Section 10 of Kurtz and Protter [7], we have the following

Lemma 6.3. *Let Conditions 6.1 and 6.2 hold. Assume that the initial values are random variables and $\{(x_i(0), v_i(0)) : i = 1, 2, \ldots\}$ is a stationary sequence. Then existence and strong uniqueness holds for the system (6.1)–(6.2).*

We denote $E_0 = R^d \times R^d$. The state space of (6.1)-(6.2) is

$$(E_0)^\infty = \{(\vec x,\vec v) = (x_1,\ldots,x_n,\ldots;v_1,\ldots,v_n,\ldots) : (x_k,v_k) \in E_0\},$$

with the usual product topology. From the symmetry of labels in (6.1)-(6.2), if the initial value $(\vec x(0),\vec v(0))$ is exchangeable, then the solution (uniqueness follows from Lemma 6.3) $(\vec x(t),\vec v(t))$ for (6.1)-(6.2) is exchangeable for all $t > 0$. Let ρ be any weak (Schwartz distributional) solution of partial differential equation (1.1). The goal of this section is to apply the Markov mapping theorem of Kurtz [6] to show that (Theorem 6.8)

$$\rho(t) = \lim_{n\to+\infty} \sum_{k=1}^n \delta_{(x_i(t),v_i(t))}.$$

The above limit exists because that $\{(x_i(t), v_i(t)), i = 1, 2, \ldots\}$ is an exchangeable sequence for all $t > 0$. Therefore, the uniqueness of ρ follows from the uniqueness of (6.1)-(6.2) which is verified in Lemma 6.3. The method is known as particle representation method. We provide the details next.

First, we express solution to (6.1)-(6.2) in terms of a martingale problem. We consider

$$(6.4)\quad D(A_0) = \{g(\vec x,\vec v) = \Pi_{k=1}^m \varphi_k(x_k,v_k) : \varphi_k \in C_c^\infty(R^d\times R^d), m = 1,2,\ldots\}.$$

For each such $g \in D(A_0)$, we define

$$(6.5)\qquad A_0 g(\vec x,\vec v) = \sum_{k=1}^m (\Pi_{j\ne k}\varphi_j(x_j,v_j))\Big(v_k\cdot\nabla_x\varphi_k(x_k,v_k) + (\eta(\vec x,\vec v) * \nabla\bar\Phi_1(x_k) + \nabla\bar\Psi_1(x_k))\cdot\nabla_v\varphi_k(x_k,v_k) + a(\eta(\vec x,\vec v),x_k)\cdot D^2_{vv}\varphi_k(x_k,v_k)\Big),$$

where the convolution with $\nabla\bar{\Phi}_1$ is made with respect to spatial marginal of measure $\eta(\vec{x},\vec{v})$ only. In the above, the map $\eta : (E_0)^\infty \mapsto \mathcal{P}(E_0)$ is defined as follow (in the notation of [6], such map is denoted by γ): let $\gamma_0 \in \mathcal{P}(E_0)$ be arbitrary but fixed,

$$\eta(\vec{x},\vec{v}) = \lim_{n\to+\infty} n^{-1}\sum_{k=1}^{n} \delta_{(x_k,v_k)} \in \mathcal{P}(E_0), \quad (\vec{x},\vec{v}) \in (E_0)^\infty$$

if the limit on the right hand side exists in the weak convergence of probability measure sense; and $\eta(\vec{x},\vec{v}) = \gamma_0$ otherwise. By Ito's formula, any solution to the infinite system (6.1)-(6.2) is also a solution to the martingale problem

$$g(\vec{x}(t),\vec{v}(t)) - g(\vec{x}(0),\vec{v}(0)) - \int_0^t A_0 g(\vec{x}(s),\vec{v}(s))ds = \text{martingale}. \tag{6.6}$$

Let

$$\mathcal{H} = \{h(\vec{x},\vec{v}) = f(x_1,v_1,\ldots,x_m,v_m) - f(x_{\sigma_1},v_{\sigma_1},\ldots,x_{\sigma_m},v_{\sigma_m}) : \\ \text{all permutation } \sigma, f \in B((E_0)^m), m = 1,\ldots\}.$$

Assume additionally that the distribution $\nu_0 \in \mathcal{P}((E_0)^\infty)$ for $((x_k(0),v_k(0)) : k = 1,2,\ldots)$ is exchangeable. Because of the symmetry in (6.1)-(6.2), exchangeability of $((x_k(t),v_k(t)) : k = 1,2,\ldots)$ follows for all $t > 0$, therefore

$$E[h(\vec{x}(t),\vec{v}(t))] = 0, \quad t \geq 0, h \in \mathcal{H}. \tag{6.7}$$

We call $(\vec{x},\vec{v})$ a solution to the *restricted* martingale problem for $(A_0,\mathcal{H},\nu_0)$ in the sense that both (6.6) and (6.7) are satisfied.

As in the setting of finite multi-dimensional diffusion processes, we also have the following.

Lemma 6.4. *Let $(\vec{x},\vec{v})$ be a solution to the restricted martingale problem for $(A_0,\mathcal{H},\nu_0)$ with trajectory in $C_{(E_0)^\infty}[0,+\infty)$ and filtration $\{\mathcal{F}_t : t \geq 0\}$, on a probability space $(\Omega,\mathcal{F},P)$. Then there exists countably infinite i.i.d Brownian motions W_is and probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ with filtration $\tilde{\mathcal{F}}_t \supset \mathcal{F}_t$ and $(\tilde{\vec{x}},\tilde{\vec{v}})$ on probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ with filtration $\tilde{\mathcal{F}}_t$. $(\tilde{\vec{x}},\tilde{\vec{v}})$ has the same distribution as $(\vec{x},\vec{v})$, and $(\tilde{\vec{x}},\tilde{\vec{v}})$ satisfies Ito's equation (6.1)-(6.2).*

Proof. The same proof of Proposition 3.1 and Theorem 3.3 in Chapter 5 of Ethier and Kurtz [4] apply to this countably infinite system setting. □

Therefore, by Lemma 6.3, we have

Lemma 6.5. *Let the initial distribution ν_0 be exchangeable. Any solution of the restricted martingale problem for $(A_0,\mathcal{H},\nu_0)$, with trajectory in $C_{(E_0)^\infty}[0,+\infty)$, is unique.*

We now define a transition probability measure $\alpha : \mathcal{P}(E_0) \mapsto (E_0)^\infty$:

$$\alpha(\rho; d\vec{x}, d\vec{v}) = \Pi_{k=1}^\infty \rho(dx_k, dv_k).$$

We verify another condition required by Theorem 3.2, Corollaries 3.5, 3.7 of Kurtz [6].

Lemma 6.6. *The set $\eta^{-1}(\rho) = \{(\vec{x},\vec{v}) \in (E_0)^\infty : \eta(\vec{x},\vec{v}) = \rho\}$ satisfies*

$$\alpha(\rho; \eta^{-1}(\rho)) = 1. \tag{6.8}$$

Proof. We note that if we take $(E_0)^\infty$ to be a probability space endowed with the product measure $\Pi_{i=1}^\infty \rho(dx_i, dv_i)$, then (6.8) holds because of the strong law of large numbers. □

We verify some regularity properties for A_0, which is needed to apply a version of the Markov mapping theorem (Corollary 3.7) of Kurtz [6].

From (6.4), $D(A_0)$ is closed under multiplication and separates points. It is also a pre-generator in the sense of that paper because system (6.1)-(6.2) can be approximated by similar systems when the Brownian motion terms are replaced by centered i.i.d. Poisson processes. Furthermore, we have

Lemma 6.7. *There exists a countable subset $\{g_k\} \subset D(A_0)$ such that the graph of A_0 is contained in the bounded pointwise closure of the linear span of $\{(g_k, A_0 g_k)\}$.*

Proof. Noting the form (6.5), we can always approximate the φ_js by countable sequence of polynomials with rational coefficients. The collection of such polynomials (and hence any finite products of them) is countable. □

Theorem 6.8. *[Uniqueness and particle representation of solution] Let Condition 1.2 hold. Define*

$$C = \Big\{ \Big(\int g(\vec{x}, \vec{v}) \alpha(\cdot; d\vec{x}, d\vec{v}), \int A_0 g(\vec{x}, \vec{v}) \alpha(\cdot, d\vec{x}, d\vec{v}) \Big) : g \in D(A_0) \Big\}.$$

Then

1. *$D(A_0)$ consists of functions of the form*

$$f(\rho) = \int g(\vec{x}, \vec{v}) \alpha(\rho; d\vec{x}, d\vec{v}) = \Pi_{k=1}^m \langle \varphi_k, \rho \rangle,$$

and

$$\begin{aligned} Cf(\rho) &\equiv \int A_0 g(\vec{x}, \vec{v}) \alpha(\rho; d\vec{x}, d\vec{v}) \\ &= \sum_{k=1}^m (\Pi_{j \neq k} \langle \varphi_j, \rho \rangle) \Big(\langle v \cdot \nabla_x \varphi_k + (\rho * \nabla \bar{\Phi}_1 + \nabla \bar{\Psi}_1) \cdot \nabla_v \varphi_k \\ &\qquad + a(\rho, x) \cdot D^2_{vv} \varphi_k), \rho \rangle \Big). \end{aligned}$$

2. *$D(C)$ is closed under multiplication, $Cf^2 = 2fCf$ and for each $f \in D(C)$. Any solution to the martingale problem for C satisfies*

$$f(\rho(t)) - f(\rho(0)) - \int_0^t Cf(\rho(s)) ds = 0, \tag{6.9}$$

which is equivalent to (Schwartz distributional) solution of (1.1).

3. *Let $\rho \in C_{\mathcal{P}(R^d \times R^d)}[0, +\infty)$ be a solution to (6.9) with deterministic initial condition $\rho(0) = \rho_0 \in \mathcal{P}(E_0)$. Let $(\vec{x}, \vec{v})$ be the unique solution (see Lemma 6.3) to (6.1)-(6.2) with initial distribution $(\vec{x}, \vec{v}) \sim \nu_0(d\vec{x}, d\vec{v}) = \Pi_{i=1}^\infty \rho_0(dx_i, dv_i)$. Assume that Conditions 6.1, 6.2 hold. Then the particle representation*

$$\rho(t) = \eta(\vec{x}(t), \vec{v}(t)).$$

holds. Consequently, solution to (6.9), in the path space $C_{\mathcal{P}(R^d \times R^d)}[0, +\infty)$, is unique. That is, weak (Schwartz distributional) solution to (1.1) is unique in $C_{\mathcal{P}(R^d \times R^d)}[0, +\infty)$.

Proof. The first two parts of the theorem follows by direct verification. Since $Cf^2 = 2fCf$, as in Section 5, the martingale appearing in the martingale problem for C has to be zero and (6.9) follows.

Part three follows by applying Corollary 3.7 of Kurtz [6] and by noting existence and uniqueness for solution of restricted martingale problem for $(A_0, \mathcal{H}, \nu_0)$ is equivalent to existence and uniqueness for infinite system (6.1)-(6.2) with initial condition $(\vec{x}(0), \vec{v}(0)) \sim \nu_0$. □

Using the above conclusion, we can strengthen the existence result in Theorem 5.1 to the following.

Theorem 6.9. *Suppose (4.4) hold. Then under Conditions 1.2, and 6.1 and 6.2, partial differential equation (1.1) has a unique solution in* $C_{\mathcal{P}(R^d \times R^d)}[0, +\infty)$. *In addition, the stochastic processes* $\{\rho_n : n = 1, 2, \dots\}$ *in (1.12) converges in probability to such unique solution.*

References

[1] Bensoussan, Alain, Lions, Jacques-Louis and Papanicolaou, George (1978). *Asymptotic Analysis for Periodic Structures.* Studies in Mathematics and its Applications **5**. North-Holland Publishing Co., Amsterdam, New York.

[2] Cercignani, C. (1988). *The Boltzmann Equation and Its Applications.* Applied Mathematical Sciences **67**. Springer-Verlag, New York.

[3] Dawson, Donald A. and Gärtner, Jürgen (1987). Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics* **20** 247–308.

[4] Ethier, Stewart N. and Kurtz, Thomas G. (1986). *Markov Processes: Characterization and Convergence.* John Wiley and Sons, New York.

[5] Kurtz, Thomas G. (1973). Convergence of sequences of semigroups of nonlinear operators with an application to gas kinetics. *Trans. Amer. Math. Soc.* **186** 259–272.

[6] Kurtz, Thomas G. (1998). Martingale problems for conditional distributions of Markov processes. *Electronic Journal of Probability* **3** 1–29.

[7] Kurtz, Thomas G. and Protter, Philip (1996). Limit theorems for solutions of stochastic equations II. *CIME School in Probability, Springer Lecture Notes in Mathematics* **1627** 197–285.

[8] Méléard, Sylvie (1996). Asymptotic behaviour of some interacting particle systems; McKean–Vlasov and Boltzmann models. *Ibid* 42–95.

[9] Stroock, Daniel W. and Varadhan, S. R. S. (1979). *Multidimensional Diffusion Processes.* Springer, New York. Corrected Second Printing, 1997.