

## STOCHASTIC CUCKER-SMALE FLOCKING DYNAMICS OF JUMP-TYPE

MARTIN FRIESEN\*

School of Mathematics and Natural Sciences  
University of Wuppertal  
Gaußstraße 20, 42119 Wuppertal, Germany

OLEKSANDR KUTOVIY

Department of Mathematics  
Bielefeld University  
Universitätsstraße 25, 33615 Bielefeld, Germany

(Communicated by the associate editor name)

**ABSTRACT.** We present a stochastic version of the Cucker-Smale flocking dynamics described by a system of  $N$  interacting particles. The velocity alignment of particles is purely discontinuous with unbounded and, in general, non-Lipschitz continuous interaction rates. Performing the mean-field limit as  $N \rightarrow \infty$  we identify the limiting process with a solution to a nonlinear martingale problem associated with a McKean-Vlasov stochastic equation with jumps. Moreover, we show uniqueness and stability for the kinetic equation by estimating its solutions in the total variation and Wasserstein distance. Finally, we prove uniqueness in law for the McKean-Vlasov equation, i.e. we establish propagation of chaos.

**1. Introduction.** Cucker and Smale postulated in [6, 7] a model for the flocking of birds where convergence to a certain consensus (same direction and velocity in the motion of birds) was shown to depend on the spatial decay of the communication rate between the birds. Putting in abstract mathematical notations, the Cucker-Smale model describes dynamics of  $N$  particles  $(r_k, v_k) \in \mathbb{R}^{2d}$ , where  $r_k$  stands for the position and  $v_k$  for the velocity of the particle with number  $k = 1, \dots, N$ . The time evolution is obtained from the system of ordinary differential equations

$$\begin{cases} \frac{dr_k}{dt} = v_k, \\ \frac{dv_k}{dt} = \frac{1}{N} \sum_{j=1}^N \psi(r_k - r_j)(v_j - v_k). \end{cases} \quad (1)$$

Here  $0 \leq \psi(r) = \psi(-r)$  describes the communication rate between the particles. The natural communication rate is regular and it is given by  $\psi(r) = a(1 + |r|^2)^{-b/2}$ ,  $a, b > 0$ . A more challenging example of a communication rate has singular nature and usually it is defined by  $\psi(r) = a|r|^{-b}$ , see e.g. [23].

---

2010 *Mathematics Subject Classification.* Primary: 35Q83, 60F05; Secondary: 60K35.

*Key words and phrases.* Flocking, Cucker-Smale dynamics, mean-field equation, McKean-Vlasov stochastic equation, propagation of chaos, Total variation distance, Wasserstein distance.

\* Corresponding author: Martin Friesen.

The particular form of (1) implies that the mean velocity is conserved, i.e.

$$v_c := \frac{1}{N} \sum_{k=1}^N v_k(t) = \frac{1}{N} \sum_{k=1}^N v_k(0), \quad \forall t \geq 0.$$

Based on Lyapunov functional techniques corresponding to certain dissipative differential inequalities, the time-asymptotic flocking property

$$\lim_{t \rightarrow \infty} \sum_{k=1}^N |v_k(t) - v_c|^2 = 0 \quad \text{and} \quad \sup_{t \geq 0} \sum_{k=1}^N |r_k(t) - r_c(t)|^2 < \infty \quad (2)$$

was studied in [16], where  $r_c(t) := \frac{1}{N} \sum_{k=1}^N r_k(t) = r_c(0) + tv_c$  denotes the center of mass.

In many cases one seeks to study properties of the particle dynamics in terms of their associated mean-field equations. For the classical Cucker-Smale dynamics the corresponding mean-field equation was (formally) derived in [17] by taking the mean-field limit  $N \rightarrow \infty$ . It was shown that the resulting particle density  $f_t(r, v)$  should solve the kinetic equation

$$\frac{\partial f_t(r, v)}{\partial t} + v \cdot (\nabla_r f_t)(r, v) = \operatorname{div}_v [f_t(r, v) Q(f_t)(r, v)], \quad (3)$$

where the forcing term is given by

$$Q(f_t)(r, v) = \int_{\mathbb{R}^{2d}} (v - w) \psi(r - q) f_t(q, w) dq dw.$$

Unfortunately this formulation only makes sense for solutions possessing enough regularity. For less regular solutions one studies measure-valued solutions  $\mu_t(dr, dv)$  obtained from the kinetic equation (in the weak formulation)

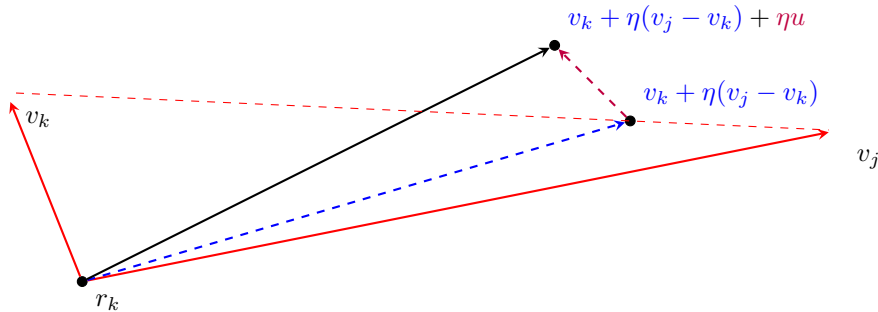
$$\frac{d}{dt} \int_{\mathbb{R}^{2d}} g(r, v, t) \mu_t(dr, dv) = \int_{\mathbb{R}^{2d}} \left( \frac{\partial g(r, v, t)}{\partial t} + B(\mu_t)g(r, v, t) \right) \mu_t(dr, dv), \quad (4)$$

where  $g$  is a compactly supported, continuously differentiable function and  $B(\mu_t) = B_0 + B_1(\mu_t)$  is given by

$$\begin{aligned} B_0 g(r, v, t) &= v \cdot (\nabla_r g)(r, v, t) \\ B_1(\mu_t) g(r, v, t) &= -(\nabla_v g)(r, v, t) \cdot \int_{\mathbb{R}^{2d}} \psi(r - q) (v - w) \mu_t(dq, dw). \end{aligned} \quad (5)$$

Note that, if  $\mu_t(dr, dv)$  has a density  $f_t(r, v)$  for each  $t \geq 0$ , then, using integration by parts, (4) is (formally) equivalent to (3). Existence and uniqueness for measure solutions to (4) was established under the constraint that  $\mu_t$  has compact support (see [16]). For different aspects of this model we refer to [27, 15], while other related models have been studied in [1, 24, 14, 5].

In this work we propose a stochastic version of the Cucker-Smale model where, roughly speaking,  $B(\mu)$  in (5) is replaced by a pure jump operator of mean-field type in the velocity component. Such replacement makes the dynamics stochastic and is motivated by the possibility to describe the effect that certain particles (e.g. birds) may deviate spontaneously from the consensus by moving in the "opposite" direction. This model and its various levels of description (particle dynamics, kinetic equation, and McKean-Vlasov stochastic equation) are introduced and discussed in the next section. Our main results are then formulated in Section 3, where the transition from particle dynamics to the kinetic equation as well as McKean-Vlasov equation is given. Such a transition is based on the mean-field limit and related to

FIGURE 1. Change of velocities with  $\eta = 3/4$ 

propagation of chaos, see [28, 22]. The remaining sections are devoted to the proof of our main results.

Let us mention that the stochastic model investigated in this work shares some (formal) similarities with the Boltzmann-Enskog model studied in [2], and more recently in [12]. In this work we use some techniques developed in [12], however, the model we study is significantly different in regard to its properties. While collisions in the Boltzmann model have a Maxwellian as invariant distribution (and hence are continuously distributed), the flocking behaviour leads to a certain consensus in which all velocities tend to become aligned (see (2)). This effect is (for the classical Cucker-Smale dynamics) also manifested by the role of entropy, see [17, Section 6].

## 2. Stochastic Cucker-Smale dynamics of jump-type.

**2.1. The particle dynamics.** Let  $N \geq 2$  be the number of interacting particles  $(r_k, v_k) \in \mathbb{R}^{2d}$ ,  $k = 1, \dots, N$ . Each particle, say  $(r_k, v_k)$ , may interact with another particle, say  $(r_j, v_j)$ , and the interaction results in a transition of its velocity  $v_k$  to a new velocity  $v^*(v_k, v_j, u)$  computed from

$$v^*(v_k, v_j, u) = v_k + \eta(v_j - v_k) + \eta u = \eta v_j + (1 - \eta)v_k + \eta u. \quad (6)$$

Here  $\eta \in (0, 1]$  parameterizes the point  $\eta v_j + (1 - \eta)v_k$  of the line joining  $v_k$  and  $v_j$ , while  $u \in \mathbb{R}^d$  is the deviation from this convex combination. We suppose that this deviation  $u$  is distributed according to a symmetric probability distribution  $a(u)du$ , i.e.  $a(u) = a(-u)$  is integrable and normalized to 1. Such transition of velocities is summarized in Figure 1 and includes the following important cases:

- (a) If  $\eta = 1$ , then  $v^*(v, w, u) = w + u$ , i.e. the particle takes the velocity of its partner (up to an error of order  $u$ ).
- (b) If  $\eta = \frac{1}{2}$ , then  $v^*(v, w, u) = \frac{v+w}{2} + \frac{u}{2} = v^*(w, v, u)$ , i.e. the particle takes a new velocity given by the mean of incoming velocities (up to an error of order  $\frac{u}{2}$ ).
- (c) The limiting case  $\eta \searrow 0$  corresponds to the classical Cucker-Smale model.

The rate of this velocity alignment is supposed to be proportional to  $\psi(r_k - r_j)\sigma(v_k - v_j)$ , where  $\psi, \sigma \geq 0$  are symmetric functions on  $\mathbb{R}^d$  meaning that  $\psi(r) = \psi(-r)$  and  $\sigma(v) = \sigma(-v)$ . Putting all together, we investigate a  $N$ -particle Markov process in

the phase space  $\mathbb{R}^{2dN}$  given, for  $F \in C_c^1(\mathbb{R}^{2dN})$ , by the Markov generator

$$(LF)(\mathbf{r}, \mathbf{v}) = \sum_{k=1}^N v_k \cdot (\nabla_{r_k} F)(\mathbf{r}, \mathbf{v}) + \frac{1}{N} \sum_{k,j=1}^N \psi(r_k - r_j) \sigma(v_k - v_j) \int_{\mathbb{R}^d} \left[ F(\mathbf{r}, \mathbf{v} + \mathbf{e}_k(v_{k_j}^* - v_k)) - F(\mathbf{r}, \mathbf{v}) \right] a(u) du, \quad (7)$$

where  $v_{k_j}^* = v^*(v_k, v_j, u)$ ,  $(\mathbf{r}, \mathbf{v}) \in \mathbb{R}^{2dN}$  and  $\mathbf{e}_k \in \mathbb{R}^{dN \times dN}$  denotes the  $N \times 1$  matrix whose entries are  $d \times d$  matrices such that  $\mathbf{e}_k = (0_{d \times d}, \dots, 0_{d \times d}, 1_{d \times d}, 0_{d \times d}, \dots, 0_{d \times d})^\top$  with the identity matrix  $1_{d \times d}$  being at the position  $k$ , and  $0_{d \times d}$  denoting the  $d \times d$  zero matrix. The following are our minimal conditions assumed throughout this work:

- (A)  $\psi \geq 0$  is continuous, bounded and symmetric, i.e.  $\psi(r) = \psi(-r)$ .
- (B)  $\sigma \geq 0$  is continuous, symmetric (i.e.  $\sigma(v) = \sigma(-v)$ ), and there exist constants  $c_\sigma > 0$  and  $\gamma \in [0, 2]$  such that

$$\sigma(u) \leq c_\sigma (1 + |u|^2)^{\gamma/2}, \quad u \in \mathbb{R}^d.$$

- (C)  $a \geq 0$  is a symmetric probability density on  $\mathbb{R}^d$ .

For most of the results we also assume that  $a$  has some finite moments, i.e.

$$\lambda_{2p} := \int_{\mathbb{R}^d} (1 + |u|^2)^{p/2} a(u) du < \infty \quad (8)$$

holds for some  $p \geq 0$ . The precise value of  $p$  will be specified in the corresponding statements.

**2.2. Mean-field kinetic equation.** Below we introduce the kinetic equation obtained from the  $N$ -particle dynamics when taking the mean-field limit  $N \rightarrow \infty$ . Precise convergence statements when  $N \rightarrow \infty$  are given in Section 3. Denote by  $\mathcal{P}(\mathbb{R}^{2d})$  the space of probability measures over  $\mathbb{R}^{2d}$  and for  $g \in C_b(\mathbb{R}^{2d})$  set

$$\mathcal{Q}g(r, v; q, w) = \int_{\mathbb{R}^d} \left[ g(r, v^*(v, w, u)) - g(r, v) \right] \psi(r - q) \sigma(v - w) a(u) du.$$

The mean particle distribution is expected to satisfy the following definition.

**Definition 2.1.** Let  $\mu_0 \in \mathcal{P}(\mathbb{R}^{2d})$ . A weak solution to

$$\begin{aligned} \int_{\mathbb{R}^{2d}} g(r, v) \mu_t(dr, dv) &= \int_{\mathbb{R}^{2d}} g(r, v) \mu_0(dr, dv) + \int_0^t \int_{\mathbb{R}^{2d}} v \cdot (\nabla_r g)(r, v) \mu_s(dr, dv) ds \\ &+ \int_0^t \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} \mathcal{Q}g(r, v; q, w) \mu_s(dr, dv) \mu_s(dq, dw) ds, \end{aligned} \quad (9)$$

is a family  $(\mu_t)_{t \geq 0} \subset \mathcal{P}(\mathbb{R}^{2d})$  satisfying

$$\sup_{s \in [0, t]} \int_{\mathbb{R}^{2d}} |v|^\gamma \mu_s(dr, dv) < \infty, \quad \forall t > 0 \quad (10)$$

and such that (9) holds for all  $g \in C_c^1(\mathbb{R}^{2d})$  and  $t > 0$ .

Note that the additional restriction (10) is necessary in order to guarantee that the integrals are well-defined.

**Remark 1.** Suppose that  $\mu_t(dr, dv) = f_t(r, v)drdv$ . Then using integration by parts, one finds for each  $g \in C_c^1(\mathbb{R}^{2d})$

$$\int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} \mathcal{Q}g(r, v; q, w) f_t(r, v) f_t(q, w) drdv dqdw = \int_{\mathbb{R}^{2d}} g(r, v) Q(f_t)(r, v) drdv, \quad (11)$$

where  $Q(f_t)(r, v)$  is the nonlinear integral operator given by

$$\begin{aligned} \int_{\mathbb{R}^{2d} \times \mathbb{R}^d} & \left[ f_t(r, w) f_t \left( q, \frac{v - (1 - \eta)w - \eta u}{\eta} \right) \sigma \left( w - \frac{v - (1 - \eta)w - \eta u}{\eta} \right) \frac{1}{\eta} \right. \\ & \left. - f_t(r, v) f_t(q, w) \sigma(v - w) \right] \psi(r - q) a(u) dudqdw. \end{aligned}$$

Since  $g$  is arbitrary we conclude that the density  $(f_t)_{t \geq 0}$  solves the weak form of the kinetic equation

$$\frac{\partial f_t(r, v)}{\partial t} + v \cdot (\nabla_r f_t)(r, v) = Q(f_t)(r, v). \quad (12)$$

Note that in the limiting case  $\eta \rightarrow 0$  the velocity displacement becomes local again. Hence the case  $\sigma \equiv \eta^{-1}$  with  $\eta \rightarrow 0$  corresponds to the classical Cucker-Smale model. Since the main objective of this work is devoted to the rigorous derivation of the kinetic equation and associated mean-field process, we postpone the study of (12) and its limiting cases for future research.

**2.3. Mean-field process.** While solutions to (9) only describe the mean particle density of the model, the possibility to describe also finite-dimensional distributions is related to the construction of a stochastic process whose time-marginals solve (9). Such construction is for instance well-known for the space-homogeneous Boltzmann equation (see [29]) and has led to many new insights, see [9, 10] and the references therein. An extension of these ideas to the space-inhomogeneous setting (for the Boltzmann-Enskog model) has been recently obtained in [2, 12].

Motivated by this progress, we study its analogue for the kinetic equation (9). For this reason we first reformulate (9) in terms of nonlinear Markov generators. For given  $\mu \in \mathcal{P}(\mathbb{R}^{2d})$ , set

$$A(\mu)g(r, v) = v \cdot \nabla_r g(r, v) + \int_{\mathbb{R}^{2d}} \mathcal{Q}g(r, v; q, w) \mu(dq, dw). \quad (13)$$

Then (9) takes the form

$$\int_{\mathbb{R}^{2d}} g(r, v) \mu_t(dr, dv) = \int_{\mathbb{R}^{2d}} g(r, v) \mu_0(dr, dv) + \int_0^t \int_{\mathbb{R}^{2d}} A(\mu_s)g(r, v) \mu_s(dr, dv) ds. \quad (14)$$

The stochastic process associated to this kinetic equation is then obtained from the nonlinear martingale problem with generator (13).

**Definition 2.2.** Let  $\mu_0 \in \mathcal{P}(\mathbb{R}^{2d})$ . A solution to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$  is a probability measure  $\mu$  on the Skorokhod space  $D(\mathbb{R}_+; \mathbb{R}^{2d})$  such that the following conditions are satisfied

- (i)  $\mu((r(0), v(0)) \in E) = \mu_0(E)$ , for all Borel sets  $E \subset \mathbb{R}^{2d}$ .
- (ii) For all  $t > 0$  it holds

$$\sup_{s \in [0, t]} \mathbb{E}_\mu(|v(s)|^\gamma) < \infty. \quad (15)$$

where  $\mathbb{E}_\mu$  denotes the expectation with respect to  $\mu$  and  $(r, v)$  is the canonical coordinate process on the Skorokhod space.

(iii) For each  $g \in C_c^1(\mathbb{R}^{2d})$ ,

$$g(r(t), v(t)) - g(r(0), v(0)) - \int_0^t (A(\mu_s)g)(r(s), v(s))ds, \quad t \geq 0, \quad (16)$$

is a martingale with respect to  $\mu$ , where  $\mu_s$  denotes the time-marginal of  $\mu$ .

As before, condition (15) is used to guarantee that the integral in (16) makes sense. For additional details and general theory on martingale problems we refer to [8].

**Remark 2.** Let  $\mu$  be a solution to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ . Denote by  $(\mu_t)_{t \geq 0}$  the time-marginals of  $\mu$ . Taking expectations in (16) we find that  $(\mu_t)_{t \geq 0}$  is a weak solution to (9).

While the notion of martingale problems is adequate for the study of convergence and compactness, for other purposes it is more natural to describe the law  $\mu$  as a weak solution to a McKean-Vlasov stochastic equation specified in the following definition.

**Definition 2.3.** A weak solution to the mean-field SDE

$$\begin{cases} R(t) &= R(0) + \int_0^t V(s)ds, \\ V(t) &= V(0) + \int_0^t \int_{[0,1] \times \mathbb{R}^d \times \mathbb{R}_+} \widehat{\alpha}(V(s-), R(s), r_s(\xi), w_s(\xi), u, z) \mathcal{N}(ds, d\xi, du, dz), \end{cases} \quad (17)$$

where

$$\widehat{\alpha} = \eta(u + w_s(\xi) - V(s-)) \mathbb{1}_{[0, \psi(R(s) - r_s(\xi))\sigma(V(s-) - w_s(\xi))]}(z)$$

consists of

- (i) a stochastic basis  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  with the usual conditions,
- (ii) an  $(\mathcal{F}_t)_{t \geq 0}$ -adapted Poisson random measure  $\mathcal{N}$  on  $\mathbb{R}_+ \times [0, 1] \times \mathbb{R}^d \times \mathbb{R}_+$  defined on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  with the compensator  $\widehat{\mathcal{N}}(ds, d\xi, du, dz) = ds d\xi a(u) du dz$ ,
- (iii) an  $(\mathcal{F}_t)_{t \geq 0}$ -adapted, càdlàg process  $(R, V)$  in  $\mathbb{R}^{2d}$  satisfying (17) a.s.
- (iv) a measurable process  $(r_t(\xi), w_t(\xi))$  defined on  $([0, 1], \mathcal{B}([0, 1]), d\xi)$  such that  $(r_t, w_t)$  has the same law on  $([0, 1], \mathcal{B}([0, 1]), d\xi)$  as  $(R(t), V(t))$  on  $(\Omega, \mathcal{F}, \mathbb{P})$ , for each  $t \geq 0$ .

Let  $(R, V)$  be a weak solution to the mean-field SDE. Applying the Itô formula one finds that its law is a solution to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ . Conversely, each solution  $\mu$  to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$  can be represented as a weak solution to the mean-field SDE (17). Now we give a precise formulation of these statements in the following lemma.

**Lemma 2.4.** *The following assertions hold.*

(a) *Let  $(R, V)$  be a weak solution to the mean-field SDE (17) satisfying*

$$\sup_{s \in [0, t]} \mathbb{E}(|V(s)|^\gamma) < \infty, \quad \forall t > 0. \quad (18)$$

*Then the law of  $(R, V)$  on the Skorokhod space  $D(\mathbb{R}_+; \mathbb{R}^{2d})$  solves the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ .*

(b) *Let  $\mu$  be a solution to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ . Then there exists a weak solution  $(R, V)$  to the mean-field SDE (17) such that  $(R, V)$  has law  $\mu$ .*

The proof of this statement is mainly based on the equivalence between weak solutions to stochastic equations and solutions to martingale problems, see [18, 20]. Some additional arguments are given in the appendix.

**3. Statement of results.** Here and below we suppose that conditions (A) - (C) are satisfied. Set  $\langle u \rangle := (1 + |u|^2)^{\frac{1}{2}}$ ,  $u \in \mathbb{R}^d$ . Then  $1 \leq \langle u \rangle \leq 1 + |u|$  and this function satisfies the elementary inequalities

$$\langle u + w \rangle \leq \sqrt{2} \min\{\langle u \rangle + \langle w \rangle, \langle u \rangle \langle w \rangle\}, \quad (19)$$

where  $u, w \in \mathbb{R}^d$ . Moreover, it is easy to see that

$$\langle \eta u + (1 - \eta)w \rangle \leq \eta \langle u \rangle + (1 - \eta) \langle w \rangle. \quad (20)$$

Both estimates will be frequently used throughout this work.

**3.1. Construction of particle dynamics.** Denote by  $\mathcal{P}(\mathbb{R}^{2dN})$  the space of all Borel probability measures over  $\mathbb{R}^{2dN}$ . Below we provide sufficient conditions such that the martingale problem with generator  $L$  and domain  $C_c^1(\mathbb{R}^{2dN})$  is well-posed, i.e. for each  $\rho \in \mathcal{P}(\mathbb{R}^{2dN})$  there exists a unique probability measure  $\mathbb{P}_\rho$  over the Skorokhod space  $D(\mathbb{R}_+; \mathbb{R}^{2dN})$  such that  $\mathbb{P}_\rho((\mathbf{r}(0), \mathbf{v}(0)) \in E) = \rho(E)$  for all Borel sets  $E \subset \mathbb{R}^{2dN}$  and

$$F(\mathbf{r}(t), \mathbf{v}(t)) - F(\mathbf{r}(0), \mathbf{v}(0)) - \int_0^t LF(\mathbf{r}(s), \mathbf{v}(s)) ds$$

is a  $\mathbb{P}_\rho$ -martingale for each  $F \in C_c^1(\mathbb{R}^{2dN})$ . Here  $(\mathbf{r}(t), \mathbf{v}(t))$  denotes the canonical coordinate process in the Skorokhod space  $D(\mathbb{R}_+; \mathbb{R}^{2dN})$ . We start with the simpler case where condition (B) holds with  $\gamma = 0$ .

**Theorem 3.1.** *Suppose that  $\gamma = 0$  and  $\eta \in (0, 1]$ . Then for each  $N \geq 2$  and each  $\rho \in \mathcal{P}(\mathbb{R}^{2dN})$  the martingale problem  $(L, C_c^1(\mathbb{R}^{2dN}), \rho)$  has a unique solution.*

The proof of this statement is a direct consequence of the classical perturbation theory for martingale problems (see [8, Section 10]) and the observation that  $L$  can be decomposed into a linear dissipative generator (the first order term) and a bounded pure jump generator.

If  $\eta = 1$  in (6), then we may also consider  $\gamma \in (0, 2]$  resulting in the following theorem.

**Theorem 3.2.** *Suppose that (8) holds for  $p := 2$ . Let  $\gamma \in (0, 2]$  and  $\eta = 1$ . Then for each  $N \geq 2$  and each  $\rho \in \mathcal{P}(\mathbb{R}^{2dN})$  satisfying*

$$\int_{\mathbb{R}^{2dN}} \sum_{j=1}^N |v_j|^4 d\rho(r, v) < \infty \quad (21)$$

*the martingale problem  $(L, C_c^1(\mathbb{R}^{2dN}), \rho)$  has a unique solution.*

Since in such a case  $LF$  is not a bounded function, even if  $F \in C_c^1(\mathbb{R}^{2dN})$ , the desired result does not immediately follow from the classical theory of martingale problems. The proof of Theorem 3.2 is given in Section 3 and is based on an additional approximation argument combined with moment inequalities so that we may apply Theorem D.2 from the appendix.

The study of singular rates  $\psi(r) \approx r^{-\alpha}$  with  $\alpha > 0$  is an interesting mathematical problem. Since such a model would create singular transition rates, it seems natural to study in such a case solutions possessing additional regularity, see e.g. [19, 26].

**3.2. Mean-field limit  $N \rightarrow \infty$  and propagation of chaos.** For each  $N \geq 2$ , let  $(R_k^N, V_k^N)_{k=1, \dots, N}$  be the Markov process with the phase space  $\mathbb{R}^{2dN}$  and the generator  $L$ , see Theorem 3.1 and Theorem 3.2. Below we study the mean-field limit  $N \rightarrow \infty$  for the sequence of empirical measures

$$\mu^{(N)} = \frac{1}{N} \sum_{k=1}^N \delta_{(R_k^N, V_k^N)}.$$

Since the moment estimates we derive in Section 4 are different for the cases  $\gamma = 0$  and  $\gamma \in (0, 2]$ , we consider these cases separately. Below we start with the simpler case  $\gamma = 0$ .

**Theorem 3.3.** *Let  $\gamma = 0$  and  $\eta \in (0, 1]$ . Suppose that (8) holds for some  $2p \geq 1$  and  $\mu_0 \in \mathcal{P}(\mathbb{R}^{2d})$  satisfies*

$$\int_{\mathbb{R}^{2d}} (|r| + |v|^{2p}) \mu_0(dr, dv) < \infty.$$

*Then, there exists a unique weak solution  $(R, V)$  to the mean-field SDE (17). Let  $\mu$  be the law of  $(R, V)$ . Then*

$$\frac{1}{N} \sum_{j=1}^N \delta_{(R_j^N, V_j^N)} \longrightarrow \mu, \quad N \rightarrow \infty \quad (22)$$

*in law on the space of probability measures over the Skorokhod space  $D(\mathbb{R}_+; \mathbb{R}^{2d})$ . Moreover, there exists a constant  $C > 0$  such that*

$$\mathbb{E} \left( \sup_{s \in [0, t]} |V(s)|^{2p} \right) \leq \int_{\mathbb{R}^{2d}} \langle v \rangle^{2p} \mu_0(dr, dv) e^{Ct}, \quad t \geq 0. \quad (23)$$

Proving this result, we also show that the kinetic equation (9) is well-posed. Hence (22) is equivalent to propagation of chaos (see [28]). Let us also mention [22, 19, 26] for other related recent developments on propagation of chaos.

In the case  $\gamma = 0$  we may also prove that the obtained solution propagates exponential moments.

**Corollary 1.** *Suppose that  $\gamma = 0$  and  $\eta \in (0, 1]$ .*

(a) *Given  $\eta = 1$ , suppose that there exist  $\delta > 0$  and  $\kappa \in (0, 1]$  such that*

$$c(\delta, \kappa) = \int_{\mathbb{R}^d} e^{\delta|u|^\kappa} a(u) du < \infty. \quad (24)$$

*Let  $\mu_0 \in \mathcal{P}(\mathbb{R}^{2d})$  be such that*

$$\int_{\mathbb{R}^{2d}} (|r| + e^{\delta|v|^\kappa}) \mu_0(dr, dv) < \infty.$$

*Then the unique weak solution  $(R, V)$  to the mean-field SDE (17) satisfies*

$$\mathbb{E} \left( \sup_{s \in [0, t]} e^{\delta|V(s)|^\kappa} \right) \leq \int_{\mathbb{R}^{2d}} e^{\delta|v|^\kappa} d\mu_0(r, v) e^{Ct}, \quad t \geq 0$$

*for some constant  $C > 0$ .*

(b) *If  $\eta \in (0, 1)$ , then assertion (a) still holds, provided  $\kappa = 1$ .*

For the case  $\gamma \in (0, 2]$ , we have to restrict our study to  $\eta = 1$ . The main reason for this are the moment estimates derived in Section 4. Moreover, since it is not known if uniqueness holds for the kinetic equation (9) when  $\gamma \neq 0$ , we can only conclude that the sequence of empirical measures is tight and that each limit provides a weak solution to the mean-field SDE (17).

**Theorem 3.4.** *Given  $\gamma \in [0, 2]$  and  $\eta = 1$  assume that (8) holds for some  $2p \geq \max\{4, 1 + 2\gamma\}$ . Let  $\mu_0 \in \mathcal{P}(\mathbb{R}^{2d})$  satisfy*

$$\int_{\mathbb{R}^{2d}} (|r| + |v|^{2p}) \mu_0(dr, dv) < \infty.$$

*Then there exists a weak solution  $(R, V)$  to the mean-field SDE (17). Moreover, there exists a constant  $C = C(\psi, \sigma, a, p) > 0$  such that*

$$\mathbb{E}(\langle V(t) \rangle^{2p}) \leq \begin{cases} C\mathbb{E}(\langle V(0) \rangle^{2p})(1 + t^{\frac{2p}{2-\gamma}}), & \gamma \neq 2 \\ \mathbb{E}(\langle V(0) \rangle^{2p})e^{Ct}, & \gamma = 2 \end{cases}, \quad t \geq 0, \quad (25)$$

*and, there exists another constant  $C' = C'(\psi, \sigma, a) > 0$  such that for  $t \geq 0$*

$$\mathbb{E} \left( \sup_{s \in [0, t]} \langle V(s) \rangle^{2p-\gamma} \right) \leq \mathbb{E}(\langle V(0) \rangle^{2p-\gamma}) + C'2^{2p} \int_0^t \mathbb{E}(\langle V(s) \rangle^{2p}) ds. \quad (26)$$

**4. The particle dynamics.** Fix  $N \geq 2$  and let  $\gamma \in [0, 2]$  be given as in condition (B). In this section we first establish some moment inequalities, then provide a pathwise description for the  $N$ -particle process with the generator  $L$ , and finally prove some moment estimates for the particle process with constants independent of  $N$  and  $\eta \in (0, 1]$ .

**4.1. Moment inequalities for the generator.** We start with a moment inequality where corresponding constants are independent of  $\eta \in (0, 1]$  and  $N \geq 2$ .

**Lemma 4.1.** *Suppose that (8) holds for  $p \geq \frac{1}{2}$ . Then for each  $k \in \{1, \dots, N\}$*

$$\begin{aligned} & \frac{1}{N} \sum_{j=1}^N \psi(r_k - r_j) \sigma(v_k - v_j) \int_{\mathbb{R}^d} |\langle \eta v_j + (1 - \eta)v_k + \eta u \rangle^{2p} - \langle v_k \rangle^{2p}| a(u) du \\ & \leq 3\lambda_{2p} 2^{p+4} \|\psi\|_{\infty} c_{\sigma} \left( \langle v_k \rangle^{2p+\gamma} + \frac{1}{N} \sum_{j=1}^N \langle v_j \rangle^{2p+\gamma} \right). \end{aligned}$$

*Proof.* Using (19), (20) and taking into account  $\eta \in (0, 1]$  we get

$$\begin{aligned} \langle \eta v_j + (1 - \eta)v_k + \eta u \rangle^{2p} + \langle v_k \rangle^{2p} & \leq 2^p \langle \eta v_j + (1 - \eta)v_k \rangle^{2p} \langle u \rangle^{2p} + \langle v_k \rangle^{2p} \langle u \rangle^{2p} \\ & \leq 2^p \langle u \rangle^{2p} (\langle v_j \rangle^{2p} + \langle v_k \rangle^{2p}) + \langle u \rangle^{2p} \langle v_k \rangle^{2p} \\ & \leq 2^{p+1} \langle u \rangle^{2p} (\langle v_j \rangle^{2p} + \langle v_k \rangle^{2p}). \end{aligned}$$

Using  $\sigma(v_k - v_j) \leq c_\sigma \langle v_k - v_j \rangle^\gamma \leq c_\sigma 2^{\frac{3}{2}\gamma} (\langle v_k \rangle^\gamma + \langle v_j \rangle^\gamma)$  we obtain

$$\begin{aligned}
& \frac{1}{N} \sum_{j=1}^N \psi(r_k - r_j) \sigma(v_k - v_j) \int_{\mathbb{R}^d} |\langle \eta v_j + (1 - \eta)v_k + \eta u \rangle^{2p} - \langle v_k \rangle^{2p}| a(u) du \\
& \leq \lambda_{2p} 2^{p+1+\frac{3}{2}\gamma} \frac{\|\psi\|_\infty c_\sigma}{N} \sum_{j=1}^N (\langle v_k \rangle^\gamma + \langle v_j \rangle^\gamma) (\langle v_j \rangle^{2p} + \langle v_k \rangle^{2p}) \\
& = \lambda_{2p} 2^{p+1+\frac{3}{2}\gamma} \frac{\|\psi\|_\infty c_\sigma}{N} \sum_{j=1}^N (\langle v_k \rangle^\gamma \langle v_j \rangle^{2p} + \langle v_k \rangle^{\gamma+2p} + \langle v_j \rangle^{2p+\gamma} + \langle v_j \rangle^\gamma \langle v_k \rangle^{2p}) \\
& \leq 3\lambda_{2p} 2^{p+1+\frac{3}{2}\gamma} \|\psi\|_\infty c_\sigma \left( \langle v_k \rangle^{2p+\gamma} + \frac{1}{N} \sum_{j=1}^N \langle v_j \rangle^{2p+\gamma} \right),
\end{aligned}$$

where we have used Young's inequality

$$\langle v_j \rangle^{2p} \langle v_k \rangle^\gamma \leq \frac{2p}{2p+\gamma} \langle v_j \rangle^{2p+\gamma} + \frac{\gamma}{2p+\gamma} \langle v_k \rangle^{2p+\gamma}. \quad (27)$$

□

Next we investigate moment inequalities in the case  $\eta = 1$ .

**Lemma 4.2.** *Suppose that (8) holds for some  $p \geq 2$ . Then*

$$\begin{aligned}
& \frac{1}{N^2} \sum_{k,j=1}^N \psi(r_k - r_j) \sigma(v_k - v_j) \int_{\mathbb{R}^d} (|v_j + u|^{2p} - |v_k|^{2p}) a(u) du \\
& \leq 3\lambda_{2p} 2^{3p+5} \frac{\|\psi\|_\infty c_\sigma}{N} \sum_{j=1}^N \langle v_j \rangle^{2p-2+\gamma}.
\end{aligned}$$

*Proof.* By the mean-value Theorem we get

$$\begin{aligned}
|v_j + u|^{2p} &= (|v_j|^2 + |u|^2 + 2(v_j \cdot u))^p \\
&= (|v_j|^2 + |u|^2)^p + 2p(|v_j|^2 + |u|^2)^{p-1}(v_j \cdot u) \\
&\quad + 4p(p-1)(v_j \cdot u)^2 \int_0^1 (1-t) (|v_j|^2 + |u|^2 + 2t(v_j \cdot u))^{p-2} dt
\end{aligned}$$

For the last integral we get by  $2|v_j||u| \leq |v_j|^2 + |u|^2$  and  $(a+b)^q \leq 2^q(a^q + b^q)$  for  $q \geq 0$  and  $a, b \geq 0$

$$\begin{aligned}
& \left| 4p(p-1)(v_j \cdot u)^2 \int_0^1 (1-t) (|v_j|^2 + |u|^2 + 2t(v_j \cdot u))^{p-2} dt \right| \\
& \leq 4p(p-1)(|v_j|^2 + |u|^2 + 2|v_j||u|)^{p-2} |v_j|^2 |u|^2 \\
& \leq p(p-1)2^p (|v_j|^2 + |u|^2)^{p-2} |v_j|^2 |u|^2 \\
& \leq p(p-1)2^{2p-2} (|v_j|^{2p-2} |u|^2 + |v_j|^2 |u|^{2p-2}) \\
& \leq p(p-1)2^{2p-1} \langle u \rangle^{2p} \langle v_j \rangle^{2p-2}.
\end{aligned}$$

Let  $k_p = \lfloor \frac{p+1}{2} \rfloor$  where  $\lfloor x \rfloor \in \mathbb{Z}$  is defined by  $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$ , set  $\binom{p}{l} = \frac{p(p-1)\cdots(p-l+1)}{l!}$ , for  $l \geq 1$ , and  $\binom{p}{0} = 1$ . Then we obtain by the fractional binomial expansion Lemma B.1 (see [3, Lemma 2] for a proof)

$$\begin{aligned} (|v_j|^2 + |u|^2)^p &\leq |u|^{2p} + |v_j|^{2p} + \sum_{l=1}^{k_p} \binom{p}{l} (|v_j|^{2l}|u|^{2p-2l} + |v_j|^{2p-2l}|u|^{2l}) \\ &\leq \langle u \rangle^{2p} + |v_j|^{2p} + \langle u \rangle^{2p} \sum_{l=1}^{k_p} \binom{p}{l} (\langle v_j \rangle^{2k_p} + \langle v_j \rangle^{2p-2}) \\ &\leq \langle u \rangle^{2p} + |v_j|^{2p} + 2^{p+1} \langle u \rangle^{2p} \langle v_j \rangle^{2p-2} \end{aligned}$$

where we have used  $k_p \leq p-1$  and  $\sum_{l=1}^{k_p} \binom{p}{l} \leq 2^p$ . Using the symmetry of  $a$  we have  $\int_{\mathbb{R}^d} (v_j \cdot u) a(u) du = 0$  and hence obtain

$$\begin{aligned} &\int_{\mathbb{R}^d} (|v_j + u|^{2p} - |v_k|^{2p}) a(u) du \\ &\leq \int_{\mathbb{R}^d} ((|v_j|^2 + |u|^2)^p - |v_k|^{2p}) a(u) du + p(p-1)2^{2p-1} \lambda_{2p} \langle v_j \rangle^{2p-2} \\ &\leq |v_j|^{2p} - |v_k|^{2p} + \lambda_{2p} + \lambda_{2p} 2^{p+1} \langle v_j \rangle^{2p-2} + p(p-1)2^{2p-1} \lambda_{2p} \langle v_j \rangle^{2p-2} \\ &\leq |v_j|^{2p} - |v_k|^{2p} + \lambda_{2p} 2^{3p+2} \langle v_j \rangle^{2p-2} \end{aligned}$$

where we have used  $p(p-1) \leq 2^p$  so that  $1 + 2^{p+1} + p(p-1)2^{p-1} \leq 2^{3p+2}$ . By symmetry we obtain

$$\sum_{k,j=1}^N \psi(r_k - r_j) \sigma(v_k - v_j) (|v_j|^{2p} - |v_k|^{2p}) = 0 \quad (28)$$

and hence

$$\begin{aligned} &\frac{1}{N^2} \sum_{k,j=1}^N \psi(r_k - r_j) \sigma(v_k - v_j) \int_{\mathbb{R}^d} (|v_j + u|^{2p} - |v_k|^{2p}) a(u) du \\ &\leq \lambda_{2p} 2^{3p+2+\frac{3}{2}\gamma} \frac{\|\psi\|_{\infty} c_{\sigma}}{N^2} \sum_{k,j=1}^N (\langle v_k \rangle^{\gamma} + \langle v_j \rangle^{\gamma}) \langle v_j \rangle^{2p-2} \\ &= \lambda_{2p} 2^{3p+2+\frac{3}{2}\gamma} \frac{\|\psi\|_{\infty} c_{\sigma}}{N^2} \sum_{k,j=1}^N (\langle v_k \rangle^{2p-2+\gamma} + 2 \langle v_j \rangle^{2p-2+\gamma}) \\ &= 3 \lambda_{2p} 2^{3p+2+\frac{3}{2}\gamma} \frac{\|\psi\|_{\infty} c_{\sigma}}{N} \sum_{j=1}^N \langle v_j \rangle^{2p-2+\gamma}. \end{aligned}$$

where we have used Young's inequality

$$\begin{aligned} \langle v_k \rangle^{\gamma} \langle v_j \rangle^{2p-2} &\leq \frac{\gamma}{2p-2+\gamma} \langle v_k \rangle^{2p-2+\gamma} + \frac{2p-2}{2p-2+\gamma} \langle v_j \rangle^{2p-2+\gamma} \\ &\leq \langle v_k \rangle^{2p-2+\gamma} + \langle v_j \rangle^{2p-2+\gamma}. \end{aligned}$$

The assertion is proved.  $\square$

Finally we give an estimate on the exponential moments, provided that  $\gamma = 0$ . In the particular case  $\eta = 1$  we obtain the following.

**Lemma 4.3.** *Assume that  $\gamma = 0$ ,  $\eta = 1$  and suppose that there exist  $\delta > 0$  and  $\kappa \in (0, 1]$  satisfying (24). Then*

$$\frac{1}{N^2} \sum_{k,j=1}^N \sigma(v_k - v_j) \int_{\mathbb{R}^d} \left| e^{\delta\langle v_j+u \rangle^\kappa} - e^{\delta\langle v_k \rangle^\kappa} \right| a(u) du \leq \frac{4c_\sigma e^\delta c(\delta, \kappa)}{N} \sum_{j=1}^N e^{\delta\langle v_j \rangle^\kappa}.$$

*Proof.* Using the inequalities  $\langle v_j + u \rangle \leq 1 + |v_j| + |u|$  and  $(1+a)^\kappa \leq 1 + a^\kappa$ ,  $a \geq 0$ , we obtain

$$\begin{aligned} \left| e^{\delta\langle v_j+u \rangle^\kappa} - e^{\delta\langle v_k \rangle^\kappa} \right| &\leq e^{\delta\langle v_j+u \rangle^\kappa} + e^{\delta\langle v_k \rangle^\kappa} \\ &\leq e^\delta e^{\delta|v_j|^\kappa} e^{\delta|u|^\kappa} + e^{\delta\langle v_k \rangle^\kappa} \\ &\leq e^\delta e^{\delta|u|^\kappa} \left( e^{\delta\langle v_j \rangle^\kappa} + e^{\delta\langle v_k \rangle^\kappa} \right) \end{aligned}$$

and hence

$$\begin{aligned} &\frac{1}{N^2} \sum_{k,j=1}^N \sigma(v_k - v_j) \int_{\mathbb{R}^d} \left| e^{\delta\langle v_j+u \rangle^\kappa} - e^{\delta\langle v_k \rangle^\kappa} \right| a(u) du \\ &\leq \frac{e^\delta c(\delta, \kappa)}{N^2} \sum_{k,j=1}^N \sigma(v_k - v_j) \left( e^{\delta\langle v_j \rangle^\kappa} + e^{\delta\langle v_k \rangle^\kappa} \right) \\ &\leq \frac{2c_\sigma e^\delta c(\delta, \kappa)}{N} \sum_{j=1}^N e^{\delta\langle v_j \rangle^\kappa}. \end{aligned}$$

□

For the case where  $\eta \in (0, 1)$  we have the following.

**Lemma 4.4.** *Assume that  $\gamma = 0$ ,  $\eta \in (0, 1)$ , and suppose that there exist  $\delta > 0$  such that (24) holds for  $\kappa = 1$ . Then*

$$\begin{aligned} &\frac{1}{N^2} \sum_{k,j=1}^N \sigma(v_k - v_j) \int_{\mathbb{R}^d} \left| e^{\delta\langle \eta v_j + (1-\eta)v_k + \eta u \rangle} - e^{\delta\langle v_k \rangle} \right| a(u) du \\ &\leq \frac{2c_\sigma e^\delta c(\delta, 1)}{N} \sum_{j=1}^N e^{\delta\langle v_j \rangle}. \end{aligned}$$

*Proof.* Using the inequality  $\langle \eta v_j + (1-\eta)v_k + \eta u \rangle \leq 1 + \eta|v_j| + (1-\eta)|v_k| + |u|$  we obtain

$$\begin{aligned} \left| e^{\delta\langle \eta v_j + (1-\eta)v_k + \eta u \rangle} - e^{\delta\langle v_k \rangle} \right| &\leq e^{\delta\langle \eta v_j + (1-\eta)v_k + \eta u \rangle} + e^{\delta\langle v_k \rangle} \\ &\leq e^\delta e^{\delta\eta|v_j|} e^{\delta(1-\eta)|v_k|} e^{\delta|u|} + e^{\delta\langle v_k \rangle} \\ &\leq 2e^\delta e^{\delta|u|} \left( e^{\delta\langle v_j \rangle} + e^{\delta\langle v_k \rangle} \right) \end{aligned}$$

where we have used Young's inequality to obtain

$$e^{\delta\eta|v_j|} e^{\delta(1-\eta)|v_k|} \leq \eta e^{\delta|v_j|} + (1-\eta) e^{\delta|v_k|} \leq e^{\delta\langle v_j \rangle} + e^{\delta\langle v_k \rangle}.$$

This implies the assertion. □

**4.2. Pathwise description of dynamics.** In order to provide a construction of the Markov process and, in particular, to study the mean-field limit  $N \rightarrow \infty$ , it is useful to give a pathwise description of the Markov process associated to  $L$  in terms of stochastic differential equations. Namely, a weak solution  $(\mathbf{R}, \mathbf{V}) = (R_1, \dots, R_N, V_1, \dots, V_N)$  to the system of stochastic equations

$$\begin{cases} \mathbf{R}(t) = \mathbf{R}(0) + \int_0^t \mathbf{V}(s) ds, \\ \mathbf{V}(t) = \mathbf{V}(0) + \int_0^t \int_{\{1, \dots, N\}^2 \times \mathbb{R}^d \times \mathbb{R}_+} \mathbf{G}(\mathbf{R}(s), \mathbf{V}(s-), u, l, l', z) \mathcal{N}(ds, dl, dl', du, dz), \end{cases} \quad (29)$$

where  $\mathbf{e}_l$  is the same as in the definition of (7) and

$$\mathbf{G}(\mathbf{R}, \mathbf{V}, u, l, l', z) = \mathbf{e}_l \eta(u + V_{l'} - V_l) \mathbb{1}_{[0, \psi(R_l(s) - R_{l'}(s)) \sigma(V_l(s-) - V_{l'}(s-))]}(z), \quad (30)$$

consists of a stochastic basis  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  with the usual conditions, a  $(\mathcal{F}_t)_{t \geq 0}$ -adapted càdlàg process  $(\mathbf{R}, \mathbf{V})$  in  $\mathbb{R}^{2dN}$  and a  $(\mathcal{F}_t)_{t \geq 0}$ -adapted Poisson random measure  $\mathcal{N}$  on  $\mathbb{R}_+ \times \{1, \dots, N\}^2 \times \mathbb{R}^d \times \mathbb{R}_+$  with the compensator

$$\tilde{\mathcal{N}}(ds, dl, dl', du, dz) = ds \otimes \left( \frac{1}{N} \sum_{j, k=1}^N \delta_j(dl) \otimes \delta_k(dl') \right) \otimes (a(u) du) \otimes dz. \quad (31)$$

Let us heuristically explain equation (29) in more detail. In order to describe a jump at time  $s$  we have to chose uniformly two particles  $(l, l') \in \{1, \dots, N\}^2$ , a random parameter  $u \in \mathbb{R}^d$  distributed according to  $a(u) du$  describing the deviation from the deterministic model, and finally another auxilliary parameter  $z \in \mathbb{R}_+$  whose intensity is  $dz$ . All this parameters are chosen with intensity (31). If at the jump time  $s$  we have  $\psi(R_l(s) - R_{l'}(s)) \sigma(V_l(s-) - V_{l'}(s-)) \leq z$ , then no jump occurs. Suppose now that  $\psi(R_l(s) - R_{l'}(s)) \sigma(V_l(s-) - V_{l'}(s-)) > z$ . Then the velocity of the particle system changes according to  $\mathbf{V}(s) \mapsto \mathbf{V}(s) + \mathbf{G}(\mathbf{R}(s), \mathbf{V}(s-), u, l, l', z)$ . The particular form of (30) shows that in this case the transition results only in the change of velocity of the particle  $l$  according to

$$V_l(s-) \mapsto v^*(V_l(s-), V_{l'}(s-), u) = V_l(s-) + \eta(V_{l'}(s-) - V_l(s-)) + \eta u.$$

The next lemma is a particular case of the Itô formula and shows that  $L$  is indeed the generator of the process obtained from (29).

**Lemma 4.5.** *Let  $(\mathbf{R}, \mathbf{V})$  be a weak solution to (29). Then for each  $F \in C^1(\mathbb{R}^{2dN})$  satisfying*

$$\sup_{|\mathbf{r}| + |\mathbf{v}| \leq M} \int_{\mathbb{R}^d} |F(\mathbf{r}, \mathbf{v} + \mathbf{e}_k(v^*(v_k, v_j, u) - v_k)) - F(\mathbf{r}, \mathbf{v})| a(u) du < \infty, \quad \forall M > 0 \quad (32)$$

the formula

$$F(\mathbf{R}(t), \mathbf{V}(t)) = F(\mathbf{R}(0), \mathbf{V}(0)) + \int_0^t LF(\mathbf{R}(s), \mathbf{V}(s)) ds + M_F(t)$$

holds. Here  $LF$  is defined as in (7) and  $(M_F(t))_{t \geq 0}$  is a local martingale given by

$$M_F(t) = \int_0^t \int_{\{1, \dots, N\}^2 \times \mathbb{R}^d \times \mathbb{R}_+} \Delta F \tilde{\mathcal{N}}(ds, dl, dl', du, dz)$$

with  $\tilde{\mathcal{N}} = \mathcal{N} - \hat{\mathcal{N}}$  and

$$\Delta F = F(\mathbf{R}(s), \mathbf{V}(s-) + \mathbf{G}(\mathbf{R}(s), \mathbf{V}(s-), u, l, l', z)) - F(\mathbf{R}(s), \mathbf{V}(s-)).$$

*Proof.* Applying the Itô formula to  $F(\mathbf{R}(t), \mathbf{V}(t))$  yields

$$\begin{aligned} F(\mathbf{R}(t), \mathbf{V}(t)) &= F(\mathbf{R}(0), \mathbf{V}(0)) + \int_0^t \mathbf{V}(s) \cdot (\nabla_{\mathbf{R}} F)(\mathbf{R}(s), \mathbf{V}(s)) ds \\ &\quad + \int_0^t \int_{\{1, \dots, N\}^2 \times \mathbb{R}^d \times \mathbb{R}_+} \Delta F \widehat{\mathcal{N}}(ds, dl, dl', du, dz) \\ &\quad + \int_0^t \int_{\{1, \dots, N\}^2 \times \mathbb{R}^d \times \mathbb{R}_+} \Delta F \widetilde{\mathcal{N}}(ds, dl, dl', du, dz). \end{aligned}$$

By direct computation, one shows that

$$\begin{aligned} &\int_0^t \mathbf{V}(s) \cdot (\nabla_{\mathbf{R}} F)(\mathbf{R}(s), \mathbf{V}(s)) ds + \int_0^t \int_{\{1, \dots, N\}^2 \times \mathbb{R}^d \times \mathbb{R}_+} \Delta F \widehat{\mathcal{N}}(ds, dl, dl', du, dz) \\ &= \int_0^t LF(\mathbf{R}(s), \mathbf{V}(s)) ds, \end{aligned}$$

while in view of (32) one has

$$\begin{aligned} &\int_0^t \int_{\{1, \dots, N\}^2 \times \mathbb{R}^d \times \mathbb{R}_+} |\Delta F| \widehat{\mathcal{N}}(ds, dl, dl', du, dz) \\ &= \frac{1}{N} \sum_{k, j=1}^N \int_0^t \psi(R_k(s) - R_j(s)) \sigma(V_k(s) - V_j(s)) \\ &\quad \cdot \int_{\mathbb{R}^d} |F(\mathbf{R}(s), \mathbf{V}(s) + \mathbf{e}_k(v^*(V_k(s), V_j(s), u) - V_k(s))) - F(\mathbf{R}(s), \mathbf{V}(s))| a(u) du ds \end{aligned}$$

where the latter expression is a.s. finite. Hence all integrals above are a.s. finite. The assertion is proved.  $\square$

The following is a standard result in the theory of martingale problems and stochastic equations, see [18, Theorem A.1] and the references therein.

**Proposition 1.** *Let  $\rho \in \mathcal{P}(\mathbb{R}^{2dN})$  and let  $\mathbb{P}_\rho$  be a solution to the martingale problem  $(L, C_c^1(\mathbb{R}^{2dN}), \rho)$ , i.e. for any  $F \in C_c^1(\mathbb{R}^{2dN})$ ,*

$$F(\mathbf{r}(t), \mathbf{v}(t)) - F(\mathbf{r}(0), \mathbf{v}(0)) - \int_0^t (LF)(\mathbf{r}(s), \mathbf{v}(s)) ds, \quad t \geq 0 \quad (33)$$

*is a martingale with respect to  $\mathbb{P}_\rho$ , where  $(\mathbf{r}(t), \mathbf{v}(t))$  the coordinate process in the Skorokhod space  $D(\mathbb{R}_+; \mathbb{R}^{2dN})$  and  $\mathbb{P}_\rho$  is a law on  $D(\mathbb{R}_+; \mathbb{R}^{2dN})$ . Then there exists a weak solution  $(\mathbf{R}, \mathbf{V})$  to (29) such that the law of  $(\mathbf{R}, \mathbf{V})$  is precisely  $\mathbb{P}_\rho$ .*

*Proof.* Using the fact that  $|LF(\mathbf{r}, \mathbf{v})| \leq C \sum_{k=1}^N \langle v_k \rangle^2$ , we conclude that (33) is a local martingale for any  $F \in C_b^1(\mathbb{R}^{2dN})$ . The assertion is now a consequence of the equivalence between martingale problems and weak solutions to stochastic equations, see, e.g., [18, Theorem A.1].  $\square$

**4.3. Proof of Theorem 3.2.** In this part we give a full proof of Theorem 3.2.

*Proof of Theorem 3.2.* Let us show that we can apply Theorem D.2 to deduce the assertion. Take  $g \in C^\infty(\mathbb{R}_+)$  such that  $\mathbb{1}_{[0,1]} \leq g \leq \mathbb{1}_{[0,2]}$  and set

$$g_m(\mathbf{v}) = g\left(\frac{\sum_{k=1}^N |v_k|^2}{m^2}\right), \quad \mathbf{v} = (v_1, \dots, v_N) \in \mathbb{R}^{dN}.$$

Let  $L_m$  be the Markov operator given by  $L$  with  $\sigma(v_k - v_j)$  replaced by  $g_m(\mathbf{v})\sigma(v_k - v_j)$ . Then for each  $F \in C_c^1(\mathbb{R}^{2dN})$  we can find a constant  $C = C(F, \psi, \sigma) > 0$  (independent of  $m$ ) such that

$$|L_m F(\mathbf{r}, \mathbf{v})|, |LF(\mathbf{r}, \mathbf{v})| \leq C \sum_{j=1}^N \langle v_j \rangle^\gamma. \quad (34)$$

*Step 1.* Let  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  be a stochastic basis and let  $(\mathbf{R}(0), \mathbf{V}(0)) \in \mathbb{R}^{2dN}$  be a random variable with some given law  $\mu \in \mathcal{P}(\mathbb{R}^{2dN})$ . Let  $\mathcal{N}_m$  be a Poisson random measure on  $\Omega$  with compensator

$$\widehat{\mathcal{N}}_m(ds, dl, dl', du, dz) = ds \otimes \left( \frac{1}{N} \sum_{j,k=1}^N \delta_j(dl) \otimes \delta_k(dl') \right) \otimes (a(u)du) \otimes dz$$

on  $\mathbb{R}_+ \times \{1, \dots, N\}^2 \times \mathbb{R}^d \times [0, c_m]$  (for some constant  $c_m > 0$  large enough such that  $\psi(r_l - r_{l'})\sigma(v_l - v_{l'})g_m(\mathbf{v}) \leq c_m$  for all  $(\mathbf{r}, \mathbf{v}) \in \mathbb{R}^{2dN}$ ). Consider the system of stochastic equations

$$\begin{cases} \mathbf{R}^m(t) &= \mathbf{R}(0) + \int_0^t \mathbf{V}^m(s) ds, \\ \mathbf{V}^m(t) &= \mathbf{V}(0) + \int_0^t \int_{\{1, \dots, N\}^2 \times \mathbb{R}^d \times [0, c_m]} \mathbf{G}^m(\mathbf{R}^m(s), \mathbf{V}^m(s-), u, l, l', z) d\mathcal{N}_m, \end{cases} \quad (35)$$

with  $d\mathcal{N}_m = \mathcal{N}_m(ds, dl, dl', du, dz)$ ,  $\mathbf{G}^m = \mathbf{G}^m(\mathbf{R}^m, \mathbf{V}^m, u, l, l', z)$ , and

$$\mathbf{G}^m = \mathbf{e}_l \eta(u + V_l^m - V_{l'}^m) \mathbb{1}_{[0, \psi(R_l^m(s) - R_{l'}^m(s))g_m(\mathbf{V}^m(s-))\sigma(V_l^m(s-) - V_{l'}^m(s-))]}(z).$$

Since  $\widehat{\mathcal{N}}_m((0, t] \times \{1, \dots, N\}^2 \times \mathbb{R}^d \times [0, c_m]) < \infty$ , for all  $t > 0$ , it follows that (35) can be uniquely solved from jump to jump. Since  $L_m F$  is bounded for each  $m$  and each  $F \in C_c^1(\mathbb{R}^{2dN})$  we conclude from [21] that the martingale problem  $(L_m, C_c^1(\mathbb{R}^{2dN}), \mu)$  has for each  $\mu \in \mathcal{P}(\mathbb{R}^{2dN})$  a unique solution whose law can be obtained from (35).

*Step 2.* Suppose that  $(\mathbf{R}(0), \mathbf{V}(0))$  has law  $\rho$  satisfying (21). Define  $Q_4(\mathbf{v}) = \frac{1}{N} \sum_{k=1}^N \langle v_k \rangle^4$  and observe that  $Q_4$  satisfies (32). Below we show that there exists a constant  $C > 0$  (independent of  $m$ ) such that

$$\mathbb{E}(Q_4(\mathbf{V}^m(t))) \leq \mathbb{E}(Q_4(\mathbf{V}(0))) e^{Ct}. \quad (36)$$

Indeed, it follows from Lemma 4.5

$$Q_4(\mathbf{V}^m(t)) = Q_4(\mathbf{V}(0)) + \int_0^t L_m Q_4(\mathbf{R}^m(s), \mathbf{V}^m(s)) ds + M_m(t), \quad (37)$$

where  $(M_m(t))_{t \geq 0}$  is a local martingale. Applying Lemma 4.2 gives

$$L_m Q_4(\mathbf{R}^m(s), \mathbf{V}^m(s)) \leq C \frac{1}{N} \sum_{k=1}^N \langle V_k^m(s) \rangle^{2+\gamma} \leq C Q_4(\mathbf{V}^m(s)),$$

where the constant  $C$  is independent of  $m$ . Let  $\tau_k$  be the sequence of stopping times that localizes  $(M_m(t))_{t \geq 0}$ , i.e.  $(M_m(t \wedge \tau_k))_{t \geq 0}$  is a true martingale for each  $k$  and  $\tau_k \nearrow \infty$  as  $k \rightarrow \infty$ . Evaluating (37) at  $t \wedge \tau_k$  and taking expectations gives

$$\begin{aligned} \mathbb{E}(Q_4(\mathbf{V}^m(t \wedge \tau_k))) &= \mathbb{E}(Q_4(\mathbf{V}(0))) + \mathbb{E} \left( \int_0^{t \wedge \tau_k} L_m Q_4(\mathbf{R}^m(s), \mathbf{V}^m(s)) ds \right) \\ &\leq \mathbb{E}(Q_4(\mathbf{V}(0))) + C \int_0^t \mathbb{E}(Q_4(\mathbf{V}^m(s \wedge \tau_k))) ds. \end{aligned}$$

Applying Gronwall's lemma and letting  $k \rightarrow \infty$  yields (36).

*Step 3.* Let us show that

$$\sup_{m \geq 1} \mathbb{E} \left( \sup_{s \in [0, t]} Q_2(\mathbf{V}^m(s)) \right) < \infty, \quad \forall t > 0 \quad (38)$$

holds for  $Q_2(\mathbf{v}) = \frac{1}{N} \sum_{k=1}^N \langle v_k \rangle^2$ . Indeed, using (35) gives

$$\begin{aligned} Q_2(\mathbf{V}^m(t)) &= Q_2(\mathbf{V}(0)) \\ &+ \int_0^t \int_{\{1, \dots, N\}^2 \times \mathbb{R}^d \times [0, c_m]} [Q_2(\mathbf{V}^m(s-) + \mathbf{G}^m) - Q_2(\mathbf{V}^m(s))] d\mathcal{N}_m \end{aligned}$$

and hence

$$\begin{aligned} \sup_{s \in [0, t]} Q_2(\mathbf{V}^m(s)) &\leq Q_2(\mathbf{V}(0)) \\ &+ \int_0^t \int_{\{1, \dots, N\}^2 \times \mathbb{R}^d \times [0, c_m]} |Q_2(\mathbf{V}^m(s-) + \mathbf{G}^m) - Q_2(\mathbf{V}^m(s))| d\mathcal{N}_m, \end{aligned}$$

where  $\mathbf{G}^m = \mathbf{G}^m(\mathbf{R}^m(s), \mathbf{V}^m(s-), u, l, l', z)$  and we have used the fact that the stochastic integral is defined pathwise. Taking expectations gives

$$\mathbb{E} \left( \sup_{s \in [0, t]} Q_2(\mathbf{V}^m(s)) \right) \leq \mathbb{E}(Q_2(\mathbf{V}(0))) + \int_0^t \mathbb{E}(\mathcal{H}_m(s)) ds$$

where  $\mathcal{H}_m(s)$  satisfies

$$\begin{aligned} \mathcal{H}_m(s) &= \frac{1}{N^2} \sum_{k, j=1}^N \psi(R_k^m(s) - R_j^m(s)) g_m(\mathbf{V}^m(s)) \sigma(V_k^m(s) - V_j^m(s)) \\ &\quad \cdot \int_{\mathbb{R}^d} |\langle \eta V_j^m(s) + (1 - \eta)V_k^m(s) + \eta u \rangle^2 - \langle V_k^m(s) \rangle^2| a(u) du \\ &\leq \frac{C}{N} \sum_{k=1}^N \langle V_k^m(s) \rangle^{2+\gamma} \\ &\leq C Q_4(\mathbf{V}^m(s)), \end{aligned}$$

and the constant  $C$  is independent of  $m$  and is given by Lemma 4.1. This gives

$$\mathbb{E} \left( \sup_{s \in [0, t]} Q_2(\mathbf{V}^m(s)) \right) \leq \mathbb{E}(Q_2(\mathbf{V}(0))) + C \int_0^t \mathbb{E}(Q_4(\mathbf{V}^m(s))) ds$$

and in view of (36) we deduce (38).

*Step 4.* Using (36) and (38) from previous steps combined with  $Q_2(\mathbf{v})^2 \leq Q_4(\mathbf{v})$  we may apply Theorem D.2 and conclude that the martingale problem for  $(L, C_c^1(\mathbb{R}^{2dN}), \rho)$  has a unique solution  $\mathbb{P}_\rho$  which satisfies

$$\sup_{s \in [0, t]} \mathbb{E}_\rho (Q_2(\mathbf{v}(s))^2) ds < \infty, \quad \forall t > 0,$$

where  $\mathbb{E}_\rho$  denotes the integration w.r.t.  $\mathbb{P}_\rho$  and  $(\mathbf{r}(t), \mathbf{v}(t))$  the coordinate process in the Skorokhod space  $D(\mathbb{R}_+; \mathbb{R}^{2dN})$ .  $\square$

**4.4. Moment estimates for the  $N$ -particle process.** Here and below we always assume that one of the following cases is satisfied:

- (i)  $\gamma = 0$ ,  $\eta \in (0, 1]$ , and  $\rho \in \mathcal{P}(\mathbb{R}^{2dN})$ .
- (ii)  $\gamma \in [0, 1]$ ,  $\eta = 1$ , (8) holds for  $p = 2$ , and  $\rho \in \mathcal{P}(\mathbb{R}^{2dN})$  satisfies (21).

In both cases we have seen that there exists a unique solution to (29) defined on a stochastic basis  $(\Omega^N, \mathcal{F}^N, (\mathcal{F}_t^N)_{t \geq 0}, \mathbb{P}^N)$  with the usual conditions, which we denote by  $X_k^N := (R_k^N, V_k^N)$ ,  $k = 1, \dots, N$ .

We call  $\rho \in \mathcal{P}(\mathbb{R}^{2dN})$  symmetric, if for any permutation  $\tau$  of  $\{1, \dots, N\}$  and any bounded measurable function  $F : \mathbb{R}^{2dN} \rightarrow \mathbb{R}$

$$\int_{\mathbb{R}^{2dN}} F(x_1, \dots, x_N) d\rho(x_1, \dots, x_N) = \int_{\mathbb{R}^{2dN}} F(x_{\tau(1)}, \dots, x_{\tau(N)}) d\rho(x_1, \dots, x_N).$$

The following corollary shows that the particle trajectories are indistinguishable.

**Corollary 2.** *Let  $\rho$  be symmetric, then  $X_1^N, \dots, X_N^N$  are exchangeable as elements in  $D(\mathbb{R}_+; \mathbb{R}^{2d})$ , i.e. for any permutation  $\tau$  of  $\{1, \dots, N\}$  and any bounded measurable function  $F : D(\mathbb{R}_+; \mathbb{R}^{2dN}) \rightarrow \mathbb{R}$*

$$\mathbb{E}(F(X_1^N, \dots, X_N^N)) = \mathbb{E}(F(X_{\tau(1)}^N, \dots, X_{\tau(N)}^N)). \quad (39)$$

*In particular,  $(R_k^N, V_k^N)$ ,  $k = 1, \dots, N$ , are identically distributed as elements in the Skorokhod space  $D(\mathbb{R}_+; \mathbb{R}^{2d})$ .*

*Proof.* Since  $L$  maps symmetric functions onto symmetric functions, the assertion follows from uniqueness of the martingale problem  $(L, C_c^1(\mathbb{R}^{2dN}), \rho)$ .  $\square$

Below we prove some moment estimates (uniform in  $N$  and  $\eta$ ) for the unique solution to (29). We start with the case  $\eta = 1$  and  $\gamma \in [0, 2]$ .

**Corollary 3.** *Let  $\eta = 1$ ,  $\gamma \in [0, 2]$  and suppose that (8) holds for some  $p \geq 2$ . Let  $\rho \in \mathcal{P}(\mathbb{R}^{2dN})$  be symmetric with*

$$\int_{\mathbb{R}^{2dN}} \sum_{j=1}^N |v_j|^{2p} d\rho(\mathbf{r}, \mathbf{v}) < \infty.$$

*Then there exists a constant  $C = C(\psi, \sigma) > 0$  (independent of  $N$ ) such that, for  $\gamma \in [0, 2)$  and  $t \geq 0$ ,*

$$\begin{aligned} & \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(t) \rangle^{2p} \right) \\ & \leq 2^{p-1} 2^{\frac{2p}{2-\gamma}} \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(0) \rangle^{2p} \right) + \frac{1}{2} \left( C_p \frac{2-\gamma}{p} \right)^{\frac{2p}{2-\gamma}} t^{\frac{2p}{2-\gamma}}, \end{aligned}$$

*where  $C_p = C \lambda_{2p} 2^{4p}$ , and, for  $\gamma = 2$ ,*

$$\mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(t) \rangle^{2p} \right) \leq 2^p \mathbb{E} \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(0) \rangle^{2p} \right) e^{C_p t}, \quad t \geq 0. \quad (40)$$

Moreover, there exists another constant  $C' > 0$  such that

$$\begin{aligned} \mathbb{E}^N \left( \sup_{s \in [0, t]} \langle V_1^N(s) \rangle^{2p-\gamma} \right) &\leq \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(0) \rangle^{2p-\gamma} \right) \\ &\quad + C' \lambda_{2p} 2^{2p} \int_0^t \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(s) \rangle^{2p} \right) ds. \end{aligned} \quad (41)$$

*Proof.* Both estimates can be shown in the same way as (36). Indeed, it follows from Lemma 4.2 and Lemma 4.5

$$\begin{aligned} &\mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N |V_j^N(t)|^{2p} \right) \\ &\leq \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N |V_j^N(0)|^{2p} \right) + C \lambda_{2p} 2^{3p} \int_0^t \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(s) \rangle^{2p-2+\gamma} \right) ds \\ &\leq \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N |V_j^N(0)|^{2p} \right) + C \lambda_{2p} 2^{3p} \int_0^t \left[ \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(s) \rangle^{2p} \right) \right]^{1-\frac{2-\gamma}{2p}} ds \end{aligned}$$

where we have used twice the Jensen inequality. Next observe that, by  $1 + |v|^{2p} \leq (v)^{2p} \leq 2^p(1 + |v|^{2p})$  and the previous estimate we have

$$\begin{aligned} \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(t) \rangle^{2p} \right) &\leq 2^p + 2^p \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N |V_j^N(t)|^{2p} \right) \\ &\leq 2^p \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(0) \rangle^{2p} \right) \\ &\quad + C_p \int_0^t \left[ \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(s) \rangle^{2p} \right) \right]^{1-\frac{2-\gamma}{2p}} ds. \end{aligned}$$

For  $\gamma = 2$  we apply the Gronwall lemma, for  $\gamma \in [0, 2)$  we may apply a nonlinear version of the Gronwall lemma stated in the appendix. To be more rigorous one has to consider the above estimates first for the variables  $\mathbf{V}^{N,m}(t) := \mathbf{V}^N(t \wedge \tau_m)$  where  $\tau_m$  is a stopping time chosen in such a way that  $\mathbf{V}^N(t \wedge \tau_m -)$  is bounded. Obtaining the desired estimates for  $\mathbf{V}^{N,m}(t)$  (with constants independent of  $m$ ), one may then pass to the limit  $m \rightarrow \infty$ . Since such type of arguments are rather standard and have been performed for the proof of (36), we leave the details for the reader.

Concerning estimate (41) we proceed similarly as in the proof of (36) to find that

$$\begin{aligned} &\sup_{s \in [0, t]} \langle V_1^N(s) \rangle^{2p-\gamma} \leq \langle V_1^N(0) \rangle^{2p-\gamma} \\ &\quad + \int_0^t \int_{\{1, \dots, N\}^2 \times \mathbb{R}^d \times \mathbb{R}_+} |\langle V_1^N(s-) + G_1 \rangle^{2p-\gamma} - \langle V_1(s-) \rangle^{2p-\gamma}| d\mathcal{N}, \end{aligned}$$

where  $G_1 = G_1(\mathbf{R}(s), \mathbf{V}(s-), u, l, l', z)$ . Taking expectations gives

$$\mathbb{E}^N \left( \sup_{s \in [0, t]} \langle V_1^N(s) \rangle^{2p-\gamma} \right) \leq \mathbb{E}^N (\langle V_1^N(0) \rangle^{2p-\gamma}) + \int_0^t \mathbb{E}^N (\mathcal{H}(s)) ds,$$

where by Lemma 4.1

$$\begin{aligned} \mathcal{H}(s) &= \frac{1}{N} \sum_{j=1}^N \psi(R_1(s) - R_j(s)) \sigma(V_1(s) - V_j(s)) \\ &\quad \cdot \int_{\mathbb{R}^d} |\langle \eta V_j(s) + (1-\eta)V_1(s) + \eta u \rangle^{2p-\gamma} - \langle V_1(s) \rangle^{2p-\gamma}| a(u) du \\ &\leq C \lambda_{2p} 2^{2p} \left( \langle V_1(s) \rangle^{2p} + \frac{1}{N} \sum_{j=1}^N \langle V_j(s) \rangle^{2p} \right). \end{aligned}$$

This gives

$$\begin{aligned} &\mathbb{E}^N \left( \sup_{s \in [0, t]} \langle V_1^N(s) \rangle^{2p-\gamma} \right) \\ &\leq \mathbb{E}^N (\langle V_1^N(0) \rangle^{2p-\gamma}) + C \lambda_{2p} 2^{2p} \int_0^t \left( \langle V_1(s) \rangle^{2p} + \frac{1}{N} \sum_{j=1}^N \langle V_j(s) \rangle^{2p} \right) ds \\ &= \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(0) \rangle^{2p-\gamma} \right) + 2C \lambda_{2p} 2^{2p} \int_0^t \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(s) \rangle^{2p} \right) ds, \end{aligned}$$

where we have used that all particles are indistinguishable. This proves the assertion.  $\square$

Using similar arguments combined with Lemma 4.1 and Lemma 4.4 yield propagation of exponential moments.

**Corollary 4.** *Suppose that  $\gamma = 0$ ,  $\eta = 1$ , and there exist  $\delta > 0$  and  $\kappa \in (0, 1]$  such that (24) holds. Then there exists a constant  $C = C(\psi, \sigma) > 0$  such that*

$$\mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N e^{\delta \langle V_j^N(t) \rangle^\kappa} \right) + \mathbb{E}^N \left( \sup_{s \in [0, t]} e^{\delta \langle V_1^N(s) \rangle^\kappa} \right) \leq \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N e^{\delta \langle V_j^N(0) \rangle^\kappa} \right) e^{Ct}.$$

A similar assertion holds also for  $\eta \in (0, 1)$ , provided  $\kappa = 1$ .

**5. The mean-field limit  $N \rightarrow \infty$ .** In this section we perform the limit  $N \rightarrow \infty$  and identify the corresponding limiting process, i.e. we prove Theorem 3.4 and partially Theorem 3.3. The exponential moment estimates can be deduced by similar arguments given in this section. Finally, the proof of Theorem 3.3 is then complete once we have also shown the uniqueness results from Section 6.

For each  $N \geq 2$ , let  $\rho^{(N)} \in \mathcal{P}(\mathbb{R}^{2dN})$  be given by

$$\rho^{(N)}(dr_1, dv_1, \dots, dr_N, dv_N) = \bigotimes_{k=1}^N \mu_0(dr_k, dv_k)$$

and denote by  $(R_k^N, V_k^N)_{k=1, \dots, N}$  the unique weak solution to (29) defined on a stochastic basis  $(\Omega_N, \mathcal{F}^N, (\mathcal{F}_t^N)_{t \geq 0}, \mathbb{P}^N)$  with the usual conditions. Let  $\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))$

be the space of probability measures over the Skorokhod space  $D(\mathbb{R}_+; \mathbb{R}^{2d})$ , similarly let  $\mathcal{P}(\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d})))$  be the space of probability measures over  $\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))$  equipped with the weak topology. Define a sequence of empirical measures

$$\mu^{(N)} = \frac{1}{N} \sum_{k=1}^N \delta_{(R_k^N, V_k^N)}, \quad (42)$$

i.e.  $\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))$ -valued random variables and let  $\pi^{(N)} \in \mathcal{P}(\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d})))$  be the law of  $\mu^{(N)}$ . The proof of Theorem 3.4 and partially Theorem 3.3 consists of the following two steps

- Step 1. Prove that  $\pi^{(N)}$  is relatively compact and show that each limit is supported on processes having the desired moment bounds.  
 Step 2. Prove that each limit  $\pi^{(\infty)}$  of a subsequence of  $\pi^{(N)}$  is supported on solutions to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ .

**5.1. Compactness and moment estimates.** We first show that  $(\pi^{(N)})_{N \geq 2}$  is relatively compact.

**Proposition 2.**  $(\pi^{(N)})_{N \geq 2}$  is relatively compact in  $\mathcal{P}(\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d})))$ .

*Proof.* In view of [28, Proposition 2.2], see also Corollary 2, it suffices to show that  $(R_1^N, V_1^N)$  is tight in  $D(\mathbb{R}_+; \mathbb{R}^{2d})$ . First we observe that

$$\sup_{t \in [0, T]} \mathbb{E}^N (|R_1^N(t)|) \leq \sup_{N \geq 2} \mathbb{E}^N (|R_1^N(0)|) + T \sup_{N \geq 2} \mathbb{E}^N \left( \sup_{t \in [0, T]} |V_1^N(t)| \right) < \infty,$$

where the right-hand side is finite due to the moment estimates of previous section. We seek to apply the Aldous criterion (see e.g. [20]). For each  $N \geq 2$  let  $S^N, T^N$  be  $(\mathcal{F}_t^N)_{t \geq 0}$  stopping times such that for  $M \in \mathbb{N}$  and  $\delta \in (0, 1]$  we have  $S^N \leq T^N \leq S^N + \delta$  and  $S^N, T^N \leq M$ . Then

$$\mathbb{E}^N (|R_1^N(T^N) - R_1^N(S^N)|) \leq \delta \sup_{N \geq 2} \mathbb{E}^N \left( \sup_{\tau \in [0, M]} |V_1^N(\tau)| \right)$$

and similarly by (29) and the exchangeability of the particles

$$\begin{aligned} & \mathbb{E}^N (|V_1^N(T^N) - V_1^N(S^N)|) \\ & \leq \frac{C}{N} \sum_{j=1}^N \mathbb{E}^N \left( \int_{S^N}^{T^N} \int_{\mathbb{R}^d} |u + (V_j^N(\tau) - V_1^N(\tau))| (\langle V_1^N(\tau) \rangle^\gamma + \langle V_j^N(\tau) \rangle^\gamma) a(u) du d\tau \right) \\ & \leq C \mathbb{E}^N \left( \int_{S^N}^{T^N} \left( \langle V_1^N(\tau) \rangle^{1+\gamma} + \frac{1}{N} \sum_{j=1}^N \langle V_j^N(\tau) \rangle^{1+\gamma} \right) d\tau \right) \\ & \leq C \delta \sup_{N \geq 2} \mathbb{E}^N \left( \sup_{\tau \in [0, M]} \langle V_1^N(\tau) \rangle^{1+\gamma} \right) < \infty \end{aligned}$$

where the last term is finite due to the moment estimates on the process. This proves the assertion.  $\square$

For  $\nu \in \mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))$  let  $\nu_t \in \mathcal{P}(\mathbb{R}^{2d})$  be the time-marginal at time  $t \geq 0$  and, for  $q \geq 0$ , set

$$\|\nu_t\|_q := \int_{\mathbb{R}^{2d}} \langle v \rangle^q \nu_t(dr, dv).$$

The next lemma provides moment estimates for the limits of the empirical measure.

**Lemma 5.1.** *There exists a constant  $C = C(\psi, \sigma) > 0$  such that for all  $t \geq 0$  we have, for  $\gamma \in [0, 2)$ ,*

$$\begin{aligned} & \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} \|\nu_t\|_{2p} d\pi^{(\infty)}(\nu) \\ & \leq 2^{p-1} 2^{\frac{2p}{2-\gamma}} \int_{\mathbb{R}^{2d}} \langle v \rangle^{2p} \mu_0(dr, dv) + \frac{1}{2} \left( C_p \frac{2-\gamma}{p} \right)^{\frac{2p}{2-\gamma}} t^{\frac{2p}{2-\gamma}}, \end{aligned}$$

with  $C_p = C\lambda_{2p}2^{4p}$  and, for  $\gamma = 2$ ,

$$\int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} \|\nu_t\|_{2p} d\pi^{(\infty)}(\nu) \leq 2^p \left( \int_{\mathbb{R}^{2d}} \langle v \rangle^{2p} \mu_0(dr, dv) \right) e^{C_p t}.$$

*Proof.* By approximation and the Lemma of Fatou we get

$$\begin{aligned} \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} \|\nu_t\|_{2p} d\pi^{(\infty)}(\nu) & \leq \sup_{N \geq 2} \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} \|\nu_t\|_{2p} d\pi^{(N)}(\nu) \\ & = \sup_{N \geq 2} \mathbb{E}^N \left( \frac{1}{N} \sum_{j=1}^N \langle V_j^N(t) \rangle^{2p} \right). \end{aligned}$$

The assertion now follows from our known moment estimates.  $\square$

From this we readily deduce, after we have completed Step 2, the desired moment estimate (25). Estimates (26) and (23) can be now deduced from the Itô formula.

**5.2. Identifying the limit.** The following shows that each limit point  $\pi^{(\infty)}$  of a subsequence of  $(\pi^{(N)})_{N \geq 2}$  is supported on solutions to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ .

**Proposition 3.** *Let  $\pi^{(\infty)} \in \mathcal{P}(\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d})))$  be any weak limit of a subsequence of  $(\pi^{(N)})_{N \geq 2}$ . Then  $\pi^{(\infty)}$ -a.a.  $\mu \in \mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))$  solve the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ .*

The rest of this section is devoted to the proof of this proposition. It is not difficult to see that the complement of

$$D_\mu = \{t > 0 \mid \mu((r, v) \in D(\mathbb{R}_+; \mathbb{R}^{2d}) : (r(t), v(t)) = (r(t-), v(t-))) = 1\}$$

is at most countable and the coordinate function  $(r, v) \mapsto (r(t), v(t))$  is  $\mu$ -a.s. continuous, for any  $t \in D_\mu$  and any  $\mu \in \mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))$ . Moreover, we can show that also the complement of

$$D(\pi^{(\infty)}) = \left\{ t > 0 \mid \pi^{(\infty)}(\mu \in \mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d})) : t \in D_\mu) = 1 \right\}$$

is at most countable.

Let  $0 \leq t_1, \dots, t_m \leq s \leq t$  with  $t_1, \dots, t_m, s, t \in D(\pi^{(\infty)})$ ,  $m \in \mathbb{N}$ ,  $g_1, \dots, g_m \in C_b(\mathbb{R}^{2d})$  and  $g \in C_c^1(\mathbb{R}^{2d})$ . For  $(r, v) \in D(\mathbb{R}_+; \mathbb{R}^{2d})$  and  $\mu \in \mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))$  let  $H(\mu; r, v)$  be given by

$$\left( g(r(t), v(t)) - g(r(s), v(s)) - \int_s^t (A(\mu_\tau)g)(r(\tau), v(\tau)) d\tau \right) \prod_{j=1}^m g_j(r(t_j), v(t_j)) \quad (43)$$

and define

$$F(\mu) := \int_{D(\mathbb{R}_+; \mathbb{R}^{2d})} H(\mu; r, v) \mu(dr, dv). \quad (44)$$

It is clear that  $\mu$  is a solution to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ , provided  $\mu((r(0), v(0)) \in \cdot) = \mu_0$ , (15) holds and  $F(\mu) = 0$ . Since, by Lemma 5.1,  $\pi^{(\infty)}$ -a.a.  $\mu$  satisfy (15) and  $\mu((r(0), v(0)) \in \cdot) = \mu_0$ , it suffices to show that

- (a)  $\lim_{N \rightarrow \infty} \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F(\mu)|^2 d\pi^{(N)}(\mu) = 0$ ,
- (b)  $\lim_{N \rightarrow \infty} \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F(\mu)| d\pi^{(N)}(\mu) = \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F(\mu)| d\pi^{(\infty)}(\mu)$ ,

where for simplicity of notation  $\pi^{(N)}$  denotes the subsequence converging weakly to  $\pi^{(\infty)}$ . Let us first prove (a).

**Lemma 5.2.** *Assertion (a) is satisfied.*

*Proof.* Let  $\tilde{\mathcal{N}}(ds, dl, dl', du, dz)$  be the compensated Poisson random measure on  $\mathbb{R}_+ \times \{1, \dots, N\}^2 \times \mathbb{R}^d$  and let  $M_{s,t}^{N,k}$  be given by

$$\int_s^t \int_E (g(R_k^N(\tau), V_k^N(\tau-)) + G_k - g(R_k^N(\tau), V_k^N(\tau-))) \tilde{\mathcal{N}}(d\tau, dl, dl', du, dz),$$

where  $E := \{1, \dots, N\}^2 \times \mathbb{R}^d \times \mathbb{R}_+$  and  $G_k = G_k(\mathbf{R}^N(\tau), \mathbf{V}^N(\tau-), u, l, l', z) \in \mathbb{R}^d$  denotes  $\mathbf{G} = (G_1, \dots, G_N) \in (\mathbb{R}^d)^N$  given by (30). Then

$$\begin{aligned} (A(\mu^{(N)})g)(R_k^N, V_k^N) &= V_k^N \cdot (\nabla_r g)(R_k^N, V_k^N) \\ &+ \frac{1}{N} \sum_{j=1}^N \psi(R_k^N - R_j^N) \sigma(V_k^N - V_j^N) \\ &\int_{\mathbb{R}^d} (g(R_k^N, v^*(V_k^N, V_j^N, u)) - g(R_k^N, V_k^N)) a(u) du \end{aligned}$$

and from the Itô formula one immediately obtains

$$g(R_k^N(t), V_k^N(t)) = g(R_k^N(s), V_k^N(s)) + \int_s^t (A(\mu_\tau^{(N)})g)(R_k^N(\tau), V_k^N(\tau)) d\tau + M_{s,t}^{N,k}.$$

This shows that

$$F(\mu^{(N)}) = \frac{1}{N} \sum_{k=1}^N H(\mu^{(N)}; R_k^N, V_k^N) = \frac{1}{N} \sum_{k=1}^N M_{s,t}^{N,k} \prod_{j=1}^m g_j(R_k^N(t_j), V_k^N(t_j)).$$

For the Doob-Meyer process of  $M_{s,t}^{N,k}$  we obtain

$$\begin{aligned} \langle M_{s,t}^{N,k} \rangle &= \frac{1}{N} \sum_{j=1}^N \int_s^t \int_{\mathbb{R}^d} (g(R_k^N, V_j^N + u) - g(R_k^N, V_k^N))^2 \\ &\quad \psi(R_k^N - R_j^N) \sigma(V_k^N - V_j^N) a(u) du d\tau \\ &\leq \frac{C}{N} \sum_{j=1}^N \int_s^t (\langle V_k^N(\tau) \rangle^\gamma + \langle V_j^N(\tau) \rangle^\gamma) d\tau. \end{aligned}$$

Using the moment estimates of previous section we obtain  $\mathbb{E}^N(\langle M_{s,t}^{N,k} \rangle) \leq C$  for all  $k = 1, \dots, N$  and some constant  $C = C(\psi, \sigma, a, g)$  independent of  $N$ . Using the particular form of  $\mathbf{G}$  (see (30)), we obtain for the covariation process  $\langle M_{s,t}^{N,k}, M_{s,t}^{N,j} \rangle = 0$  for all  $k \neq j$ . Hence we conclude from the properties of the processes  $\langle M_{s,t}^{N,k} \rangle$  and

$$\begin{aligned}
& \langle M_{s,t}^{N,k}, M_{s,t}^{N,j} \rangle \\
&= \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F(\nu)|^2 d\pi^{(N)}(\nu) \\
&= \frac{1}{N^2} \sum_{k \neq j} \mathbb{E}^N \left( M_{s,t}^{N,k} M_{s,t}^{N,j} \prod_{l_1=1}^m g_{l_1}(R_k^N(t_{l_1}), V_k^N(t_{l_1})) \prod_{l_2=1}^m g_{l_2}(R_k^N(t_{l_2}), V_k^N(t_{l_2})) \right) \\
&\quad + \frac{1}{N^2} \sum_{k=1}^N \mathbb{E}^N \left( (M_{s,t}^{N,k})^2 \prod_{l=1}^m g_l(R_k^N(t_l), V_k^N(t_l))^2 \right) \\
&= \frac{1}{N^2} \sum_{k \neq j} \mathbb{E}^N \left( \langle M_{s,t}^{N,k}, M_{s,t}^{N,j} \rangle \prod_{l_1=1}^m g_{l_1}(R_k^N(t_{l_1}), V_k^N(t_{l_1})) \prod_{l_2=1}^m g_{l_2}(R_k^N(t_{l_2}), V_k^N(t_{l_2})) \right) \\
&\quad + \frac{1}{N^2} \sum_{k=1}^N \mathbb{E}^N \left( \langle M_{s,t}^{N,k} \rangle \prod_{l=1}^m g_l(R_k^N(t_l), V_k^N(t_l))^2 \right) \\
&\leq \frac{C(\psi, \sigma, a, g, g_1, \dots, g_m)}{N},
\end{aligned}$$

which proves the assertion.  $\square$

Next we prove assertion (b).

**Lemma 5.3.** *Assertion (b) is satisfied.*

*Proof.* Take  $\varphi \in C^\infty(\mathbb{R}_+)$  with  $\mathbb{1}_{[0,1]} \leq \varphi \leq \mathbb{1}_{[0,2]}$ . For  $R > 0$  and  $\nu \in \mathcal{P}(\mathbb{R}^{2d})$  let

$$\begin{aligned}
(A_R(\nu)g)(r, v) &= v \cdot (\nabla_r g)(r, v) \\
&\quad + \int_{\mathbb{R}^{2d} \times \mathbb{R}^d} \varphi \left( \frac{|w|^2}{R^2} \right) (g(r, v^*(v, w, u)) - g(r, v)) \\
&\quad \quad \quad \psi(r - q) \sigma(v - w) \nu(dq, dw) a(u) du.
\end{aligned}$$

Then it is not difficult to see that

$$\mathcal{P}(\mathbb{R}^{2d}) \times \mathbb{R}^{2d} \ni (\nu, r, v) \mapsto (A_R(\nu)g)(r, v)$$

is jointly continuous where  $\mathcal{P}(\mathbb{R}^{2d})$  is endowed with the topology of weak convergence. Moreover one can show that for some constant  $C = C(\psi, \sigma, a, g)$

$$\begin{aligned}
|A_R(\nu)g(r, v) - A(\nu)g(r, v)| &\leq C \int_{\mathbb{R}^{2d}} \mathbb{1}_{\{|w| > R\}} \sigma(v - w) \nu(dw) \\
&\leq \frac{C}{R^{1/2}} \|\nu\|_{\gamma + \frac{1}{2}} \langle v \rangle^\gamma.
\end{aligned} \tag{45}$$

Let  $H_R$  be defined by (43) with  $A$  replaced by  $A_R$  and define  $F_R(\mu)$  by (44) with  $H$  replaced by  $H_R$ . Then we obtain

$$\begin{aligned}
& \left| \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F(\mu)| d\pi^{(N)}(\mu) - \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F(\mu)| d\pi^{(\infty)}(\mu) \right| \\
&\leq \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F(\mu) - F_R(\mu)| d\pi^{(N)}(\mu) + \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F_R(\mu) - F(\mu)| d\pi^{(\infty)}(\mu) \\
&\quad + \left| \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F_R(\mu)| d\pi^{(N)}(\mu) - \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F_R(\mu)| d\pi^{(\infty)}(\mu) \right| \\
&= I_1 + I_2.
\end{aligned}$$

Using (45) we obtain for  $T > t$  and some constant  $C = C(g, g_1, \dots, g_m, \psi, \sigma, a)$

$$\begin{aligned} & |F(\mu) - F_R(\mu)| \\ & \leq C \int_s^t \int_{D(\mathbb{R}_+; \mathbb{R}^{2d})} |(A(\mu_\tau)g)(r(\tau), v(\tau)) - (A_R(\mu_\tau)g)(r(\tau), v(\tau))| \mu(dr, dv) d\tau \\ & \leq \frac{C}{R^{1/2}} \int_s^t \|\mu_\tau\|_{\gamma+\frac{1}{2}} \|\mu_\tau\|_\gamma d\tau \\ & \leq \frac{C}{R^{1/2}} \int_s^t \|\mu_\tau\|_{2\gamma+1} d\tau, \end{aligned}$$

where we have used that  $\|\mu_\tau\|_{\gamma+\frac{1}{2}} \|\mu_\tau\|_\gamma \leq \|\mu_\tau\|_{\gamma+\frac{1}{2}}^2 \leq \|\mu_\tau\|_{2\gamma+1}$ . Using the moment estimates from Corollary 3 and Lemma 5.1 we find a constant  $C > 0$  such that  $\sup_{N \geq 2} I_1 \leq CR^{-1/2}$ . Hence it remains to prove that  $I_2 \rightarrow 0$  as  $N \rightarrow \infty$  for any fixed  $R > 0$ .

Fix  $R > 0$  and recall that  $\varphi$  is a smooth function on  $\mathbb{R}_+$  satisfying  $\mathbb{1}_{[0,1]} \leq \varphi \leq \mathbb{1}_{[0,2]}$ . Define

$$\begin{aligned} H_{R,m}^1(\mu; x) & := \varphi \left( \frac{\sup_{\tau \in [s,t]} \langle v(\tau) \rangle^2}{m^2} \right) H_R(\mu; x), \\ H_{R,m}^2(\mu; x) & := \left( 1 - \varphi \left( \frac{\sup_{\tau \in [s,t]} \langle v(\tau) \rangle^2}{m^2} \right) \right) H_R(\mu; x), \end{aligned}$$

and let  $F_{R,m}^j$  be given by (44) with  $H$  replaced by  $H_{R,m}^j$ ,  $j = 1, 2$ . Then we obtain  $I_2 \leq J_1 + J_2$ , where

$$J_j = \left| \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F_{R,m}^j(\mu)| d\pi^{(N)}(\mu) - \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} |F_{R,m}^j(\mu)| d\pi^{(\infty)}(\mu) \right|.$$

For any  $m \geq 1$  and  $\mu \in \mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))$  we find a constant  $C$  independent of  $N$  and  $m$  such that

$$\begin{aligned} |F_{R,m}^2(\mu)| & \leq \int_{D(\mathbb{R}_+; \mathbb{R}^{2d})} \mathbb{1}_{\left\{ \sup_{\tau \in [s,t]} \langle v(\tau) \rangle > m \right\}} |H_R(\mu; r, v)| \mu(dr, dv) \\ & \leq C \int_{D(\mathbb{R}_+; \mathbb{R}^{2d})} \int_s^t \mathbb{1}_{\left\{ \sup_{\tau \in [s,t]} \langle v(\tau) \rangle > m \right\}} \langle v(\tau) \rangle^\gamma d\tau \mu(dr, dv) \\ & \leq \frac{C}{m} \int_{D(\mathbb{R}_+; \mathbb{R}^{2d})} \sup_{\tau \in [s,t]} \langle v(\tau) \rangle^{1+\gamma} \mu(dr, dv). \end{aligned}$$

The moment estimates from Corollary 3 and a similar application of the Lemma of Fatou as in Lemma 5.1 gives

$$J_2 \leq \frac{C}{m} \int_{\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))} \int_{D(\mathbb{R}_+; \mathbb{R}^{2d})} \sup_{\tau \in [s,t]} \langle v(\tau) \rangle^{1+\gamma} \mu(dr, dv) d(\pi^{(N)} + \pi^{(\infty)})(\mu) \leq \frac{C}{m}.$$

Hence it suffices to show that  $J_1 \rightarrow 0$  as  $N \rightarrow \infty$  for each fixed  $R, m$ .

Note that  $H_{R,m}^1$  is bounded and jointly continuous in  $(\mu, r, v)$ . Hence  $F_{R,m}^1$  is continuous and bounded on  $\mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))$ . Using the weak convergence  $\pi^{(N)} \rightarrow \pi^{(\infty)}$  as  $N \rightarrow \infty$  we conclude that also  $J_1 \rightarrow 0$  as  $N \rightarrow \infty$ , for each fixed  $R, m$ .  $\square$

**6. Uniqueness for bounded coefficients.** In this section we study uniqueness for the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$  as well as for the kinetic equation (9) for the case  $\gamma = 0$ . The following is our main result in this case.

**Theorem 6.1.** *Suppose that  $\gamma = 0$  and  $\eta \in (0, 1]$ . Then for each  $\mu_0 \in \mathcal{P}(\mathbb{R}^{2d})$  there exists at most one solution to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ . In particular, there exists at most one weak solution to the mean-field SDE (17).*

The proof of this theorem is deduced from the following considerations.

**6.1. Uniqueness for the time-marginals.** In this section we study uniqueness and stability for the time-marginals, i.e. solutions to (14). More precisely, we prove an a priori bound for any two solutions to (14) with respect to the total variation distance

$$\|\mu - \nu\|_{\text{TV}} = \sup \{ \langle g, \mu - \nu \rangle : g \in B(\mathbb{R}^{2d}), \|g\|_\infty \leq 1 \},$$

where  $B(\mathbb{R}^{2d})$  denotes the space of all bounded measurable functions on  $\mathbb{R}^{2d}$ . The proof of such bound relies on a mild formulation of (14) given below.

**Lemma 6.2.** *Let  $(\mu_t)_{t \geq 0} \subset \mathcal{P}(\mathbb{R}^{2d})$  be given. Then  $(\mu_t)_{t \geq 0}$  satisfies (14) if and only if*

$$\langle g, \mu_t \rangle = \langle S(t)g, \mu_0 \rangle + \int_0^t \langle \mathcal{Q}S(t-s)g, \mu_s \otimes \mu_s \rangle ds \quad (46)$$

holds for all  $g \in C_c^1(\mathbb{R}^{2d})$ , where  $S(t-s)g(r, v) = g(r + (t-s)v, v)$ . Moreover, (46) naturally extends to all  $g \in B(\mathbb{R}^{2d})$ .

*Proof.* If  $(\mu_t)_{t \geq 0}$  is a solution to (14), then

$$\frac{d}{ds} \langle S(t-s)g, \mu_s \rangle = \int_{\mathbb{R}^{2d} \times \mathbb{R}^{2d}} \mathcal{Q}S(t-s)g(r, v; q, w) \mu_s(dr, dv) \mu_s(dq, dw)$$

and hence integrating over  $[0, t]$  gives (46). Conversely, suppose that  $(\mu_t)_{t \geq 0}$  satisfies (46). Differentiating (46) readily yields (14). For the last part we use

$$|\mathcal{Q}S(t-s)g(r, v; q, w)| \leq 2\|g\|_\infty \|\psi\|_\infty \|\sigma\|_\infty. \quad (47)$$

combined with dominated convergence and standard density arguments.  $\square$

The following is our main estimate for solutions to (14) in the case  $\gamma = 0$ .

**Theorem 6.3.** *Suppose that  $\gamma = 0$  and  $\eta \in (0, 1]$ . Let  $(\mu_t)_{t \geq 0}$  and  $(\nu_t)_{t \geq 0}$  be two solutions to (14). Then*

$$\|\mu_t - \nu_t\|_{\text{TV}} \leq \|\mu_0 - \nu_0\|_{\text{TV}} \exp(4\|\psi\|_\infty \|\sigma\|_\infty t), \quad t \geq 0.$$

*Proof.* Let  $g \in B(\mathbb{R}^{2d})$  be such that  $\|g\|_\infty \leq 1$ . Then, by (46),

$$\begin{aligned} \langle g, \mu_t - \nu_t \rangle &= \langle S(t)g, \mu_0 - \nu_0 \rangle + \int_0^t \langle \mathcal{Q}S(t-s)g, \mu_s \otimes \mu_s - \nu_s \otimes \nu_s \rangle ds \\ &= \langle S(t)g, \mu_0 - \nu_0 \rangle + \int_0^t \langle \mathcal{Q}S(t-s)g, \mu_s \otimes (\mu_s - \nu_s) \rangle ds \\ &\quad + \int_0^t \langle \mathcal{Q}S(t-s)g, (\mu_s - \nu_s) \otimes \nu_s \rangle ds \\ &\leq \|\mu_0 - \nu_0\|_{\text{TV}} + 4\|\psi\|_\infty \|\sigma\|_\infty \int_0^t \|\mu_s - \nu_s\|_{\text{TV}} ds, \end{aligned}$$

where we have used  $\|S(t)g\|_\infty \leq 1$  and (47) to obtain

$$\begin{aligned} \langle \mathcal{Q}S(t-s)g, (\mu_s - \nu_s) \otimes \nu_s \rangle &\leq \sup_{(q,w) \in \mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} (\mathcal{Q}S(t-s)g)(r, v; q, w) (\mu_s - \nu_s)(dr, dv) \\ &\leq 2\|\psi\|_\infty \|\sigma\|_\infty \|\mu_s - \nu_s\|_{\text{TV}} \end{aligned}$$

and similarly

$$\langle \mathcal{Q}S(t-s)g, (\mu_s - \nu_s) \otimes \nu_s \rangle \leq 2\|\psi\|_\infty \|\sigma\|_\infty \|\mu_s - \nu_s\|_{\text{TV}}.$$

Taking the supremum over all  $g \in B(\mathbb{R}^{2d})$  with  $\|g\|_\infty \leq 1$  and then applying the Gronwall lemma yields the assertion.  $\square$

**6.2. Uniqueness in law for the Vlasov-McKean equation.** Below we prove that the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$  has at most one solution.

**Proposition 4.** *Suppose that  $\gamma = 0$  and let  $\mu_0 \in \mathcal{P}(\mathbb{R}^{2d})$ . Then there exists at most one solution  $\mu \in \mathcal{P}(D(\mathbb{R}_+; \mathbb{R}^{2d}))$  to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ .*

*Proof.* Let  $\mu, \tilde{\mu}$  be solutions to the nonlinear martingale problem  $(A, C_c^1(\mathbb{R}^{2d}), \mu_0)$ . Then their time-marginals  $(\mu_t)_{t \geq 0}$  and  $(\tilde{\mu}_t)_{t \geq 0}$  both solve (14) and hence coincide, i.e.  $\mu_t = \tilde{\mu}_t$ , for all  $t \geq 0$ . Consequently  $\mu$  and  $\tilde{\mu}$  are both solutions to the linear time-inhomogeneous martingale problem  $(A_t, C_c^1(\mathbb{R}^{2d}), \mu_0)$ , i.e.

$$g(x(t)) - g(x(0)) - \int_0^t (A(\mu_s)g)(x(s)) ds, \quad t \geq 0, \quad g \in C_c^1(\mathbb{R}^{2d}), \quad (48)$$

is a martingale with respect to  $\mu$  and  $\tilde{\mu}$ , where  $A_t := A(\mu_t) = A(\tilde{\mu}_t)$ . Using the same argument as in the proof of Theorem 6.3, one easily shows that there exists at most one solution  $(\rho_t)_{t \geq 0}$  to the time-inhomogeneous Fokker-Planck equation

$$\langle g, \rho_t \rangle = \langle g, \rho_0 \rangle + \int_0^t \langle A(\mu_s)g, \rho_s \rangle, \quad t \geq 0, \quad g \in C_c^1(\mathbb{R}^{2d}).$$

Applying [8, p.184, Theorem 4.2] we conclude that the time-inhomogeneous martingale problem (48) is well-posed and hence  $\mu = \tilde{\mu}$ .  $\square$

**7. Uniqueness for kinetic equation when  $\gamma \in (0, 2]$ .** In this section we provide some sufficient condition for uniqueness and stability of solutions to (14) in the case where  $\gamma \in (0, 2]$ . As before, it is not difficult to see that any solution to (14) still satisfies the mild formulation (46).

**7.1. Estimate on the total variation distance.** For  $\delta > 0$  let

$$\mathcal{U}(\gamma, \delta) = \left\{ (\mu_t)_{t \geq 0} \mid \sup_{t \in [0, T]} \mathcal{C}_\gamma(\delta, \mu_t) < \infty, \quad \forall T > 0 \right\}$$

where

$$\mathcal{C}_\gamma(\delta, \mu_t) := \int_{\mathbb{R}^{2d}} e^{\delta \langle v \rangle^\gamma} \mu_t(dr, dv). \quad (49)$$

The following is the main result on uniqueness and stability for (14).

**Theorem 7.1.** Fix  $\delta > 0$ . Then there exists a constant  $C = C(\psi, \sigma, \delta) > 0$  such that any two solutions  $(\mu_t)_{t \geq 0}, (\nu_t)_{t \geq 0} \in \mathcal{U}(\gamma, \delta)$  to (14) satisfy

$$\begin{aligned} \|\mu_t - \nu_t\|_{\text{TV}} &\leq \|\mu_0 - \nu_0\|_{\text{TV}} \\ &\quad + C \int_0^t \mathcal{C}_\gamma(\delta, \mu_s + \nu_s)^2 \|\mu_s - \nu_s\|_{\text{TV}} (1 + |\ln(\|\mu_s - \nu_s\|_{\text{TV}})|) ds. \end{aligned}$$

In particular, the following assertions hold

(a) There exists at most one solution to (14) in  $\mathcal{U}(\gamma, \delta)$ .

(b) Let  $\mu_0, \mu_0^{(n)} \in \mathcal{P}(\mathbb{R}^{2d})$  with

$$\|\mu_0 - \mu_0^{(n)}\|_{\text{TV}} \longrightarrow 0, \quad n \rightarrow \infty$$

and let  $(\mu_t)_{t \geq 0}$  and  $(\mu_t^{(n)})_{t \geq 0}$  be two solutions to (14) with initial condition  $\mu_0$  and  $\mu_0^{(n)}$ , respectively. Suppose that there exists  $\delta > 0$  such that

$$\sup_{n \geq 1} \sup_{t \in [0, T]} \mathcal{C}_\gamma(\delta, \mu_t + \mu_t^{(n)}) < \infty, \quad \forall T > 0.$$

Then, for any  $t \geq 0$ ,

$$\|\mu_t - \mu_t^{(n)}\|_{\text{TV}} \longrightarrow 0, \quad n \rightarrow \infty.$$

*Proof.* Let  $g \in B(\mathbb{R}^{2d})$  be such that  $\|g\|_\infty \leq 1$ . Using the mild formulation (46) we obtain

$$\begin{aligned} \langle g, \mu_t - \nu_t \rangle &= \langle S(t)g, \mu_0 - \nu_0 \rangle + \int_0^t \langle \mathcal{Q}S(t-s)g, \mu_s \otimes \mu_s - \nu_s \otimes \nu_s \rangle ds \\ &= \int_0^t \langle \mathcal{Q}S(t-s)g, \mu_s \otimes (\mu_s - \nu_s) \rangle ds + \int_0^t \langle \mathcal{Q}S(t-s)g, (\mu_s - \nu_s) \otimes \nu_s \rangle ds. \end{aligned}$$

Let  $\varphi$  be a smooth function on  $\mathbb{R}_+$  such that  $\mathbb{1}_{[0,1]} \leq \varphi \leq \mathbb{1}_{[0,2]}$  and set  $\varphi_R(w) := \varphi\left(\frac{\langle w \rangle^2}{R^2}\right)$ . Using the definition of  $\mathcal{Q}$  and  $(1 - \varphi_R(w)) \leq \mathbb{1}_{\{\langle w \rangle > R\}}$  we obtain

$$\begin{aligned} &|\langle \mathcal{Q}S(t-s)g, \mu_s \otimes (\mu_s - \nu_s) \rangle| \\ &\leq \int_{\mathbb{R}^{4d}} \varphi_R(w) |(\mathcal{Q}S(t-s)g)(r, v; q, w)| d\mu_s(r, v) d(\mu_s - \nu_s)(q, w) \\ &\quad + \int_{\mathbb{R}^{4d}} \mathbb{1}_{\{\langle w \rangle > R\}} |(\mathcal{Q}S(t-s)g)(r, v; q, w)| d\mu_s(r, v) d(\mu_s + \nu_s)(q, w) \\ &\leq C \|\mu_s\|_\gamma R^\gamma \|\mu_s - \nu_s\|_{\text{TV}} + C \|\mu_s\|_\gamma \int_{\mathbb{R}^{4d}} \mathbb{1}_{\{\langle w \rangle > R\}} \langle w \rangle^\gamma d(\mu_s + \nu_s)(q, w). \end{aligned}$$

For the last term we use similar arguments to [11] and [13]. Namely, using  $\langle w \rangle^\gamma \leq C e^{\frac{\delta}{2} \langle w \rangle^\gamma}$  for some constant  $C > 0$  large enough, we get

$$\begin{aligned} &\int_{\mathbb{R}^{4d}} \mathbb{1}_{\{\langle w \rangle > R\}} \langle w \rangle^\gamma d(\mu_s + \nu_s)(q, w) \\ &\leq C \int_{\mathbb{R}^{4d}} \mathbb{1}_{\{\langle w \rangle > R\}} e^{-\frac{\delta}{2} \langle w \rangle^\gamma} e^{\delta \langle w \rangle^\gamma} d(\mu_s + \nu_s)(q, w) \\ &\leq C e^{-\frac{\delta}{2} R^\gamma} \mathcal{C}_\gamma(\delta, \mu_s + \nu_s). \end{aligned}$$

Taking  $R^\gamma = \frac{2}{\delta} |\ln(\|\mu_s - \nu_s\|_{\text{TV}})|$  we deduce

$$\begin{aligned} &\langle \mathcal{Q}S(t-s)g, \mu_s \otimes (\mu_s - \nu_s) \rangle \\ &\leq C \mathcal{C}_\gamma(\delta, \mu_s + \nu_s) \|\mu_s\|_\gamma \|\mu_s - \nu_s\|_{\text{TV}} (1 + |\ln(\|\mu_s - \nu_s\|_{\text{TV}})|). \end{aligned}$$

Proceeding in the same way we can also show that

$$\begin{aligned} & \langle \mathcal{Q}S(t-s)g, (\mu_s - \nu_s) \otimes \nu_s \rangle \\ & \leq CC_\gamma(\delta, \mu_s + \nu_s) \|\nu_s\|_\gamma \|\mu_s - \nu_s\|_{\text{TV}} (1 + |\ln(\|\mu_s - \nu_s\|_{\text{TV}})|). \end{aligned}$$

Since  $\|\mu_s\|_\gamma + \|\nu_s\|_\gamma \leq CC_\gamma(\delta, \mu_s + \nu_s)$ , the desired inequality follows by taking the supremum over all  $g \in B(\mathbb{R}^{2d})$  with  $\|g\|_\infty \leq 1$ . Uniqueness and stability is a direct consequence of the a priori estimate we have shown, i.e. one may apply a generalization of the Gronwall inequality stated in the appendix.  $\square$

**7.2. Estimate on the Wasserstein distance.** In this part we prove estimates for solutions to (14) with respect to the Wasserstein distance

$$d(\mu, \nu) = \sup_{\|g\|_0 \leq 1} \frac{|g(r, v) - g(\tilde{r}, \tilde{v})|}{|r - \tilde{r}| + |v - \tilde{v}|}, \quad \|g\|_0 := \sup_{(r, v) \neq (\tilde{r}, \tilde{v})} \frac{|g(r, v) - g(\tilde{r}, \tilde{v})|}{|r - \tilde{r}| + |v - \tilde{v}|},$$

where  $\mu, \nu \in \mathcal{P}(\mathbb{R}^{2d})$  are supposed to have finite first moments. Since particles are transported by the transport operator  $v \cdot \nabla_r$ , it is more natural to use the shifted Wasserstein distance

$$d_t(\mu, \nu) = d(S(-t)^* \mu, S(-t)^* \nu), \quad t \geq 0,$$

where  $S(t)g(r, v) = g(r + vt, v)$  and  $S(t)^*$  is the adjoint operator defined by the relation

$$\langle S(t)g, \mu \rangle = \langle g, S(t)^* \mu \rangle, \quad g \in B(\mathbb{R}^{2d}), \quad \mu \in \mathcal{P}(\mathbb{R}^{2d}).$$

Below we will use another characterization of the shifted distance in terms of optimal couplings described as follows.

Introduce a one-parameter family of metrics on  $\mathbb{R}^{2d}$

$$|(r, v) - (\tilde{r}, \tilde{v})|_t := |(r - vt) - (\tilde{r} - \tilde{v}t)| + |v - \tilde{v}|, \quad t \geq 0$$

and related to this metrics define the time-dependent Lipschitz norms

$$\|g\|_t = \sup_{(r, v) \neq (\tilde{r}, \tilde{v})} \frac{|g(r, v) - g(\tilde{r}, \tilde{v})|}{|(r, v) - (\tilde{r}, \tilde{v})|_t}.$$

Given  $\mu, \nu \in \mathcal{P}(\mathbb{R}^{2d})$ , a coupling  $H$  of  $(\mu, \nu)$  is a probability measure on  $\mathbb{R}^{4d}$  such that its marginals are given by  $\mu$  and  $\nu$ , respectively. Let  $\mathcal{H}(\mu, \nu)$  the space of all such couplings. The reader may consult [30] for additional details on couplings and related Wasserstein distance.

**Proposition 5.** *Let  $\mu, \nu \in \mathcal{P}(\mathbb{R}^{2d})$  satisfy  $\int_{\mathbb{R}^{2d}} (|r| + |v|)(\mu + \nu)(dr, dv) < \infty$  and fix  $t \geq 0$ . Then there exists  $H_t \in \mathcal{H}(\mu, \nu)$  such that*

$$\begin{aligned} d_t(\mu, \nu) &= \sup_{\|\psi\|_0 \leq 1} \langle S(-t)\psi, \mu - \nu \rangle \\ &= \sup_{\|\psi\|_t \leq 1} \langle \psi, \mu - \nu \rangle = \int_{\mathbb{R}^{4d}} |(r, v) - (\tilde{r}, \tilde{v})|_t dH_t(r, v; \tilde{r}, \tilde{v}). \end{aligned} \tag{50}$$

*Proof.* The first equality follows from the definition of  $S(t)^*$ , the second equality from the definition of the norms  $\|\cdot\|_t$  while the third equality is a particular case of the Kantorovich-duality (see [30]).  $\square$

The following is our main coupling estimate for the Wasserstein distance  $d_t$ .

**Proposition 6.** *Suppose that  $\int_{\mathbb{R}^{2d}} |u|a(u)du < \infty$  and let  $\mu_0, \nu_0 \in \mathcal{P}(\mathbb{R}^{2d})$  satisfy*

$$\int_{\mathbb{R}^{2d}} (|r| + |v|)(\mu_0 + \nu_0)(dr, dv) < \infty.$$

Let  $(\mu_t)_{t \geq 0}$  and  $(\nu_t)_{t \geq 0}$  be two solutions to (14) satisfying

$$\int_0^T \int_{\mathbb{R}^{2d}} (|r| + |v|^{1+\gamma}) (\mu_t + \nu_t)(dr, dv)dt < \infty, \quad \forall T > 0.$$

For  $t \geq 0$ , let  $H_t \in \mathcal{H}(\mu_t, \nu_t)$  be such that

$$d_t(\mu_t, \nu_t) = \int_{\mathbb{R}^{4d}} |(r, v) - (\tilde{r}, \tilde{v})|_t dH_t(r, v; \tilde{r}, \tilde{v}). \quad (51)$$

Then for each  $T > 0$  there exists  $C > 0$  (independent of  $\mu_t, \nu_t$ ) such that, for any  $t \in [0, T]$ ,

$$d_t(\mu_t, \nu_t) \leq d_0(\mu_0, \nu_0) + C(T, a, \psi) \int_0^t \int_{\mathbb{R}^{8d}} \Lambda(r, v, q, w; \tilde{r}, \tilde{v}, \tilde{q}, \tilde{w}) dH_s^0 dH_s^1 ds,$$

where  $dH_s^0 = dH_s(r, v; \tilde{r}, \tilde{v})$ ,  $dH_s^1 = dH_s(q, w; \tilde{q}, \tilde{w})$  and

$$\begin{aligned} \Lambda(r, v, q, w; \tilde{r}, \tilde{v}, \tilde{q}, \tilde{w}) &= (\langle v \rangle + \langle w \rangle + \langle \tilde{v} \rangle + \langle \tilde{w} \rangle) |\sigma(v - w)\psi(r - q) - \sigma(\tilde{v} - \tilde{w})\psi(\tilde{r} - \tilde{q})| \\ &\quad + (|(r, w) - (\tilde{r}, \tilde{w})|_s + |(r, v) - (\tilde{r}, \tilde{v})|_s) \min\{\sigma(v - w), \sigma(\tilde{v} - \tilde{w})\}. \end{aligned}$$

*Proof.* It is not difficult to see that both solutions still satisfy the mild formulation (46) for any  $g$  with  $\|g\|_0 \leq 1$ . For simplicity of notation, let  $\tilde{\psi} = \psi(\tilde{r} - \tilde{q})$ ,  $\tilde{\sigma} = \sigma(\tilde{v} - \tilde{w})$  and similarly  $\psi = \psi(r - q)$  and  $\sigma = \sigma(v - w)$ . Using the definition of  $dH_s^0 dH_s^1$  together with  $x = x \wedge y + (x - y)_+$ , for  $x, y \geq 0$ , we obtain

$$\begin{aligned} &\langle S(-t)g, \mu_t - \nu_t \rangle - \langle g, \mu_0 - \nu_0 \rangle \\ &= \int_0^t \langle \mathcal{Q}S(-s)g, \mu_s \otimes \mu_s - \nu_s \otimes \nu_s \rangle ds \\ &= \int_0^t \int_{\mathbb{R}^{8d}} [(\mathcal{Q}S(-s)g)(r, v; \tilde{r}, \tilde{v}) - (\mathcal{Q}S(-s)g)(q, w; \tilde{q}, \tilde{w})] dH_s^0 dH_s^1 ds \\ &= \int_0^t \int_{\mathbb{R}^{9d}} \left\{ (S(-s)g(r, v^*(v, w, u)) - S(-s)g(r, v)) \psi \sigma \right. \\ &\quad \left. - (S(-s)g(\tilde{r}, v^*(\tilde{v}, \tilde{w}, u)) - S(-s)g(\tilde{r}, \tilde{v})) \tilde{\psi} \tilde{\sigma} \right\} a(u) dudH_s^0 dH_s^1 ds \\ &= \int_0^t \int_{\mathbb{R}^{9d}} \left\{ S(-s)g(r, v^*(v, w, u)) - S(-s)g(\tilde{r}, v^*(\tilde{v}, \tilde{w}, u)) \right. \\ &\quad \left. + S(-s)g(\tilde{r}, \tilde{v}) - S(-s)g(r, v) \right\} (\psi \sigma \wedge \tilde{\psi} \tilde{\sigma}) a(u) dudH_s^0 dH_s^1 ds \\ &\quad + \int_0^t \int_{\mathbb{R}^{9d}} (S(-s)g(r, v^*(v, w, u)) - S(-s)g(r, v)) (\psi \sigma - \tilde{\psi} \tilde{\sigma})_+ a(u) dudH_s^0 dH_s^1 ds \\ &\quad + \int_0^t \int_{\mathbb{R}^{9d}} (S(-s)g(\tilde{r}, v^*(\tilde{v}, \tilde{w}, u)) - S(-s)g(\tilde{r}, \tilde{v})) (\tilde{\psi} \tilde{\sigma} - \psi \sigma)_+ a(u) dudH_s^0 dH_s^1 ds \\ &= J_1 + J_2 + J_3. \end{aligned}$$

Using  $\|S(-s)g\|_s \leq 1$  we obtain

$$\begin{aligned}
J_2 + J_3 &\leq \int_0^t \int_{\mathbb{R}^{9d}} \left\{ |(r, v^*(v, w, u)) - (r, v)|_s \right. \\
&\quad \left. + |(\tilde{r}, v^*(\tilde{v}, \tilde{w}, u)) - (\tilde{r}, \tilde{v})|_s \right\} |\psi\sigma - \tilde{\psi}\tilde{\sigma}| a(u) dudH_s^0 dH_s^1 ds \\
&\leq \int_0^t (1+s) \int_{\mathbb{R}^{9d}} \left( |v^*(v, w, u) - v| \right. \\
&\quad \left. + |v^*(\tilde{v}, \tilde{w}, u) - \tilde{v}| \right) |\psi\sigma - \tilde{\psi}\tilde{\sigma}| a(u) dudH_s^0 dH_s^1 ds \\
&\leq C \int_0^t \int_{\mathbb{R}^{8d}} (\langle v \rangle + \langle w \rangle + \langle \tilde{v} \rangle + \langle \tilde{w} \rangle) |\psi\sigma - \tilde{\psi}\tilde{\sigma}| dH_s^0 dH_s^1 ds
\end{aligned}$$

where we have used  $|v^*(v, w, u) - v| + |v^*(\tilde{v}, \tilde{w}, u) - \tilde{v}| \leq C\langle u \rangle (\langle v \rangle + \langle w \rangle + \langle \tilde{v} \rangle + \langle \tilde{w} \rangle)$  in the last inequality. Using again  $\|S(-s)g\|_s \leq 1$  gives

$$\begin{aligned}
&S(-s)g(r, v^*(v, w, u)) - S(-s)g(\tilde{r}, v^*(\tilde{v}, \tilde{w}, u)) \\
&\leq |(r, v^*(v, w, u)) - (\tilde{r}, v^*(\tilde{v}, \tilde{w}, u))|_s \\
&= C|(r, v) - (\tilde{r}, \tilde{v})|_s + C|(r, w) - (\tilde{r}, \tilde{w})|_s,
\end{aligned}$$

and

$$S(-s)g(\tilde{r}, \tilde{v}) - S(-s)g(r, v) \leq |(r, v) - (\tilde{r}, \tilde{v})|_s.$$

Hence  $J_1$  is estimated by

$$\begin{aligned}
J_1 &\leq C \int_0^t \int_{\mathbb{R}^{9d}} (|(r, w) - (\tilde{r}, \tilde{w})|_s + |(r, v) - (\tilde{r}, \tilde{v})|_s) (\psi\sigma \wedge \tilde{\psi}\tilde{\sigma}) a(u) dudH_s^0 dH_s^1 ds \\
&\leq C \int_0^t \int_{\mathbb{R}^{8d}} (|(r, w) - (\tilde{r}, \tilde{w})|_s + |(r, v) - (\tilde{r}, \tilde{v})|_s) (\sigma \wedge \tilde{\sigma}) dH_s^0 dH_s^1 ds
\end{aligned}$$

which proves the assertion.  $\square$

The following gives the main estimate for this section.

**Theorem 7.2.** *Suppose that  $\int_{\mathbb{R}^{2d}} |u|a(u)du < \infty$  and assume that  $\psi, \sigma$  are globally Lipschitz continuous. Then for each  $\delta > 0$  and  $T > 0$  there exists a constant  $C = C(T, \delta, a, \psi, \sigma)$  such that for all  $\mu_0, \nu_0 \in \mathcal{P}(\mathbb{R}^{2d})$  any two solutions  $(\mu_t)_{t \geq 0}, (\nu_t)_{t \geq 0}$  to (14) satisfying*

$$\mathcal{C}_\gamma(T, \mu + \nu, \delta) = \sup_{t \in [0, T]} \int_{\mathbb{R}^{2d}} \left( e^{\delta|v|^{1+\gamma}} + |r|^{1+\delta} \right) d(\mu_t + \nu_t)(r, v) < \infty \quad (52)$$

it holds that

$$d_t(\mu_t, \nu_t) \leq d_0(\mu_0, \nu_0) + CC_\gamma(T, \mu + \nu, \delta)^2 \int_0^t d_s(\mu_s, \nu_s) (1 + |\ln(d_s(\mu_s, \nu_s))|) ds.$$

*Proof.* It is easily seen that Proposition 6 is applicable in this case. Let us start with the first term in  $\Lambda$ . Using the elementary inequality

$$c_{a,b}|x^{a+b} - y^{a+b}| \leq (x^a + y^a)|x^b - y^b| \leq C_{a,b}|x^{a+b} - y^{a+b}|, \quad x, y \geq 0, \quad a, b > 0$$

we obtain

$$\begin{aligned} & |\sigma(v-w)\psi(r-q) - \sigma(\tilde{v}-\tilde{w})\psi(\tilde{r}-\tilde{q})| \\ & \leq \sigma(v-w)|\psi(r-q) - \psi(\tilde{r}-\tilde{q})| + \psi(\tilde{r}-\tilde{q})|\sigma(v-w) - \sigma(\tilde{v}-\tilde{w})| \\ & \leq C(\langle v \rangle^\gamma + \langle w \rangle^\gamma)(|r-\tilde{r}| + |q-\tilde{q}|) + C(|v-\tilde{v}| + |w-\tilde{w}|) \end{aligned}$$

and hence

$$\begin{aligned} & (\langle v \rangle + \langle w \rangle + \langle \tilde{v} \rangle + \langle \tilde{w} \rangle)|\sigma(v-w)\psi(r-q) - \sigma(\tilde{v}-\tilde{w})\psi(\tilde{r}-\tilde{q})| \\ & \leq C(\langle v \rangle^{1+\gamma} + \langle w \rangle^{1+\gamma} + \langle \tilde{v} \rangle^{1+\gamma} + \langle \tilde{w} \rangle^{1+\gamma})(|r-\tilde{r}| + |q-\tilde{q}| + |v-\tilde{v}| + |w-\tilde{w}|) \\ & \leq C(\langle w \rangle^{1+\gamma} + \langle \tilde{w} \rangle^{1+\gamma})(|r-\tilde{r}| + |v-\tilde{v}|) \\ & \quad + C(\langle v \rangle^{1+\gamma} + \langle \tilde{v} \rangle^{1+\gamma})(|q-\tilde{q}| + |w-\tilde{w}|) \\ & \quad + C(\langle v \rangle^{1+\gamma} + \langle \tilde{v} \rangle^{1+\gamma})(|r-\tilde{r}| + |v-\tilde{v}|) \\ & \quad + C(\langle w \rangle^{1+\gamma} + \langle \tilde{w} \rangle^{1+\gamma})(|q-\tilde{q}| + |w-\tilde{w}|). \end{aligned}$$

Hence using that  $H_s^0, H_s^1 \in \mathcal{H}(\mu_s, \nu_s)$  we obtain

$$\begin{aligned} & \int_{\mathbb{R}^{8d}} (\langle v \rangle + \langle w \rangle + \langle \tilde{v} \rangle + \langle \tilde{w} \rangle)|\sigma(v-w)\psi(r-q) - \sigma(\tilde{v}-\tilde{w})\psi(\tilde{r}-\tilde{q})| dH_s^0 dH_s^1 \\ & \leq C(\|\mu_s\|_{1+\gamma} + \|\nu_s\|_{1+\gamma}) \int_{\mathbb{R}^{4d}} (|r-\tilde{r}| + |v-\tilde{v}|) dH_s(r, v; \tilde{r}, \tilde{v}) \\ & \quad + C \int_{\mathbb{R}^{4d}} (\langle v \rangle^{1+\gamma} + \langle \tilde{v} \rangle^{1+\gamma})(|r-\tilde{r}| + |v-\tilde{v}|) dH_s(r, v; \tilde{r}, \tilde{v}) \\ & \leq CC_\gamma(T, \mu + \nu, \delta) d_s(\mu_s, \nu_s) \\ & \quad + CC_\gamma(T, \mu + \nu, \delta) \int_{\mathbb{R}^{4d}} (\langle v \rangle^{1+\gamma} + \langle \tilde{v} \rangle^{1+\gamma}) |(r, v) - (\tilde{r}, \tilde{v})|_s dH_s(r, v; \tilde{r}, \tilde{v}) \\ & \leq CC_\gamma(T, \mu + \nu, \delta) d_s(\mu_s, \nu_s)(1 + |\ln(d_s(\mu_s, \nu_s))|), \end{aligned}$$

where we have used  $|r-\tilde{r}| + |v-\tilde{v}| \leq (1+T)|(r, v) - (\tilde{r}, \tilde{v})|_s$ , (51) and similar arguments to the proof of Theorem 7.1 (see also [13] and [11]) to obtain

$$\begin{aligned} & \int_{\mathbb{R}^{4d}} (\langle v \rangle^{1+\gamma} + \langle \tilde{v} \rangle^{1+\gamma}) |(r, v) - (\tilde{r}, \tilde{v})|_s dH_s(r, v; \tilde{r}, \tilde{v}) \\ & \leq CC_\gamma(T, \mu + \nu, \delta) d_s(\mu_s, \nu_s)(1 + |\ln(d_s(\mu_s, \nu_s))|). \end{aligned}$$

For the second term in  $\Lambda$  we use

$$\begin{aligned} |(r, w) - (\tilde{r}, \tilde{w})|_s & \leq |r-\tilde{r}| + (1+s)|w-\tilde{w}| \\ & \leq (1+T)|(r, v) - (\tilde{r}, \tilde{v})|_s + (1+T)|(q, w) - (\tilde{q}, \tilde{w})|_s \end{aligned}$$

combined with  $\min\{\sigma(v-w), \sigma(\tilde{v}-\tilde{w})\} \leq \sigma(v-w) \leq C\langle v \rangle^\gamma \langle w \rangle^\gamma$  to obtain

$$\begin{aligned} & \int_{\mathbb{R}^{8d}} (|(r, w) - (\tilde{r}, \tilde{w})|_s + |(r, v) - (\tilde{r}, \tilde{v})|_s) \min\{\sigma(v-w), \sigma(\tilde{v}-\tilde{w})\} dH_s^0 dH_s^1 \\ & \leq C \int_{\mathbb{R}^{8d}} (|(r, v) - (\tilde{r}, \tilde{v})|_s + |(q, w) - (\tilde{q}, \tilde{w})|_s) \min\{\sigma(v-w), \sigma(\tilde{v}-\tilde{w})\} dH_s^0 dH_s^1 \\ & \leq C(\|\mu_s\|_{1+\gamma} + \|\nu_s\|_{1+\gamma}) \int_{\mathbb{R}^{4d}} \langle v \rangle^\gamma |(r, v) - (\tilde{r}, \tilde{v})|_s dH_s(r, v; \tilde{r}, \tilde{v}) \\ & \leq CC_\gamma(T, \mu + \nu, \delta)^2 d_s(\mu_s, \nu_s)(1 + |\ln(d_s(\mu_s, \nu_s))|). \end{aligned}$$

Applying the general coupling inequality and then above estimates proves the assertion.  $\square$

**Remark 3.** Using again Lemma C.1 from the Appendix we may deduce from above estimate uniqueness and stability with respect to the Wasserstein distance.

**Appendix A. Proof of Lemma 2.4.** (a) Applying the Itô formula we obtain, for  $g \in C_c^1(\mathbb{R}^{2d})$ ,

$$g(R(t), V(t)) - g(R(0), V(0)) - \int_0^t (A(\mu_s)g)(R(s), V(s)) ds = M_g(t), \quad t \geq 0$$

where  $(M_g(t))_{t \geq 0}$  is a local martingale. It suffices to show that  $(M_g(t))_{t \geq 0}$  is, indeed, a martingale. For each  $g \in C_c^1(\mathbb{R}^{2d})$  we find  $C > 0$  with

$$|A(\mu_s)g(r, v)| \leq C \int_{\mathbb{R}^{2d}} \langle w \rangle^\gamma d\mu_s(q, w) \langle v \rangle^\gamma = C \|\mu_s\|_\gamma \langle v \rangle^\gamma.$$

This implies that

$$\begin{aligned} \mathbb{E} \left( \sup_{s \in [0, t]} |M_g(t)| \right) &\leq 2\|g\|_\infty + \int_0^t \mathbb{E}(|(A(\mu_s)g)(R(s), V(s))|) ds \\ &\leq 2\|g\|_\infty + C \int_0^t \|\mu_s\|_\gamma \mathbb{E}(\langle V(s) \rangle^\gamma) ds \\ &\leq 2\|g\|_\infty + t \sup_{s \in [0, t]} \|\mu_s\|_\gamma^2 < \infty, \end{aligned}$$

i.e.  $(M_g(t))_{t \geq 0}$  is a martingale (see e.g. [25, Theorem 46, p.36]).

(b) Let  $(q_t, w_t)$  be a measurable process defined on  $([0, 1], \mathcal{B}([0, 1]), d\eta)$  such that  $(q_t, w_t)$  has law  $\mu_t$ , for all  $t \geq 0$ , where  $\mu_t$  denotes the time-marginal of  $\mu$ . Using [18, Theorem A.1] gives the existence of a weak solution  $(R, V)$  to (17) such that  $(R, V)$  has law  $\mu$ .

**Appendix B. Fractional binomial expansion.** The following lemma is due to [3, Lemma 2].

**Lemma B.1.** Let  $p \geq 1$  and  $k_p = \lfloor \frac{p+1}{2} \rfloor$  where  $\lfloor x \rfloor \in \mathbb{Z}$  is defined by  $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$ , set  $\binom{p}{l} = \frac{p(p-1)\cdots(p-l+1)}{l!}$ , for  $l \geq 1$ , and  $\binom{p}{0} = 1$ . Then for all  $x, y \geq 0$

$$\sum_{k=0}^{k_p-1} \binom{p}{k} (x^k y^{p-k} + x^{p-k} y^k) \leq (x+y)^p \leq \sum_{k=0}^{k_p} \binom{p}{k} (x^k y^{p-k} + x^{p-k} y^k).$$

**Appendix C. Some variants of the Gronwall lemma.** We need the following generalization of the Gronwall inequality (see [4, Lemma 5.2.1, p. 89]) for a proof).

**Lemma C.1.** Let  $\rho$  be a nonnegative bounded function on  $[0, T]$ ,  $a \in [0, \infty)$  and  $g$  be a strictly positive and non-decreasing function on  $(0, \infty)$ . Suppose that  $\int_0^1 \frac{dx}{g(x)} = \infty$  and

$$\rho(t) \leq a + \int_0^t g(\rho(s)) ds, \quad t \in [0, T].$$

Then

- (a) If  $a = 0$ , then  $\rho(t) = 0$  for all  $t \in [0, T]$ .
- (b) If  $a > 0$ , then  $G(a) - G(\rho(t)) \leq t$  where  $G(x) = \int_x^1 \frac{dy}{g(y)}$ .

The following nonlinear generalization of the Gronwall lemma is a particular case of the Bihari-LaSalle inequality.

**Lemma C.2.** *Let  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a locally bounded function satisfying*

$$f(t) \leq f(0) + K \int_0^t f(s)^{1-\alpha} ds, \quad t \geq 0$$

for some  $K \geq 0$  and  $\alpha \in (0, 1)$ . Then for any  $t \geq 0$

$$f(t) \leq (f(0)^\alpha + \alpha K t)^{1/\alpha} \leq 2^{1/\alpha-1} f(0) + \frac{(2\alpha K)^{1/\alpha}}{2} t^{1/\alpha}.$$

**Appendix D. Some localization result.** Let  $(E, \rho)$  be a complete, separable metric space. Let  $A \subset C_b(E) \times C(E)$  be a (multi-valued) operator such that there exists  $1 \leq \psi \in C(E)$  with

$$|g| \leq K_f \psi, \quad \forall (f, g) \in A \quad (53)$$

for some  $K_f > 0$ . Set  $\mathcal{P}_\psi := \{\mu \in \mathcal{P}(E) \mid \int_E \psi(x) d\mu(x) < \infty\}$ . Here and below  $D(\mathbb{R}_+; E)$  denotes the Skorokhod space and  $x$  the canonical process on  $D(\mathbb{R}_+; E)$ .

**Definition D.1.** Let  $\mu \in \mathcal{P}_\psi$ . A solution to the martingale problem  $(A, \mu)$  is a probability measure  $\mathbb{P}_\mu$  on  $D(\mathbb{R}_+; E)$  such that

- (a)  $\mathbb{P}_\mu(x(0) \in F) = \mu(F)$  for all  $F \in \mathcal{B}(E)$ .
- (b)  $\int_0^T \mathbb{E}_\mu(\psi(x(t))) dt < \infty$  for all  $T > 0$ .
- (c) For all  $(f, g) \in A$

$$f(x(t)) - f(x(0)) - \int_0^t g(x(s)) ds, \quad t \geq 0 \quad (54)$$

is a martingale w.r.t.  $\mathbb{P}_\mu$ .

When working with martingale problems the use of localization techniques such as [8, Theorem 6.3, Corollary 6.4] is essential. However, the statements therein require that  $A \subset C_b(E) \times B(E)$ , i.e.  $\psi = 1$ . Below we give one possible extension.

**Theorem D.2.** *Let  $A \subset C_b(E) \times C(E)$  satisfy (53) and  $A_m \subset C_b(E) \times C(E)$  be such that  $|g_m| \leq K_f \psi$  holds for  $(f, g_m) \in A_m$  with a constant  $K_f > 0$  independent of  $m \geq 1$ . Suppose that there exists  $\mu \in \mathcal{P}_\psi$  such that the following conditions hold:*

- (i) *There exist open sets  $(U_m)_{m \geq 1}$  with  $\bar{U}_m \subset U_{m+1}$ ,  $\bigcup_{m \geq 1} U_m = E$  and*

$$\{(f, \mathbb{1}_{U_m} g) \mid (f, g) \in A_m\} = \{(f, \mathbb{1}_{U_m} g) \mid (f, g) \in A\}, \quad m \geq 1.$$

Moreover  $\mathbb{1}_{U_m} \psi$  is bounded for any  $m \geq 1$ .

- (ii) *The martingale problem  $(A_m, \rho)$  has for each  $\rho \in \mathcal{P}(E)$  and each  $m \geq 1$  a unique solution.*
- (iii) *We have*

$$\lim_{k \rightarrow \infty} \sup_{m \geq k} \mathbb{P}_\mu^m(\tau_k \leq T) = 0, \quad \forall T > 0$$

where  $\mathbb{P}_\mu^m$  is the unique solution to the martingale problem  $(A_m, \mu)$  and

$$\tau_k = \inf\{t > 0 \mid x(t) \notin U_k \text{ or } x(t-) \notin U_k\}$$

is a stopping time on  $D(\mathbb{R}_+; E)$ .

- (iv) *There exists  $p > 1$  such that for all  $T > 0$  there exists  $C(p, T) > 0$  satisfying*

$$\sup_{m \geq 1} \sup_{t \in [0, T]} \mathbb{E}_\mu^m(\psi(x(t))^p) \leq C(p, T),$$

where  $\mathbb{E}_\mu^m$  denotes the expectation w.r.t.  $\mathbb{P}_\mu^m$ .

Then there exists a unique solution  $\mathbb{P}_\mu$  to the martingale problem  $(A, \mu)$ . This solution satisfies

$$\sup_{t \in [0, T]} \mathbb{E}_\mu(\psi(x(t))^p) \leq C(p, T), \quad T > 0.$$

**Remark 4.** In several cases one may take  $U_m = \{x \in E \mid \psi(x) < m\}$ . In such a case condition (iii) is implied by

$$\lim_{k \rightarrow \infty} \sup_{m \geq k} \mathbb{P}_\mu^m \left( \sup_{t \in [0, T]} \psi(x(t)) \geq k \right) = 0, \quad \forall T > 0$$

or the stronger condition

$$\sup_{m \geq 1} \mathbb{E}_\mu^m \left( \sup_{t \in [0, T]} \psi(x(t)) \right) < \infty, \quad \forall T > 0.$$

*Proof. Step 1.* Let  $n \geq 1$ ,  $0 \leq t_1, \dots, t_n \leq T$  and  $H \in C_b(E^n)$ . Then (i), (ii) together with [8, Chapter 4, Theorem 6.1] yield for  $1 \leq k \leq m$

$$\mathbb{E}_\mu^m (\mathbb{1}_{\{\tau_k > T\}} H(x(t_1), \dots, x(t_n))) = \mathbb{E}_\mu^k (\mathbb{1}_{\{\tau_k > T\}} H(x(t_1), \dots, x(t_n))).$$

*Step 2.* Let us prove that  $\mathbb{P}_\mu^m \rightarrow \mathbb{P}_\mu$  weakly in  $\mathcal{P}(D(\mathbb{R}_+; E))$ . For this purpose we metrize the topology on  $\mathcal{P}(D(\mathbb{R}_+; E))$  with respect to a complete metric and prove that  $(\mathbb{P}_\mu^m)_{m \geq 1}$  is a Cauchy sequence in this metric.

Recall that the topology on  $D(\mathbb{R}_+; E)$  may be obtained from the metric

$$d(x, y) = \inf_{\lambda \in \Lambda} \left( \gamma(\lambda) \vee \int_0^\infty e^{-u} \sup_{t \geq 0} q(x(t \wedge u), y(\lambda(t) \wedge u)) du \right)$$

where  $q := \rho \wedge 1$ ,  $\gamma(\lambda) := \sup_{0 \leq s < t} \left| \log \left( \frac{\lambda(t) - \lambda(s)}{t - s} \right) \right|$  and  $\Lambda$  is the set of all strictly increasing, surjective, Lipschitz continuous functions  $\lambda : [0, \infty) \rightarrow [0, \infty)$  with  $\gamma(\lambda) < \infty$  (see [8, p.117]). For  $H : D(\mathbb{R}_+; E) \rightarrow \mathbb{R}$  let

$$\|H\|_{BL} = \|H\|_\infty + \sup_{x \neq y} \frac{|H(x) - H(y)|}{d(x, y)}.$$

Then it suffices to prove that  $(\mathbb{P}_\mu^m)_{m \geq 1} \subset \mathcal{P}(D(\mathbb{R}_+; E))$  is a Cauchy sequence w.r.t. the metric

$$d_{BL}(P, Q) = \sup_{\|H\|_{BL} \leq 1} \left| \int_{D(\mathbb{R}_+; E)} H(x) dP(x) - \int_{D(\mathbb{R}_+; E)} H(x) dQ(x) \right|.$$

Take  $H$  with  $\|H\|_{BL} \leq 1$ ,  $T > 0$ ,  $1 \leq k < m$  and set  $x^T := x(\cdot \wedge T)$ ,  $H^T(x) := H(x^T)$ . Then

$$\begin{aligned} & |\mathbb{E}_\mu^m(H) - \mathbb{E}_\mu^k(H)| \\ & \leq |\mathbb{E}_\mu^m(H^T) - \mathbb{E}_\mu^m(H)| + |\mathbb{E}_\mu^m(H^T) - \mathbb{E}_\mu^k(H^T)| + |\mathbb{E}_\mu^k(H^T) - \mathbb{E}_\mu^k(H)| \\ & =: I_1 + I_2 + I_3. \end{aligned}$$

Then by Step 1 and  $\mathbb{1}_{\{\tau_m > T\}} \geq \mathbb{1}_{\{\tau_k > T\}}$  we get

$$I_2 = \mathbb{E}_\mu^m(\mathbb{1}_{\{\tau_m \leq T\}} H^T) - \mathbb{E}_\mu^k(\mathbb{1}_{\{\tau_k \leq T\}} H^T) \leq \|H\|_\infty (\mathbb{P}_\mu^m(\tau_m \leq T) + \mathbb{P}_\mu^k(\tau_k \leq T))$$

which tends by (iii) to zero. Moreover we have

$$\begin{aligned} I_1 & = |\mathbb{E}_\mu^m(H^T) - \mathbb{E}_\mu^m(H)| \leq \mathbb{E}_\mu^m(d(x^T, x)) \\ & \leq \mathbb{E}_\mu^m \left( \int_0^\infty e^{-u} \sup_{t \geq 0} q(x(t \wedge u \wedge T), x(t \wedge u)) du \right) \leq e^{-T} \end{aligned}$$

and likewise  $I_3 \leq e^{-T}$  which completes Step 2.

*Step 3.* Let  $\mathbb{P}_\mu$  be the limit of  $\mathbb{P}_\mu^m$ . Using (iv), monotone convergence and the Lemma of Fatou one can show that

$$\sup_{t \in [0, T]} \mathbb{E}_\mu(\psi(x(t))^p) \leq C(p, T), \quad T > 0.$$

*Step 4.* Take  $g \in C(E)$  such that there exists  $K_g > 0$  with  $|g| \leq K_g \psi$ . We show that

$$\lim_{m \rightarrow \infty} \mathbb{E}_\mu^m(g(x(t))) = \mathbb{E}_\mu(g(x(t))), \quad t \in D_\mu$$

where  $D_\mu = \{t \geq 0 \mid \mathbb{P}_\mu(x(t) = x(t-)) = 1\}$ . Note that  $D_\mu^c$  is at most countable.

Let  $h_k \in C_b(E)$  be such that  $\mathbb{1}_{U_k} \leq h_k \leq \mathbb{1}_{U_{k+1}}$ ,  $k \geq 1$ . Then for  $k < m$

$$\begin{aligned} & |\mathbb{E}_\mu^m(g(x(t))) - \mathbb{E}_\mu(g(x(t)))| \\ & \leq |\mathbb{E}_\mu^m(h_k(x(t))g(x(t))) - \mathbb{E}_\mu(h_k(x(t))g(x(t)))| \\ & \quad + |\mathbb{E}_\mu^m((1 - h_k(x(t)))g(x(t)))| + |\mathbb{E}_\mu((1 - h_k(x(t)))g(x(t)))| \\ & = I_1 + I_2 + I_3. \end{aligned}$$

It suffices to show that

$$\begin{aligned} \lim_{m \rightarrow \infty} I_1 &= 0, \quad \forall k \geq 1 \\ \lim_{k \rightarrow \infty} \sup_{m \geq k} (I_2 + I_3) &= 0. \end{aligned}$$

Concerning  $I_1$  the assertion follows by Step 2 and since  $x \mapsto h_k(x(t))g(x(t))$  is bounded and  $\mathbb{P}_\mu$ -a.s. continuous on  $D(\mathbb{R}_+; E)$  for any  $k \geq 1$ . For the second property we use  $\mathbb{1}_{U_k}(1 - h_k) = 0$  so that

$$\begin{aligned} I_2 + I_3 &= |\mathbb{E}_\mu^m(\mathbb{1}_{\{\tau_k \leq t\}}(1 - h_k(x(t)))g(x(t)))| + |\mathbb{E}_\mu(\mathbb{1}_{\{\tau_k \leq t\}}(1 - h_k(x(t)))g(x(t)))| \\ & \leq K_g \mathbb{E}_\mu^m(\mathbb{1}_{\{\tau_k \leq t\}}\psi(x(t))) + K_g \mathbb{E}_\mu(\mathbb{1}_{\{\tau_k \leq t\}}\psi(x(t))) \\ & \leq K_g (\mathbb{P}_\mu^m(\tau_k \leq t))^{1 - \frac{1}{p}} \mathbb{E}_\mu^m(\psi(x(t))^p)^{\frac{1}{p}} + K_g (\mathbb{P}_\mu(\tau_k \leq t))^{1 - \frac{1}{p}} \mathbb{E}_\mu(\psi(x(t))^p)^{\frac{1}{p}}. \end{aligned}$$

For the first term we can use (iii) and (iv); for the second term this follows from  $\mathbb{P}_\mu \in \mathcal{P}(D(\mathbb{R}_+; E))$ .

*Step 5.*  $\mathbb{P}_\mu$  is a solution for the martingale problem for  $(A, \mu)$ .

Fix  $n \geq 1$ ,  $0 \leq t_1, \dots, t_n \leq s < t$  in  $D_\mu$ ,  $h_1, \dots, h_n \in C_b(E)$ ,  $(f, g) \in A$  and set

$$H := \left( f(x(t)) - f(x(s)) - \int_s^t g(x(s)) ds \right) \prod_{k=1}^n h_k(x(t_k)). \quad (55)$$

We have to show that  $\mathbb{E}_\mu(H) = 0$ . First using Steps 3 and 4 together with (iv) and dominated convergence we easily deduce

$$\mathbb{E}_\mu(H) = \lim_{m \rightarrow \infty} \mathbb{E}_\mu^m(H) = \lim_{m \rightarrow \infty} \mathbb{E}_\mu^m(\mathbb{1}_{\tau_m \leq T} H) + \lim_{m \rightarrow \infty} \mathbb{E}_\mu^m(\mathbb{1}_{\tau_m > T} H)$$

where  $t < T$ . We can find a constant  $C > 0$  such that

$$|\mathbb{E}_\mu^m(\mathbb{1}_{\tau_m \leq T} H)| \leq C \mathbb{P}_\mu^m(\tau_m \leq T) + C \mathbb{P}_\mu^m(\tau_m \leq T)^{1 - \frac{1}{p}} \sup_{t \in [0, T]} \mathbb{E}_\mu^m(\psi(x(t))^p)^{\frac{1}{p}} \quad (56)$$

and the right-hand side tends to zero as  $m \rightarrow \infty$ . Since  $(f, g) \in A$  we can find by (i)  $g_m \in C_b(E)$  such that  $(f, g_m) \in A_m$  and  $\mathbb{1}_{U_m} g = \mathbb{1}_{U_m} g_m$ . Let  $H_m$  be given by (55) with  $g$  replaced by  $g_m$ . Then, since  $\mathbb{P}_\mu^m$  is a solution to the martingale problem  $(A_m, \mu)$ , it follows  $\mathbb{E}_\mu^m(H_m) = 0$  and hence

$$\mathbb{E}_\mu^m(\mathbb{1}_{\tau_m > T} H) = \mathbb{E}_\mu^m(\mathbb{1}_{\tau_m > T} H_m) = -\mathbb{E}_\mu^m(\mathbb{1}_{\tau_m \leq T} H_m).$$

Since  $|g_m| \leq C\psi$  for some  $C > 0$  independent of  $m$ , the latter expression can be estimated in the same way as (56).

*Step 6.* It remains to show that there exists only one solution to the martingale problem  $(A, \mu)$ . Let  $\mathbb{P}'_\mu \in \mathcal{P}(D(\mathbb{R}_+; E))$  be any solution to the martingale problem  $(A, \mu)$ . Let  $n \geq 1$ ,  $0 \leq t_1, \dots, t_n \leq T$  and  $H \in C_b(E^n)$ . Then (i) and (ii) imply

$$\mathbb{E}_\mu^m(\mathbb{1}_{\{\tau_m > T\}}H(x(t_1), \dots, x(t_n))) = \mathbb{E}'_\mu(\mathbb{1}_{\{\tau_m > T\}}H(x(t_1), \dots, x(t_n))).$$

The assertion now follows from the identity

$$\begin{aligned} & \mathbb{E}_\mu^m(H(x(t_1), \dots, x(t_n))) - \mathbb{E}'_\mu(H(x(t_1), \dots, x(t_n))) \\ &= \mathbb{E}_\mu^m(\mathbb{1}_{\{\tau_m \leq T\}}H(x(t_1), \dots, x(t_n))) - \mathbb{E}'_\mu(\mathbb{1}_{\{\tau_m \leq T\}}H(x(t_1), \dots, x(t_n))) \end{aligned}$$

after taking the limit  $m \rightarrow \infty$ . □

## REFERENCES

- [1] S. M. Ahn and S.-Y. Ha, Stochastic flocking dynamics of the Cucker-Smale model with multiplicative white noises, *J. Math. Phys.*, **51** (2010), 103301, 17, URL <https://doi.org/10.1063/1.3496895>.
- [2] S. Albeverio, B. Rüdiger and P. Sundar, The Enskog process, *J. Stat. Phys.*, **167** (2017), 90–122, URL <https://doi.org/10.1007/s10955-017-1743-9>.
- [3] A. V. Bobylev, I. M. Gamba and V. A. Panferov Panferov, Moment inequalities and high-energy tails for Boltzmann equations with inelastic interactions, *J. Statist. Phys.*, **116** (2004), 1651–1682, URL <https://doi.org/10.1023/B:J0SS.0000041751.11664.ea>.
- [4] J.-Y. Chemin, Fluides parfaits incompressibles, *Astérisque*, 177.
- [5] C.-C. Chen, S.-Y. Ha and X. Zhang, The global well-posedness of the kinetic Cucker-Smale flocking model with chemotactic movements, *Commun. Pure Appl. Anal.*, **17** (2018), 505–538, URL <https://doi.org/doi:10.3934/cpaa.2018028>.
- [6] F. Cucker and S. Smale, Emergent behavior in flocks, *IEEE Trans. Automat. Control*, **52** (2007), 852–862, URL <http://dx.doi.org/10.1109/TAC.2007.895842>.
- [7] F. Cucker and S. Smale, On the mathematics of emergence, *Jpn. J. Math.*, **2** (2007), 197–227, URL <http://dx.doi.org/10.1007/s11537-007-0647-x>.
- [8] S. Ethier and T. G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiles & Sons, Inc., New York, 1986, Characterization and convergence.
- [9] N. Fournier, Finiteness of entropy for the homogeneous Boltzmann equation with measure initial condition, *Ann. Appl. Probab.*, **25** (2015), 860–897, URL <http://dx.doi.org/10.1214/14-AAP1012>.
- [10] N. Fournier and S. Mischler, Rate of convergence of the Nanbu particle system for hard potentials and Maxwell molecules, *Ann. Probab.*, **44** (2016), 589–627, URL <https://doi.org/10.1214/14-AOP983>.
- [11] N. Fournier and C. Mouhot, On the well-posedness of the spatially homogeneous Boltzmann equation with a moderate angular singularity, *Comm. Math. Phys.*, **289** (2009), 803–824, URL <http://dx.doi.org/10.1007/s00220-009-0807-3>.
- [12] M. Friesen, B. Rüdiger and P. Sundar, The Enskog process for hard and soft potentials, *NoDEA Nonlinear Differential Equations Appl.*, **26** (2019), Art. 20, 42, URL <https://doi.org/10.1007/s00030-019-0566-6>.
- [13] M. Friesen, B. Rüdiger and P. Sundar, On uniqueness for the Enskog process for hard and soft potentials, (*to appear*).
- [14] S.-Y. Ha, J. Jeong, S. E. Noh, Q. Xiao and X. Zhang, Emergent dynamics of Cucker-Smale flocking particles in a random environment, *J. Differential Equations*, **262** (2017), 2554–2591, URL <https://doi.org/10.1016/j.jde.2016.11.017>.
- [15] S.-Y. Ha and D. Levy, Particle, kinetic and fluid models for phototaxis, *Discrete Contin. Dyn. Syst. Ser. B*, **12** (2009), 77–108, URL <http://dx.doi.org/10.3934/dcdsb.2009.12.77>.
- [16] S.-Y. Ha and J.-G. Liu, A simple proof of the Cucker-Smale flocking dynamics and mean-field limit, *Commun. Math. Sci.*, **7** (2009), 297–325, URL <http://projecteuclid.org/euclid.cms/1243443982>.

- [17] S.-Y. Ha and E. Tadmor, From particle to kinetic and hydrodynamic descriptions of flocking, *Kinet. Relat. Models*, **1** (2008), 415–435, URL <http://dx.doi.org/10.3934/krm.2008.1.415>.
- [18] J. Horowitz and R. L. Karandikar, Martingale problems associated with the Boltzmann equation, in *Seminar on Stochastic Processes, 1989 (San Diego, CA, 1989)*, vol. 18 of Progr. Probab., Birkhäuser Boston, Boston, MA, 1990, 75–122.
- [19] P.-E. Jabin and Z. Wang, Quantitative estimates of propagation of chaos for stochastic systems with  $W^{-1,\infty}$  kernels, *Invent. Math.*, **214** (2018), 523–591, URL <https://doi.org/10.1007/s00222-018-0808-y>.
- [20] J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, vol. 288 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 2nd edition, Springer-Verlag, Berlin, 2003, URL <http://dx.doi.org/10.1007/978-3-662-05265-5>.
- [21] T. G. Kurtz, Equivalence of stochastic equations and martingale problems, in *Stochastic analysis 2010*, Springer Heidelberg, 2011, 113–130.
- [22] S. Mischler and C. Mouhot, Kac’s program in kinetic theory, *Invent. Math.*, **193** (2013), 1–147, URL <https://doi.org/10.1007/s00222-012-0422-3>.
- [23] P. B. Mucha and J. Peszek, The Cucker-Smale equation: singular communication weight, measure-valued solutions and weak-atomic uniqueness, *Arch. Ration. Mech. Anal.*, **227** (2018), 273–308, URL <https://doi.org/10.1007/s00205-017-1160-x>.
- [24] B. Piccoli, F. Rossi and E. Trélat, Control to flocking of the kinetic Cucker-Smale model, *SIAM J. Math. Anal.*, **47** (2015), 4685–4719, URL <https://doi.org/10.1137/140996501>.
- [25] P. E. Protter, *Stochastic integration and differential equations*, vol. 21 of Stochastic Modelling and Applied Probability, Springer-Verlag, Berlin, 2005, URL <https://doi.org/10.1007/978-3-662-10061-5>, Second edition. Version 2.1, Corrected third printing.
- [26] S. Serfaty and M. Duerinckx, Mean field limit for coulomb-type flows, *arXiv:1803.08345 [math.AP]*.
- [27] J. Shen, Cucker-Smale flocking under hierarchical leadership, *SIAM J. Appl. Math.*, **68** (2007/08), 694–719, URL <http://dx.doi.org/10.1137/060673254>.
- [28] A.-S. Sznitman, Topics in propagation of chaos, in *École d’Été de Probabilités de Saint-Flour XIX—1989*, vol. 1464 of Lecture Notes in Math., Springer, Berlin, 1991, 165–251, URL <http://dx.doi.org/10.1007/BFb0085169>.
- [29] H. Tanaka, Probabilistic treatment of the Boltzmann equation of Maxwellian molecules, *Z. Wahrsch. Verw. Gebiete*, **46** (1978/79), 67–105, URL <http://dx.doi.org/10.1007/BF00535689>.
- [30] C. Villani, *Optimal transport*, vol. 338 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Springer-Verlag, Berlin, 2009, URL <http://dx.doi.org/10.1007/978-3-540-71050-9>, Old and new.

Received xxxx 20xx; revised xxxx 20xx.

*E-mail address:* [friesen@math.uni-wuppertal.de](mailto:friesen@math.uni-wuppertal.de)

*E-mail address:* [kutoviy@math.uni-bielefeld.de](mailto:kutoviy@math.uni-bielefeld.de)