

# Macroscopic limits for stochastic partial differential equations of McKean-Vlasov type

Peter M. Kotelenez · Thomas G. Kurtz

Received: date / Accepted: date

**Abstract** A class of quasilinear stochastic partial differential equations (SPDEs), driven by spatially correlated Brownian noise, is shown to become macroscopic (i.e., deterministic), as the length of the correlations tends to 0. The limit is the solution of a quasilinear partial differential equation. The quasilinear SPDEs are obtained as a continuum limit from the empirical distribution of a large number of stochastic ordinary differential equations (SODEs), coupled through a mean-field interaction and driven by correlated Brownian noise. The limit theorems are obtained by application of a general result on the convergence of exchangeable systems of processes. We also compare our approach to SODEs with the one introduced by Kunita.

**Keywords** Stochastic partial differential equations · partial differential equations · macroscopic limit · particle systems · stochastic ordinary differential equations · exchangeable sequences

**Mathematics Subject Classification (2000)** 60H15 · 60F17 · 60J60 · 60G09 · 60K40

## 1 Introduction

Let  $N$  point particles be distributed over  $\mathbf{R}^d$ ,  $d \in \mathbf{N}$ . The position of the  $i$ -th particle at time  $t$  will be denoted  $r^i(t)$  and its mass  $m_i$ . Then the empirical mass distribution (also called the “empirical process”) at time  $t$  is given by  $X_N(t) := \sum_{i=1}^N m_i \delta_{r^i(t)}$ , where  $\delta_r$

---

This research was partially supported by NSF grant DMS 05-03983

P. M. Kotelenez  
Department of Mathematics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106  
E-mail: pxk4@cwru.edu

T. G. Kurtz  
Department of Mathematics, University of Wisconsin, 480 Lincoln Drive, Madison, WI 53706-1388  
E-mail: kurtz@math.wisc.edu

is the unit measure concentrated at  $r$ . In what follows we will choose  $m_i = \frac{1}{N}$  for  $i = 1, \dots, N$ .

Let  $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$  be a stochastic basis with right continuous filtration. All our stochastic processes are assumed to live on  $\Omega$  and be  $\mathcal{F}_t$ -adapted (including all initial conditions for the SODEs and SPDEs). The stochastic component of the displacement of  $r^i(t)$  in a short time increment should be Brownian (multiplied by some diffusion coefficient, which may depend both on  $r^i(t)$  and on  $X_N(t)$ ).

Following Kotelenetz (1995), we employ i.i.d. Gaussian standard white noise random fields  $w_l(dq, dt)$  on  $\mathbf{R}^d \times \mathbf{R}_+$ ,  $l = 1, \dots, d$ , as a stochastic perturbation for the positions of the particles. We list some of the properties of  $w_l(dr, dt)$ .

Let  $\mathcal{B}^d$  denote the Borel sets in  $\mathbf{R}^d$ , and for  $B \in \mathcal{B}^d$ , let  $|B|$  denote its Lebesgue measure. For  $i = 1, 2$ , let  $A_i \in \mathcal{B}^d$  with  $|A_i| < \infty$  and  $0 \leq s_i < t_i$ . Then  $w_l(A_i \times (s_i, t_i]) = \int_{s_i}^{t_i} \int_{A_i} w_l(dq, du)$ ,  $i = 1, 2$ , are Gaussian random variables with mean 0 and covariance  $|A_1 \cap A_2| |(s_1, t_1] \cap (s_2, t_2]|$ . For any  $A \in \mathcal{B}^d$  with  $|A| < \infty$ ,  $w_l(A, t) := w_l(A \times [0, t])$  defines a Brownian motion with variance parameter  $|A|$  adapted to  $\{\mathcal{F}_t\}$ . Stochastic integration with respect to  $w_l(dq, dt)$  follows the pattern established by Walsh (1986), where  $w_l(dq, dt)$  was used as a driving term for stochastic partial differential equations (SPDEs). Let  $\mathcal{M}_{d \times d}$  denote the  $d \times d$  matrices over  $\mathbf{R}$  with the usual matrix norm

$$|C| = \sup_{y, z \in \mathbf{R}^d} \frac{|y^T C x|}{|y| |x|}.$$

Let  $\mathcal{J}_\varepsilon(r, q, \mu, t)$  be a ‘‘nice’’  $\mathcal{M}_{d \times d}$ -valued function, jointly measurable in all arguments, depending on the position of the particle  $r$ , the spatial noise coordinate  $q$ , the empirical distribution  $\mu$ , time  $t$  and a correlation parameter  $\varepsilon > 0$ . (See Examples 1.3–1.5.) In addition to Lipschitz and measurability assumptions, ‘‘nice’’ means that the one-dimensional components of  $\mathcal{J}_\varepsilon(r, q, \mu, t)$  are square-integrable in  $q$  with respect to Lebesgue measure  $dq$ . Similar conditions are assumed for the one-dimensional components of the  $\mathbf{R}^d$ -valued function  $F_\varepsilon$ . Consider the following system of SODEs driven by  $w(dq, dt)$ :

$$\begin{aligned} dr_{\varepsilon, N}^i(t) &= F_\varepsilon(r_{\varepsilon, N}^i(t), X_{\varepsilon, N}(t), t) dt + \int \mathcal{J}_\varepsilon(r_{\varepsilon, N}^i(t), q, X_{\varepsilon, N}(t), t) w(dq, dt), \\ & \hspace{15em} (1.1) \\ r_{\varepsilon, N}^i(0) &= q_\varepsilon^i, i = 1, \dots, N, \quad X_{\varepsilon, N}(t) := \sum_{i=1}^N \frac{1}{N} \delta_{r_{\varepsilon, N}^i(t)}. \end{aligned}$$

The integration in (1.1) is taken over  $\mathbf{R}^d$ , and also in what follows, we will not indicate the integration domain, if it is  $\mathbf{R}^d$ . Under appropriate Lipschitz conditions (2.2), Kotelenetz (1995) shows that (1.1) has a unique strong Itô solution that is an  $\mathbf{R}^{dN}$ -valued diffusion process. The two-particle<sup>1</sup> and one-particle diffusion matrices

<sup>1</sup> The two-particle diffusion matrix, describing the pair correlations of the noise perturbations, is the time derivative of the mutual (tensor) quadratic variation of the noise.

are given by

$$\begin{aligned}\tilde{D}_\varepsilon(r^i, r^j, \mu, t) &:= \int \mathcal{J}_\varepsilon(r^i, q, \mu, t) \mathcal{J}_\varepsilon^T(r^j, q, \mu, t) dq \quad \forall i, j = 1, \dots, N, \\ D_\varepsilon(r, \mu, t) &:= \tilde{D}_\varepsilon(r, r, \mu, t),\end{aligned}\tag{1.2}$$

where  $C^T$  denotes the transpose of a matrix  $C$ . The entries of a matrix  $A$  will be denoted by  $A_{kl}$ .

Let  $\varphi$  be a twice continuously differentiable function on  $\mathbf{R}^d$ , and define

$$\mathcal{L}_\varepsilon \varphi(r, \mu, t) = F_\varepsilon(r, \mu, t) \cdot \nabla \varphi(r) + \frac{1}{2} \sum_{kl} D_{\varepsilon kl}(r, \mu, t) \partial_k \partial_l \varphi(r).$$

Let  $\langle \cdot, \cdot \rangle$  denote the duality between measures and continuous functions, that is,  $\langle \mu, \varphi \rangle = \int \varphi d\mu$ . Then (see Kotelenetz (1995), Kurtz and Protter (1996), and Kurtz and Xiong (1999)), under modest assumptions on the coefficients in (1.1), Itô's formula yields

$$\begin{aligned}\langle X_{\varepsilon, N}(t), \varphi \rangle &= \langle X_{\varepsilon, N}(0), \varphi \rangle + \int_0^t \langle X_{\varepsilon, N}(s), \mathcal{L}_\varepsilon \varphi(\cdot, X_{\varepsilon, N}(s), s) \rangle ds \\ &\quad + \int_{\mathbf{R}^d \times [0, t]} \langle X_{\varepsilon, N}(s), \nabla \varphi(\cdot)^T \mathcal{J}_\varepsilon(\cdot, q, X_{\varepsilon, N}(s), s) \rangle w(dq, ds)\end{aligned}\tag{1.3}$$

with quadratic variation given by

$$\begin{aligned}[\langle X_{\varepsilon, N}, \varphi \rangle]_t & \\ &= \int_0^t \sum_{i, j=1}^N \frac{1}{N^2} \sum_{k, l=1}^d (\partial_k \varphi)(r_{\varepsilon, N}^i(s)) (\partial_l \varphi)(r_{\varepsilon, N}^j(s)) \tilde{D}_{\varepsilon, k, l}(r_{\varepsilon, N}^i, r_{\varepsilon, N}^j, X_{\varepsilon, N}(s), s) ds.\end{aligned}\tag{1.4}$$

Due to the spatial correlations, we have  $N^2$  terms in the sum (1.4), divided by  $N^2$ , and the noise does not disappear in the limit as  $N \rightarrow \infty$ . Hence, we cannot expect a deterministic limit for  $X_{\varepsilon, N}(t)$ , as  $N \rightarrow \infty$  as long as  $\varepsilon > 0$  remains fixed. Moreover, the form of quasilinear SPDE in (1.3) does not depend on  $N$ , but can always be written as

$$\begin{aligned}\langle X_\varepsilon(t), \varphi \rangle &= \langle X_\varepsilon(0), \varphi \rangle + \int_0^t \langle X_\varepsilon(s), \mathcal{L}_\varepsilon \varphi(\cdot, X_\varepsilon(s), s) \rangle ds \\ &\quad + \int_{\mathbf{R}^d \times [0, t]} \langle X_\varepsilon(s), \nabla \varphi(\cdot)^T \mathcal{J}_\varepsilon(\cdot, q, X_\varepsilon(s), s) \rangle w(dq, ds).\end{aligned}\tag{1.5}$$

The solution of this equation can be extended by continuity (pass to the limit as  $N \rightarrow \infty$ ) to initial conditions given by any finite measure. See Kotelenetz (1995) and Kurtz and Xiong (1999). Typically, if the initial condition has a Lebesgue density, the solution will have a Lebesgue density for all time.

We next introduce an alternative and more traditional model with spatially uncorrelated Brownian noise, following Oelschläger (1984) and Gärtner (1988). To this end, choose a sequence of i.i.d.  $\mathbf{R}^d$ -valued standard Brownian motions  $\{\beta_i\}_{i \in \mathbf{N}}$ . Let

$\mathcal{J}_0(r, \mu, t)$  be a “nice”  $\mathcal{M}_{d \times d}$ -valued function, depending on the position of a particle, the empirical distribution, and time  $t$ .  $F_0(r, \mu, t)$  is as in (1.1). Consider stochastic ordinary differential equations (SODEs) for the displacement of  $r^i$  of the following type:

$$\begin{aligned} dr_{0,N}^i(t) &= F_0(r_{0,N}^i(t), X_{0,N}(t), t)dt + \mathcal{J}_0(r_{0,N}^i(t), X_{0,N}(t), t)d\beta_i(t) \\ r_{0,N}^i(0) &= q_0^i, i = 1, \dots, N, \quad X_{0,N}(t) := \sum_{i=1}^N \frac{1}{N} \delta_{r_{0,N}^i(t)}. \end{aligned} \quad (1.6)$$

Letting  $\mathbf{0} \in \mathcal{M}_{d \times d}$  denote the matrix with all entries being equal to 0, the two-particle diffusion matrix is given by

$$\tilde{D}_0(r^i, r^j, \mu, t) := \begin{cases} \mathcal{J}_0(r^i, \mu, t) \mathcal{J}_0^T(r^j, \mu, t), & \text{if } i = j, \\ \mathbf{0}, & \text{if } i \neq j, \end{cases} \quad (1.7)$$

and the one-particle matrix is  $D_0(r, \mu, t) = \mathcal{J}_0(r, \mu, t) \mathcal{J}_0^T(r, \mu, t)$ . Defining

$$\mathcal{L}_0 \varphi(r, \mu, t) = F_0(r, \mu, t) \cdot \nabla \varphi(r) + \frac{1}{2} \sum_{kl} D_{0,kl}(r, \mu, t) \partial_k \partial_l \varphi(r),$$

the analog of (1.3) is

$$\begin{aligned} \langle X_{0,N}(t), \varphi \rangle &= \langle X_{0,N}(0), \varphi \rangle + \int_0^t \langle X_{0,N}(s), \mathcal{L}_0 \varphi(\cdot, X_{0,N}(s), s) \rangle ds \\ &\quad + \frac{1}{N} \sum_{i=1}^N \int_0^t \nabla \varphi(r_{0,N}^i)^T \mathcal{J}_0(r_{0,N}^i, X_{0,N}(s), s) d\beta_i(s), \end{aligned} \quad (1.8)$$

and the quadratic variation is

$$\begin{aligned} &[\langle X_{0,N}, \varphi \rangle]_t \\ &= \int_0^t \sum_{i=1}^N \frac{1}{N^2} \sum_{k,l=1}^d (\partial_k \varphi)(r_{0,N}^i(t)) (\partial_l \varphi)(r_{0,N}^i(t)) \tilde{D}_{0,k,l}(r_{0,N}^i(t), r_{0,N}^i(t), X_{0,N}(t), t) dt. \end{aligned} \quad (1.9)$$

Clearly, we have  $N$  terms in the sum (1.9), divided by  $N^2$ , and it follows that the quadratic variation and the term involving the stochastic integrals in (1.8) tend to zero as  $N \rightarrow \infty$ .

Under the Lipschitz conditions defined in Section 2, the systems (1.1) and (1.6) have unique solutions, and assuming  $\{r_{\varepsilon,N}^i(0)\} \Rightarrow \{r_{\varepsilon,\infty}^i(0)\}$ , where  $\{r_{\varepsilon,\infty}^i(0)\} = \{q_\varepsilon^i\}$  is exchangeable, the results of Kurtz and Protter (1996), Section 10, imply that the solution of (1.1) converges to the solution of

$$dr_{\varepsilon,\infty}^i(t) = F_\varepsilon(r_{\varepsilon,\infty}^i(t), X_{\varepsilon,\infty}(t), t)dt + \int \mathcal{J}_\varepsilon(r_{\varepsilon,\infty}^i(t), q, X_{\varepsilon,\infty}(t), t) w(dq, dt), \quad (1.10)$$

$$r_{\varepsilon,\infty}^i(0) = q_\varepsilon^i, i = 1, 2, \dots \quad X_{\varepsilon,\infty}(t) := \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{1}{n} \delta_{r_{\varepsilon,\infty}^i(t)}$$

and the solution of (1.6) converges to the solution of

$$dr_{0,\infty}^i(t) = F_0(r_{0,\infty}^i(t), X_{0,\infty}(t), t)dt + \mathcal{J}_0(r_{0,\infty}^i(t), X_{0,\infty}(t), t)d\beta_i(t) \quad (1.11)$$

$$r_{0,\infty}^i(0) = q_0^i, i = 1, 2, \dots, X_{0,\infty}(t) := \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{1}{n} \delta_{r_{0,\infty}^i(t)}.$$

*Remark 1.1* If  $X_{\varepsilon,N}(0) \Rightarrow \mu$  in  $\mathbf{M}_1$ , the space of probability measures on  $\mathbf{R}^d$ , then one can always randomly permute the indices in (1.1) and (1.6) so that  $\{r_{\varepsilon,N}^{\sigma_i}(0)\} \Rightarrow \{r_{\varepsilon,\infty}^i(0)\}$ , where  $\{r_{\varepsilon,\infty}^i(0)\}$  is exchangeable. Consequently, convergence of the initial distribution  $X_{\varepsilon,N}(0) \Rightarrow \mu$  implies  $X_{\varepsilon,N} \Rightarrow X_{\varepsilon,\infty}$ , where  $X_{\varepsilon,\infty}$  is a solution of (1.5).

For  $\varepsilon = 0$ ,  $X_{0,N} \Rightarrow X_{0,\infty}$ , where  $X_{0,\infty}$  is a solution of the quasilinear parabolic partial differential equation (PDE) of McKean-Vlasov type given by

$$\langle X_{0,\infty}(t), \varphi \rangle = \langle \mu, \varphi \rangle + \int_0^t \langle X_{0,\infty}(s), \mathcal{L}_0 \varphi(\cdot, X_{0,\infty}(s), s) \rangle ds, \quad (1.12)$$

which is the weak form of the macroscopic McKean-Vlasov equation (or “nonlinear diffusion equation”)<sup>2</sup>

$$\frac{\partial}{\partial t} X_{0,\infty} = \frac{1}{2} \sum_{k,l=1}^d \partial_{kl}^2 (D_{0,kl}(\cdot, X_{0,\infty}, t) X_{0,\infty}) - \nabla \cdot (X_{0,\infty} F_0(\cdot, X_{0,\infty}, t)) \quad (1.13)$$

$$X_{0,\infty}(0) = \mu.$$

Under suitable assumptions, the solution of this PDE lives in the space of densities with respect to Lebesgue measure.<sup>3</sup> One may combine perturbations of each particle by its own independent Brownian motion, and the space-time perturbation as in (1.1). For some results of this type see Wang (1995), Kurtz and Protter (1996), Kurtz and Xiong (1999), and the references therein.

*Remark 1.2* The limit of  $X_{\varepsilon,N}$  as  $N \rightarrow \infty$  is “smooth” only if the initial distributions converge to a “smooth” initial condition; however, the main interest in the limit  $N \rightarrow \infty$  is the derivation of density-valued distributions. Therefore, we are calling the limit the “continuum limit” of the particle distribution for both (1.1) and (1.6). The main difference is that the continuum limit of the particle distribution for (1.1) is stochastic whereas for (1.6), it is deterministic or “macroscopic.”

To better understand the spatial correlations, let us consider the following examples:

*Example 1.3* Let  $\varepsilon > 0$  be given, and let  $|\cdot|$  denote the Euclidean norm on  $\mathbf{R}^d$ . Set  $\tilde{\Gamma}_\varepsilon(r) := \left(\frac{1}{(2\pi\varepsilon)}\right)^{\frac{d}{4}} \exp\left(-\frac{|r|^2}{4\varepsilon}\right)$  and  $\Gamma_\varepsilon(r)$  the diagonal  $d \times d$ -matrix whose entries on

<sup>2</sup> See Oelschläger (1984). Gärtner (1988) obtains the same result under local Lipschitz and linear growth conditions, making an additional assumption on the initial condition.

<sup>3</sup> See Oelschläger (1984) and Gärtner (1988).

the main diagonal are all  $\tilde{\Gamma}_\varepsilon(r)$ . Consider the system of stochastic ordinary differential equations (SODEs):

$$\begin{aligned} dr_\varepsilon^i &= \int \Gamma_\varepsilon(r_\varepsilon^i - q) w(dq, dt) \\ r_\varepsilon^i(0) &= q^i, i = 1, \dots, N, \quad X_{\varepsilon, N}(t) := \sum_{i=1}^N \frac{1}{N} \delta_{r_\varepsilon^i(t)}. \end{aligned} \quad (1.14)$$

Then the two-particle diffusion matrix is independent of  $\mu$  and  $t$ , and its entries are

$$\tilde{D}_{\varepsilon, k, l}(r_\varepsilon^i, r_\varepsilon^j) = \exp\left(-\frac{|r_\varepsilon^i - r_\varepsilon^j|^2}{8\varepsilon}\right) \delta_{kl} \quad \forall i, j = 1, \dots, N. \quad (1.15)$$

*Example 1.4* Let  $\varepsilon > 0$  be given and set  $G_\varepsilon(r) := -c_\varepsilon \nabla \left(\frac{1}{(2\pi\varepsilon)}\right)^{\frac{d}{4}} \exp\left(-\frac{|r|^2}{4\varepsilon}\right)$ , where  $c_\varepsilon > 0$ . Consider the system of stochastic ordinary differential equations (SODEs)

$$\begin{aligned} dr_\varepsilon^i &= \int G_\varepsilon(r_\varepsilon^i - q) w_1(dq, dt) \\ r_\varepsilon^i(0) &= q^i, i = 1, \dots, N, \quad X_{\varepsilon, N}(t) := \sum_{i=1}^N \frac{1}{N} \delta_{r_\varepsilon^i(t)}. \end{aligned} \quad (1.16)$$

The two-particle diffusion matrix is given by

$$\tilde{D}_\varepsilon(r_\varepsilon^i, r_\varepsilon^j) = \int G_\varepsilon(r^i - r^j - q) G_\varepsilon^T(-q) dq \quad \forall i, j = 1, \dots, N. \quad (1.17)$$

It follows from Kotelenetz (1995) that both (1.14) and (1.16) have unique (Itô) solutions. Levy's theorem implies that for each  $i$  both  $M_\varepsilon(r_\varepsilon^i, t) := \int_0^t \int \Gamma_\varepsilon(r_\varepsilon^i(s) - q) w(dq, ds)$  and  $M_\varepsilon(r_\varepsilon^i, t) := \int_0^t \int G_\varepsilon(r_\varepsilon^i(s) - q) w_1(dq, ds)$  are  $\mathbf{R}^d$ -valued Brownian motions (Ethier and Kurtz (1986), Chapter 7, Theorem 1.1). So (1.14) and (1.16) (and, more generally, (1.1)) extend the classical Einstein-Smoluchowski model from one Brownian particle to several correlated Brownian particles, where the correlation length depends on the distance of the particles from one another. Clearly,  $(M_\varepsilon(r_\varepsilon^i, t), M_\varepsilon(r_\varepsilon^j, t))$  for  $i \neq j$  is not Gaussian as an  $\mathbf{R}^{2d}$ -valued process; however, if  $|r_\varepsilon^i(t) - r_\varepsilon^j(t)| \gg \sqrt{\varepsilon}$ , the  $d$ -dimensional martingales are almost uncorrelated and the joint process should look essentially like an  $\mathbf{R}^{2d}$ -valued (resp.  $\mathbf{R}^{d \cdot N}$ -valued) Brownian motion. Example 1.4 is motivated by the derivation of correlated Brownian motions from deterministic dynamics of "large" solute and "small" solvent particles in a scaling limit (Kotelenetz (2005a,b)) and the analysis of depletion forces in colloids (Kotelenetz et al (2007))<sup>4</sup>.

*Example 1.5* Examples 1.3 and 1.4 are both of the form

$$\mathcal{J}_\varepsilon(r, q, \mu, t) = \varepsilon^{-d/2} \mathcal{J}\left(r, \frac{r-q}{\varepsilon}, \mu, t\right)$$

<sup>4</sup> In our approach the fluid for the suspended particles is being modeled by the space-time white noise random field and the interaction of the fluid particles (the solvent) and the suspended particles (the solute) is given by a smooth kernel. Using the same random field as a perturbation for all particles reflects the simple fact that real suspended particles float in the same fluid. (See Spohn (1991), Part II, Section 7.2, and also our Remark 1.6).

(after replacing  $\varepsilon$  by  $\varepsilon^2$ ). By a change of variable in (1.2), we can then write

$$\tilde{D}_\varepsilon(r^i, r^j, \mu, t) = \int \mathcal{J}(r^i, \frac{r^i}{\varepsilon} - z, \mu, t) \mathcal{J}^T(r^j, \frac{r^j}{\varepsilon} - z, \mu, t) dz,$$

assuming  $\int |\mathcal{J}(r, z, \mu, t)|^2 dz < \infty$ . Under this assumption, we note that

$$\lim_{\varepsilon \rightarrow 0} \tilde{D}_\varepsilon(r^i, r^j, \mu, t) = \begin{cases} \int \mathcal{J}(r, z, \mu, t) \mathcal{J}^T(r, z, \mu, t) dz & \text{if } r_i = r_j = r \\ \mathbf{0} & \text{if } r_i \neq r_j \end{cases}.$$

*Remark 1.6* (i) The spatial correlation of several large Brownian particles suspended in a fluid and sufficiently close to each other is an empirical fact. The simplest argument is provided by the colloid community: If two large particles get closer to each other than a multiple of the diameter of a typical small particle (which represent the medium), then the osmotic pressure on each of the large particles can no longer be isotropic, as the fluid between the two large particles gets “depleted.” (See, for example, Asakura and Oosawa (1954), Götzelmann et al (1998), as well as Kotelenez et al (2007).)

(ii) The choice of a fixed  $\varepsilon > 0$  in Kotelenez’s scaling limit enables us to study, in a stochastic continuum model, fluctuations and anisotropy which occur in a discrete “microscopic” particle picture and which are neglected in the classical approach. In this paper, we will show that under certain assumptions, the solutions of the SPDE (1.5) converge to the solutions of the (macroscopic) PDE (1.12), if the correlations tend to 0. This convergence establishes our SPDE as a mesoscopic model for the distribution of (large) suspended particles. (For an explanation of the term “mesoscopic” see van Kampen (1981), Haken (1983), and others.)

In Section 2, we introduce the Wasserstein metric and formulate the main hypotheses. In Section 3, we state and prove the macroscopic limit theorem. In the Appendix, Section 4.1 includes general results on the convergence of exchangeable systems that are used to verify the interchange of the limits  $N \rightarrow \infty$  and  $\varepsilon \rightarrow 0$ , and Section 4.2 contains lemmas giving conditions under which diffusion-like processes do not hit points. These conditions are critical in the proof that the noises driving the individual particles in the limiting models are independent. Section 4.3 includes comments on previous related work.

For  $s < t$  and  $X_{\varepsilon, \infty}$  given by (1.10), let  $\psi_{s,t}(z)$  satisfy

$$\begin{aligned} \psi_{s,t}(z) = z + \int_s^t F_\varepsilon(\psi_{s,u}(z), X_{\varepsilon, \infty}(u), u) du \\ + \int_{\mathbf{R}^d \times (s,t]} \mathcal{J}_\varepsilon(\psi_{s,u}(z), q, X_{\varepsilon, \infty}(u), u) w(dq, du). \end{aligned} \quad (1.18)$$

Then, at least formally,  $\psi_{s,t}(z) = \psi_{u,t}(\psi_{s,u}(z))$ ,  $s < u < t$ , and  $r_{\varepsilon, \infty}^i(t) = \psi_{0,t}(r_{\varepsilon, \infty}^i(0))$ . In fact, (1.10) can be reformulated as (1.18) plus the requirement that

$$\langle X_{\varepsilon, \infty}(t), \varphi \rangle = \langle X_{\varepsilon, \infty}(0), \varphi \circ \psi_{0,t} \rangle.$$

This observation demonstrates a close connection between the systems considered here and stochastic flows as studied by Kunita (1990). Kunita formulates the analog

of (1.18) in what is apparently a very different manner, writing a stochastic equation driven by what he calls a  $C$ -valued Brownian motion. In Section 4.4, we compare Kunita's approach with the Gaussian white-noise integral approach taken here.

## 2 The Wasserstein metric and basic assumptions

Let  $\rho(\cdot, \cdot)$  be a metric on  $\mathbf{R}^d$ . We are mainly interested in  $\rho(r, q) := |r - q|$ , where the latter is the Euclidean distance on  $\mathbf{R}^d$ , and in

$$\rho(r, q) := |r - q| \wedge 1,$$

where " $\wedge$ " denotes "minimum." Most statements will be made with respect to  $\rho(r, q)$  with additional comments for the Euclidean distance.  $C_{L,\infty}(\mathbf{R}^d; \mathbf{R})$  is the space of all uniformly bounded Lipschitz functions  $f$  from  $\mathbf{R}^d$  into  $\mathbf{R}$ . Define

$$\|f\|_\infty := \sup_q |f(q)|; \quad \|f\|_L := \sup_{r \neq q} \left| \frac{f(r) - f(q)}{\rho(r, q)} \right|; \quad \|f\|_{L,\infty} := \|f\|_L + \|f\|_\infty.$$

Let  $\mathbf{M}_1$  denote the Borel probability measures on  $\mathbf{R}^d$ . For  $\mu, \nu \in \mathbf{M}_1$ , we set

$$\gamma(\mu, \nu) := \sup_{\|f\|_{L,\infty} \leq 1} \left| \int f(q)(\mu(dq) - \nu(dq)) \right|. \quad (2.1)$$

$\mathbf{M}_1$ , endowed with the metric  $\gamma$ , is a complete separable metric space, and the space of finite sums of point measures with nonnegative weights and total mass 1 is dense in  $(\mathbf{M}_1, \gamma)$ . (See de Acosta (1982).) Actually,  $\gamma$  is the restriction of a norm

$$\|\zeta\|_\gamma := \sup_{\|f\|_{L,\infty} \leq 1} |\langle \zeta, f \rangle|$$

on the dual of  $C_{L,\infty}(\mathbf{R}^d; \mathbf{R})$ , that is,  $\gamma(\mu, \nu) = \|\mu - \nu\|_\gamma$  for  $\mu, \nu \in \mathbf{M}_1$ . We easily see that

$$\|\mu\|_\gamma = \mu(\mathbf{R}^d) = 1,$$

for  $\mu \in \mathbf{M}_1$ .

We assume the following Lipschitz and boundedness or growth conditions on the coefficients of our equations.

**Condition 2.1** The mappings  $(r, \mu, t) \in \mathbf{R}^d \times \mathbf{M}_1 \times [0, \infty) \rightarrow F_\varepsilon(r, \mu, t) \in \mathbf{R}^d$  and

$$(r, \mu, t) \in \mathbf{R}^d \times \mathbf{M}_1 \times [0, \infty) \rightarrow \mathcal{J}_\varepsilon(r, \cdot, \mu, t) \in L^2(\mathbf{R}^d)$$

are continuous, and there exist  $c_F, c_{\mathcal{J}} \in (0, \infty)$  such that for all  $\varepsilon > 0$ ,  $(r_1, \mu_1), (r_2, \mu_2) \in \mathbf{R}^d \times \mathbf{M}_1$ , and  $t \geq 0$ ,

$$\begin{aligned} |F_\varepsilon(r_1, \mu_1, t) - F_\varepsilon(r_2, \mu_2, t)| &\leq c_F \{\rho(r_1, r_2) + \gamma(\mu_1, \mu_2)\}, \\ \sum_{k,l=1}^d \int (\mathcal{J}_{\varepsilon,kl}(r_1, q, \mu_1, t) - \mathcal{J}_{\varepsilon,kl}(r_2, q, \mu_2, t))^2 dq & \quad (2.2) \\ &\leq c_{\mathcal{J}}^2 \{\rho^2(r_1 - r_2) + \gamma^2(\mu_1 - \mu_2)\}. \end{aligned}$$

**Condition 2.2** Either

$$\sup_{r, \mu, t, \varepsilon} (|F_\varepsilon(r, \mu, t)|^2 + \sum_{k, l=1}^d \int \mathcal{J}_{\varepsilon, kl}^2(r, q, \mu, t) dq) < \infty$$

or  $\rho(x, y) = |x - y|$  and there exists  $c > 0$  such that

$$\begin{aligned} |F_\varepsilon(r, \mu_1, t)| &\leq c(1 + |r|) \\ \int |\mathcal{J}_\varepsilon(r, q, \mu, t)|^2 dq &\leq c(1 + |r|)^2. \end{aligned} \quad (2.3)$$

Note that Condition 2.2 ensures that for  $p \geq 1$ , if  $E[|r_\varepsilon^i(0)|^p] < \infty$ , then

$$E[\sup_{s \leq t} |r_\varepsilon^i(s)|^p] < \infty. \quad (2.4)$$

We assume the same conditions for  $\varepsilon = 0$ , replacing the integral Lipschitz and linear growth conditions for  $\mathcal{J}_\varepsilon$  by the corresponding condition on  $\mathcal{J}_0$  and the stochastic driving term as in (1.6).<sup>5</sup>

Let  $\sigma_\varepsilon(r, \mu, t)$  be the nonnegative definite square root of  $D_\varepsilon(r, \mu, t)$ . We will require the following nondegeneracy condition.

**Condition 2.3** For  $T > 0$ , compact  $K \subset \mathbf{R}^d$ , and compact  $\mathcal{C} \subset \mathbf{M}_1$ ,

$$\inf_{0 < \varepsilon < 1, r \in K, 0 \leq t \leq T, \mu \in \mathcal{C}} \inf_{z \in \mathbf{R}^d, z \neq 0} \frac{z^T D_\varepsilon(r, \mu, t) z}{|z|^2} > 0 \quad (2.5)$$

or equivalently

$$\sup_{0 < \varepsilon < 1, r \in K, 0 \leq t \leq T, \mu \in \mathcal{C}} |\sigma_\varepsilon^{-1}(r, \mu, t)| < \infty. \quad (2.6)$$

We need the following convergence conditions as  $\varepsilon \rightarrow 0$ .

**Condition 2.4** For  $\delta, T > 0$ , compact  $K \subset \mathbf{R}^d$ , and compact  $\mathcal{C} \subset \mathbf{M}_1$ ,

$$\lim_{\varepsilon \downarrow 0} \sup_{|r_1 - r_2| > \delta, 0 \leq t \leq T, \mu \in \mathcal{C}} |\tilde{D}_\varepsilon(r_1, r_2, \mu, t)| = 0, \quad (2.7)$$

and

$$\lim_{\varepsilon \downarrow 0} \sup_{r \in K, 0 \leq t \leq T} \sup_{\mu \in \mathcal{C}} (|F_\varepsilon(r, \mu, t) - F_0(r, \mu, t)| + |\sigma_\varepsilon(r, \mu, t) - \mathcal{J}_0(r, \mu, t)|) = 0. \quad (2.8)$$

If  $d = 2$ , we will need an additional regularity condition. A continuous  $\mathbf{R}$ -valued process  $Y$  is an *Itô process* if it can be written as

$$Y(t) = Y(0) + M(t) + \int_0^t b(s) ds,$$

where  $M$  is a martingale with

$$[M]_t = \int_0^t a(s) ds$$

<sup>5</sup> See Kotelenez (1995, 2007) for examples.

and  $a$  and  $b$  are progressively measurable processes. We will say that  $Y$  has *locally bounded characteristics* if  $\sup_{s \leq t} (|a(s)| + |b(s)|) < \infty$  a.s. for each  $t > 0$  and  $L^1$ -*bounded characteristics* if  $E[\sup_{s \leq t} (|a(s)| + |b(s)|)] < \infty$ .

A continuous  $\mathbf{M}_1$ -valued process  $X$  is a *de Finetti process* if there exists an exchangeable family  $\{\zeta^i\}$  of  $\mathbf{R}^d$ -valued processes such that

$$\langle X(t), \varphi \rangle = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \varphi(\zeta^i(t)).$$

A de Finetti process  $X$  has  $L^1$ -*bounded characteristics* if the  $\zeta^i$  are Itô processes with  $L^1$ -bounded characteristics.

**Lemma 2.5** *If  $X$  is a de Finetti process with  $L^1$ -bounded characteristics, then for each  $\varphi \in C_b^2$ ,  $\langle X, \varphi \rangle$  is an Itô process with  $L^1$ -bounded characteristics.*

*Proof.* We can write  $\zeta^i(t) = M^i(t) + \int_0^t b^i(s) ds$ , where  $[M^i]_t = \int_0^t a^i(s) ds$  for a  $d \times d$ -matrix  $a^i = ((a_{kl}^i))$  and  $E[\sup_{s \leq t} (|b^i(s)| + |a^i(s)|)] < \infty$ . Then

$$\langle X(t), \varphi \rangle = \langle X(0), \varphi \rangle + M_\varphi(t) + \int_0^t b_\varphi(s) ds$$

where

$$M_\varphi(t) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^t \nabla \varphi(\zeta^i(s))^T dM^i(s) \quad (2.9)$$

and

$$\int_0^t b_\varphi(s) ds = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \int_0^t (\nabla \varphi(\zeta^i(s)) \cdot b^i(s) + \frac{1}{2} \sum_{kl} \partial_k \partial_l \varphi(\zeta^i(s)) a_{kl}^i(s)) ds.$$

In both cases, the convergence is in  $L^1$ . Let  $M_\varphi^n(t)$  denote the process on the right of (2.9), and note that

$$[M_\varphi^n]_{t+u} - [M_\varphi^n]_t \leq \|\nabla \varphi\|^2 \frac{1}{n^2} \sum_{1 \leq i, j \leq n} \int_t^{t+u} \frac{|a^i(s)| + |a^j(s)|}{2}.$$

The  $L^1$ -boundedness assumption implies the  $\{M_\varphi^n\}$  satisfies the convergence conditions in Theorem 2.2 of Kurtz and Protter (1991) which in turn implies that the quadratic variation of  $M_\varphi^n$  converges to the quadratic variation of  $M_\varphi$  giving the  $L^1$ -boundedness of the characteristics for  $M_\varphi$  and hence for  $\langle X, \varphi \rangle$ .  $\square$

**Condition 2.6** For every  $\mathbf{R}^2$ -valued Itô process  $R$  with  $L^1$ -bounded characteristics and every  $\mathbf{M}^1$ -valued de Finetti process  $X$  with  $L^1$ -bounded characteristics,

$$\mathcal{J}_0(R(\cdot), X(\cdot), \cdot)$$

is an Itô process with  $L^1$ -bounded characteristics.

*Remark 2.7* Condition 2.6 will be satisfied if

$$\mathcal{J}_0(r, \mu, t) = H(r, \langle \varphi_1, \mu \rangle, \dots, \langle \varphi_m, \mu \rangle, t),$$

where  $\varphi_i \in C_b^2$  and  $H$  is sufficiently smooth.

Condition 2.6 also holds if

$$\mathcal{J}_0(r, \mu, t) = \int_{\mathbf{R}^2} \Gamma(r - q) \mu(dq)$$

for  $\Gamma \in C_b^2(\mathbf{R}^2)$ , since

$$\langle X(t), \Gamma(R(t) - \cdot) \rangle = \langle X^R, \Gamma \rangle$$

where  $X^R$  is the de Finetti process given by  $\{R - \zeta_i\}$ .

The following lemma ensures relative compactness of  $\{(\{r_{\varepsilon, N}^i\}, X_{\varepsilon, N}), \varepsilon > 0, N > 0\}$ .

**Lemma 2.8** *For  $\varepsilon > 0$ , let  $N_\varepsilon = \infty$  or  $N_\varepsilon \rightarrow \infty$  as  $\varepsilon \rightarrow 0$ . Let  $\{r_{\varepsilon, N_\varepsilon}^i\}$  be a solution of (1.1) or if  $N_\varepsilon = \infty$ , a solution (1.10).*

*Suppose that Conditions 2.1 and 2.2 hold and that  $\{\{r_{\varepsilon, N_\varepsilon}^i(0)\}, \varepsilon \geq 0\}$  is relatively compact (for convergence in distribution in  $(\mathbf{R}^d)^\infty$ ). Then  $\{(\{r_{\varepsilon, N_\varepsilon}^i\}, X_{\varepsilon, N_\varepsilon})\}$  is relatively compact in  $C_{(\mathbf{R}^d)^\infty \times \mathbf{M}_1}[0, \infty)$ , and any limit point  $\{(\{\hat{r}_\infty^i\}, \hat{X}_\infty)\}$  satisfies*

$$\hat{X}_\infty(t) = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m \delta_{\hat{r}_\infty^i(t)}.$$

*Proof.* The boundedness and growth conditions on  $F_\varepsilon$  and  $\mathcal{J}_\varepsilon$  ensure that  $\{r_{\varepsilon, N_\varepsilon}^i, \varepsilon > 0\}$  is relatively compact for each fixed  $i$  and hence  $\{\{r_{\varepsilon, N_\varepsilon}^i\}, \varepsilon > 0\}$  is relatively compact in the infinite product space  $(C_{\mathbf{R}^d}[0, \infty))^\infty$  (taking  $r_{\varepsilon, N_\varepsilon}^i \equiv 0$  for  $i > N_\varepsilon$ ). But the topology on  $(C_{\mathbf{R}^d}[0, \infty))^\infty$  is the same as the topology on  $C_{(\mathbf{R}^d)^\infty}[0, \infty)$ , so  $\{\{r_{\varepsilon, N_\varepsilon}^i\}, \varepsilon > 0\}$  is relatively compact in  $C_{(\mathbf{R}^d)^\infty}[0, \infty)$ . The lemma now follows by Lemma 4.4 and Remark 4.5.  $\square$

We will need the following lemma in the case  $d = 2$ .

**Lemma 2.9** *In addition to the assumptions of Lemma 2.8, suppose that Condition 2.4 holds and that  $\sup_{\varepsilon > 0} E[(r_{\varepsilon, N_\varepsilon}^i(0))^2] < \infty$ . Then any limit point of  $\{X_{\varepsilon, N_\varepsilon}\}$  is a de Finetti process with  $L^1$ -bounded characteristics.*

*Proof.* By Lemma 2.8,  $\{(\{r_{\varepsilon, N_\varepsilon}^i\}, X_{\varepsilon, N_\varepsilon})\}$  is relatively compact, and Condition 2.4 and the continuous mapping theorem imply that any limit point will satisfy

$$\hat{r}^i(t) = \hat{r}^i(0) + \int_0^t F_0(\hat{r}^i(s), \hat{X}(s), s) ds + M^i(t),$$

where  $M^i$  is the limit in distribution of

$$M_\varepsilon^i = \int_{\mathbf{R}^d \times [0, \cdot]} \mathcal{J}_\varepsilon(r_{\varepsilon, \infty}^i(s), q, X_{\varepsilon, \infty}(s), s) w(dq, ds).$$

The moment estimates ensure that  $M^i$  is a martingale, and the quadratic variation of  $M^i$  is the limit of the quadratic variation of  $M_\varepsilon^i$  giving

$$[M^i]_t = \int_0^t \mathcal{J}_0(\hat{r}^i(s), \hat{X}(s), s) \mathcal{J}_0^T(\hat{r}^i(s), \hat{X}(s), s) ds.$$

$L^1$ -boundedness of the characteristics follows from (2.3) and (2.4).  $\square$

### 3 The macroscopic limit

Under the conditions of Section 2, for fixed  $N$ , as  $\varepsilon \rightarrow 0$ ,  $\{r_{\varepsilon, N}^i(0)\} \Rightarrow \{r_{0, N}^i(0)\}$  easily implies  $\{r_{\varepsilon, N}^i\} \Rightarrow \{r_{0, N}^i\}$ . Assuming exchangeability of  $\{r_{0, N}^i(0)\}$ , we have already observed that as  $N \rightarrow \infty$ ,  $\{r_{0, N}^i(0)\} \Rightarrow \{r_{0, \infty}^i(0)\}$  implies  $(\{r_{0, N}^i\}, X_{0, N}) \Rightarrow (\{r_{0, \infty}^i\}, X_{0, \infty})$ . Our main theorem states that the order of convergence can be interchanged.

**Theorem 3.1** *For  $\varepsilon > 0$ , let  $N_\varepsilon = \infty$  or  $N_\varepsilon \rightarrow \infty$  as  $\varepsilon \rightarrow 0$ . Let  $\{r_{\varepsilon, N_\varepsilon}^i\}$  be a solution of (1.1) or if  $N_\varepsilon = \infty$ , a solution (1.10). Assume that Conditions 2.1, 2.2, 2.3, and 2.4 hold, and in addition, if  $d = 2$ , that Condition 2.6 holds and*

$$\sup_{\varepsilon > 0} E[(r_{\varepsilon, N_\varepsilon}^i(0))^2] < \infty.$$

Suppose that  $\{r_{\varepsilon, N_\varepsilon}^i(0)\}$  is exchangeable and  $\{r_{\varepsilon, N_\varepsilon}^i(0)\} \Rightarrow \{r_{0, \infty}^i(0)\}$  where

$$P\{r_{0, \infty}^i(0) \neq r_{0, \infty}^j(0)\} = 1, \quad \text{for } i \neq j.$$

Then

$$(\{r_{\varepsilon, N_\varepsilon}^i\}, X_{\varepsilon, N_\varepsilon}) \Rightarrow (\{r_{0, \infty}^i\}, X_{0, \infty}) \text{ in } C(\mathbf{R}^d)^\infty \times \mathbf{M}_1[0, \infty), \text{ as } \varepsilon \downarrow 0. \quad (3.1)$$

*Example 3.2* Before proving Theorem 3.1 in full generality, we first consider the particular case given in Example 1.5. System (1.10) then becomes

$$\begin{aligned} dr_{\varepsilon, \infty}^i(t) &= F_\varepsilon(r_{\varepsilon, \infty}^i(t), X_{\varepsilon, \infty}(t), t) dt \\ &\quad + \int \varepsilon^{-d/2} \mathcal{J}(r_{\varepsilon, \infty}^i(t), \frac{r_{\varepsilon, \infty}^i(t) - q}{\varepsilon}, X_{\varepsilon, \infty}(t), t) w(dq, dt). \end{aligned}$$

Define

$$w_\varepsilon^i(C \times [0, t]) = \int_{\mathbf{R}^d \times [0, t]} \varepsilon^{-d/2} \mathbf{1}_C(\frac{r_{\varepsilon, \infty}^i(t) - q}{\varepsilon}) w(dq, ds),$$

and note that

$$[w_{\varepsilon, l}^i(C \times [0, \cdot]), w_{\varepsilon, k}^i(C \times [0, \cdot])]_t = \delta_{kl} \int_0^t \int \varepsilon^{-d} \mathbf{1}_C(\frac{r_{\varepsilon, \infty}^i(t) - q}{\varepsilon}) dq ds = \delta_{kl} |C| t.$$

Consequently, by Levy's theorem,  $w_\varepsilon^i$  has the same distribution as  $w$ . Writing the system in terms of the  $w_\varepsilon^i$ , we have

$$dr_{\varepsilon, \infty}^i = F_\varepsilon(r_{\varepsilon, \infty}^i(t), X_{\varepsilon, \infty}(t), t) dt + \int \mathcal{J}(r_{\varepsilon, \infty}^i(t), z, X_{\varepsilon, \infty}(t), t) w_\varepsilon^i(dz, dt).$$

For  $\varphi \in L^2(\mathbf{R}^d)$ , define

$$w_\varepsilon^i(\varphi, t) = \int_{\mathbf{R}^d \times [0, t]} \varphi(z) w_\varepsilon^i(dz, ds).$$

The collection  $\{\{w_\varepsilon^i\}, \varepsilon > 0\}$  is relatively compact in the sense that for any finite collection of indices  $(i_1, l_1), \dots, (i_m, l_m)$  and  $\varphi_{i_j, l_j} \in L^2(\mathbf{R}^d)$ ,

$$\{(w_\varepsilon^{i_1}(\varphi_{i_1, l_1}, \cdot), \dots, w_\varepsilon^{i_m}(\varphi_{i_m, l_m}, \cdot))\}$$

is relatively compact in  $\mathbf{C}_m[0, \infty)$ . At least along a subsequence, the exchangeability of  $\{(r_{\varepsilon, \infty}^i, w_\varepsilon^i)\}$  and Lemma 4.4 imply  $\{(r_{\varepsilon, \infty}^i, w_\varepsilon^i), X_{\varepsilon, \infty}\}$  converges in distribution to a process  $\{(\hat{r}_\infty^i, w^i), \hat{X}_\infty\}$ , where

$$\hat{X}_\infty(t) = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m \delta_{\hat{r}_\infty^i}(t).$$

Convergence of the  $w_\varepsilon^i$  is in the sense that (along the subsequence)

$$\{(w_\varepsilon^{i_1}(\varphi_{i_1, l_1}, \cdot), \dots, w_\varepsilon^{i_m}(\varphi_{i_m, l_m}, \cdot))\} \Rightarrow \{(w^{i_1}(\varphi_{i_1, l_1}, \cdot), \dots, w^{i_m}(\varphi_{i_m, l_m}, \cdot))\},$$

for all choices of  $\varphi_{i_j, l_j} \in L^2(\mathbf{R}^d)$  and  $m = 1, 2, \dots$

By Theorem 4.2 of Kurtz and Protter (1996), for each  $i$ ,

$$d\hat{r}_\infty^i = F_0(\hat{r}_\infty^i(t), \hat{X}_\infty(t), t)dt + \int \mathcal{J}(\hat{r}_\infty^i(t), z, \hat{X}_\infty(t), t) w^i(dz, dt), \quad (3.2)$$

Each  $w^i$  is Gaussian white noise, and if they are independent, then the solution of (3.2) has the same distribution as the solution of (1.11). Independence of the  $w^i$  is at least not immediate. Specifically, for  $\varphi_1, \varphi_2 \in L^2(\mathbf{R}^d)$ ,

$$(w_{\varepsilon, l}^i(\varphi_1, \cdot), w_{\varepsilon, k}^j(\varphi_2, \cdot)) \Rightarrow (w_l^i(\varphi_1, \cdot), w_k^j(\varphi_2, \cdot)),$$

and since

$$\begin{aligned} [w_{\varepsilon, l}^i(\varphi_1, \cdot), w_{\varepsilon, k}^j(\varphi_2, \cdot)]_t &= w_{\varepsilon, l}^i(\varphi_1, t) w_{\varepsilon, k}^j(\varphi_2, t) - \int_0^t w_{\varepsilon, l}^i(\varphi_1, s) dw_{\varepsilon, k}^j(\varphi_2, s) \\ &\quad - \int_0^t w_{\varepsilon, k}^j(\varphi_2, s) dw_{\varepsilon, l}^i(\varphi_1, s), \end{aligned}$$

by Theorem 2.2 of Kurtz and Protter (1991),

$$[w_{\varepsilon, l}^i(\varphi_1, \cdot), w_{\varepsilon, k}^j(\varphi_2, \cdot)]_t \Rightarrow [w_l^i(\varphi_1, \cdot), w_k^j(\varphi_2, \cdot)]_t.$$

Consequently,

$$[w_l^i(\varphi_1, \cdot), w_k^j(\varphi_2, \cdot)]_t = \delta_{kl} \lim_{\varepsilon \rightarrow 0} \int_0^t \int \varphi_1\left(\frac{r_{\varepsilon, \infty}^i(s)}{\varepsilon} - q\right) \varphi_2\left(\frac{r_{\varepsilon, \infty}^j(s)}{\varepsilon} - q\right) dq ds,$$

where the convergence is in distribution, and for  $i \neq j$ , the limit is zero for all choice of  $\varphi_1$  and  $\varphi_2$  if and only if  $\inf_{s \leq t} |\hat{r}_\infty^i(s) - \hat{r}_\infty^j(s)| > 0$  a.s. for each  $t > 0$ . It follows by Lévy's theorem that  $w^i$  and  $w^j$  behave as independent Gaussian white noises until

$$\tau_{ij} = \lim_{\delta \rightarrow \infty} \tau_{ij}(\delta) := \lim_{\delta \rightarrow \infty} \inf\{t : |\hat{r}_\infty^i(t) - \hat{r}_\infty^j(t)| \leq \delta\}. \quad (3.3)$$

We need to show that  $\tau_{ij} = \infty$ .

Let

$$\begin{aligned} Y(t) &= \hat{r}_\infty^i(t) - \hat{r}_\infty^j(t) \\ &= \hat{r}_\infty^i(0) - \hat{r}_\infty^j(0) + \int_0^t (F_0(\hat{r}_\infty^i(s), \hat{X}_\infty(s), s) - F_0(\hat{r}_\infty^j(s), \hat{X}_\infty(s), s)) ds \\ &\quad + \int_{\mathbf{R}^d \times [0, t]} \mathcal{J}(\hat{r}_\infty^i(s), z, \hat{X}_\infty(s), s) w^i(dz, ds) \\ &\quad - \int_{\mathbf{R}^d \times [0, t]} \mathcal{J}(\hat{r}_\infty^j(t), z, \hat{X}_\infty(t), t) w^j(dz, dt). \end{aligned}$$

Then setting

$$\begin{aligned} M(t) &= \int_{\mathbf{R}^d \times [0, t]} \mathcal{J}(\hat{r}_\infty^i(s), z, \hat{X}_\infty(s), s) w^i(dz, ds) \\ &\quad - \int_{\mathbf{R}^d \times [0, t]} \mathcal{J}(\hat{r}_\infty^j(t), z, \hat{X}_\infty(t), t) w^j(dz, dt) \end{aligned}$$

and  $D(r, \mu, t) = \int \mathcal{J}(r, z, \mu, t) \mathcal{J}^T(r, z, \mu, t) dz$ ,

$$[M]_t = \int_0^t (D(\hat{r}_\infty^i(s), \hat{X}_\infty(s), s) + D(\hat{r}_\infty^j(s), \hat{X}_\infty(s), s)) ds,$$

for  $t < \tau_{ij}$ . For  $d \geq 3$ ,  $\tau_{ij} = \infty$  a.s. by Lemma 4.7. For  $d = 2$ , under the analog of Condition 2.6, Lemmas 2.9 and 4.8 imply  $\tau_{ij} = \infty$  a.s.

*Proof.*[of Theorem 3.1] Define

$$W_\varepsilon^i(t) := \int_0^t \int \sigma_\varepsilon^{-1}(r_{\varepsilon, N_\varepsilon}^i(s), X_{\varepsilon, N_\varepsilon}(s), s) \mathcal{J}_\varepsilon(r_{\varepsilon, N_\varepsilon}^i(s), q, X_{\varepsilon, N_\varepsilon}(s), s) w(dq, ds). \quad (3.4)$$

By Levy's theorem, each  $W_\varepsilon^i$  is an  $\mathbf{R}^d$ -valued Brownian motion, and

$$\begin{aligned} r_{\varepsilon, N_\varepsilon}^i(t) &= r_{\varepsilon, N_\varepsilon}^i(0) + \int_0^t F_\varepsilon(r_{\varepsilon, N_\varepsilon}^i(s), X_{\varepsilon, N_\varepsilon}(s), s) ds \\ &\quad + \int_{\mathbf{R}^d \times [0, t]} \mathcal{J}_\varepsilon(r_{\varepsilon, N_\varepsilon}^i(s), q, X_{\varepsilon, N_\varepsilon}(s), s) w(dq, ds) \\ &= r_{\varepsilon, N_\varepsilon}^i(0) + \int_0^t F_\varepsilon(r_{\varepsilon, N_\varepsilon}^i(s), X_{\varepsilon, N_\varepsilon}(s), s) ds \\ &\quad + \int_0^t \sigma_\varepsilon(r_{\varepsilon, N_\varepsilon}^i(s), X_{\varepsilon, N_\varepsilon}(s), s) dW_\varepsilon^i(s). \end{aligned}$$

Abbreviating

$$c_\varepsilon(r, q, \mu, t) := \sigma_\varepsilon^{-1}(r, \mu, t) D_\varepsilon(r, q, \mu, t) \sigma_\varepsilon^{-1}(q, \mu, t),$$

we obtain

$$[W_\varepsilon^i, W_\varepsilon^j]_t = \int_0^t c_\varepsilon(r_{\varepsilon, N_\varepsilon}^i(s), r_{\varepsilon, N_\varepsilon}^j(s), X_{\varepsilon, N_\varepsilon}(s), s) ds, \quad (3.5)$$

where the left side of (3.5) is the tensor quadratic variation. By Lemma 2.8 and Theorem 2.7 of Kurtz and Protter (1991),  $(\{(r_{\varepsilon, N_\varepsilon}^i, W_\varepsilon^i)\}, X_{\varepsilon, N_\varepsilon})$  is relatively compact and any limit point satisfies

$$\begin{aligned} \hat{r}_\infty^i(t) &= r_{0, \infty}^i(0) + \int_0^t F_0(\hat{r}_\infty^i(s), \hat{X}_\infty(s), s) ds \\ &\quad + \int_0^t \mathcal{J}_0(\hat{r}_\infty^i(s), \hat{X}_\infty(s), s) dW^i(s) \end{aligned}$$

where

$$\hat{X}_\infty(t) = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m \delta_{\hat{r}_{i\infty}^i(t)}.$$

The  $W^i$  are Brownian motions whose independence we must still verify.

Conditions 2.3 and 2.4 imply that for any  $\delta > 0$ , any compact subset  $\mathcal{C} \subset \mathbf{M}_1$ , and any  $T > 0$ ,

$$\lim_{\varepsilon \downarrow 0} \sup_{|r-q| > \delta} \sup_{\mu \in \mathcal{C}, 0 \leq t \leq T} |c_\varepsilon(r, q, \mu, t)| = 0. \quad (3.6)$$

Consequently, as in Example 3.2,

$$\lim_{\varepsilon \rightarrow 0} [W_\varepsilon^i, W_\varepsilon^j]_t = \delta_{ij} t I$$

provided

$$\tau_{ij} = \lim_{\delta \rightarrow \infty} \tau_{ij}(\delta) := \lim_{\delta \rightarrow \infty} \inf\{t : |\hat{r}_\infty^i(t) - \hat{r}_\infty^j(t)| \leq \delta\} \quad (3.7)$$

is infinite a.s. This assertion follows by Lemmas 4.7 and 4.8.

The independence of the  $W^i$  then implies that the limit is given by the unique solution of (1.11).  $\square$

## 4 Appendix

### 4.1 Convergence of sequences of exchangeable families.

Let  $S$  be a complete, separable metric space. A family of  $S$ -valued random variables  $\{\xi_1, \dots, \xi_m\}$  is exchangeable if for every permutation  $(\sigma_1, \dots, \sigma_m)$  of  $(1, \dots, m)$ ,  $\{\xi_{\sigma_1}, \dots, \xi_{\sigma_m}\}$  has the same distribution as  $\{\xi_1, \dots, \xi_m\}$ . A sequence  $\xi_1, \xi_2, \dots$  is exchangeable if every finite subfamily  $\xi_1, \dots, \xi_m$  is exchangeable.

**Theorem 4.1** (de Finetti) *Let  $\xi_1, \xi_2, \dots$  be an exchangeable sequence of  $S$ -valued random variables. Then there is a  $\mathcal{P}(S)$ -valued random variable  $\Xi$  such that*

$$\Xi = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m \delta_{\xi_i} \quad \text{a.s.}$$

and, conditioned on  $\Xi$ ,  $\xi_1, \xi_2, \dots$  are iid with distribution  $\Xi$ , that is, for each  $f \in B(S^m)$ ,  $m = 1, 2, \dots$ ,

$$E[f(\xi_1, \dots, \xi_m) | \Xi] = \langle f, \Xi^m \rangle.$$

We will refer to  $\Xi$  as the de Finetti measure for  $\xi_1, \xi_2, \dots$ . The next result focuses on convergence of a sequence of exchangeable families.

**Lemma 4.2** For  $n = 1, 2, \dots$ , let  $\{\xi_1^n, \dots, \xi_{N_n}^n\}$  be exchangeable,  $S$ -valued random variables. (We allow  $N_n = \infty$ .) Let  $\Xi^n$  be the corresponding empirical measure,

$$\Xi^n = \frac{1}{N_n} \sum_{i=1}^{N_n} \delta_{\xi_i^n},$$

where if  $N_n = \infty$ , we mean

$$\Xi^n = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m \delta_{\xi_i^n}.$$

Assume that  $N_n \rightarrow \infty$  and that for each  $m = 1, 2, \dots$ ,  $\{\xi_1^n, \dots, \xi_m^n\} \Rightarrow \{\xi_1, \dots, \xi_m\}$  in  $S^m$ . Then  $\{\xi_i\}$  is exchangeable and setting  $\xi_i^n = s_0 \in S$  for  $i > N_n$ ,

$$\{\Xi^n, \xi_1^n, \xi_2^n, \dots\} \Rightarrow \{\Xi, \xi_1, \xi_2, \dots\}$$

in  $\mathcal{P}(S) \times S^\infty$ , where  $\Xi$  is the de Finetti measure for  $\{\xi_i\}$ . If for each  $m$ ,

$$\{\xi_1^n, \dots, \xi_m^n\} \rightarrow \{\xi_1, \dots, \xi_m\}$$

in probability in  $S^m$ , then  $\Xi^n \rightarrow \Xi$  in probability in  $\mathcal{P}(S)$ .

The converse also holds in the sense that  $\Xi^n \Rightarrow \Xi$  implies

$$\{\xi_1^n, \dots, \xi_m^n\} \Rightarrow \{\xi_1, \dots, \xi_m\}.$$

*Proof.* The exchangeability of  $\{\xi_i\}$  follows immediately from the exchangeability of  $\{\xi_i^{N_n}\}$ . Assuming  $m+k \leq N_n$ , exchangeability implies

$$\begin{aligned} & E[f(\xi_1^n, \dots, \xi_{m+k}^n)] \\ &= E\left[\frac{1}{(N_n - m) \cdots (N_n - m - k + 1)} \sum_{\{i_1, \dots, i_k\} \subset \{m+1, \dots, N_n\}} f(\xi_1^n, \dots, \xi_m^n, \xi_{i_1}^n, \dots, \xi_{i_k}^n)\right] \\ &= E\left[\int_{S^k} f(\xi_1^n, \dots, \xi_m^n, s_1, \dots, s_k) \Xi^n(ds_1) \cdots \Xi^n(ds_k)\right] + O\left(\frac{1}{N_n}\right), \end{aligned}$$

and hence if  $f \in \bar{C}(S^{m+k})$ ,

$$\begin{aligned} & \lim_{n \rightarrow \infty} E\left[\int_{S^k} f(\xi_1^n, \dots, \xi_m^n, s_1, \dots, s_k) \Xi^n(ds_1) \cdots \Xi^n(ds_k)\right] \\ &= E[f(\xi_1, \dots, \xi_{m+k})] \\ &= E\left[\int_{S^k} f(\xi_1, \dots, \xi_m, s_1, \dots, s_k) \Xi(ds_1) \cdots \Xi(ds_k)\right], \end{aligned}$$

where the second equality follows by exchangeability. Since the space of functions on  $\mathcal{P}(S) \times S^\infty$  of the form

$$F(\mu, x_1, \dots, x_m) = \int_{S^k} f(x_1, \dots, x_m, s_1, \dots, s_k) \mu(ds_1) \cdots \mu(ds_k)$$

form a convergence determining class, the first part of the lemma follows.

If for each  $m$ ,  $\{\xi_1^n, \dots, \xi_m^n\} \rightarrow \{\xi_1, \dots, \xi_m\}$  in probability, then

$$\Xi_n^{(m)} \equiv \frac{1}{m} \sum_{i=1}^m \delta_{\xi_i^n} \rightarrow \frac{1}{m} \sum_{i=1}^m \delta_{\xi_i}$$

in probability in  $\mathcal{P}(S)$ , and the convergence of  $\Xi_n$  to  $\Xi$  follows by approximation, that is, by exchangeability, for each  $\varepsilon > 0$  and  $\varphi \in \bar{C}(S)$ ,

$$\lim_{m \rightarrow \infty} \sup_n P\{|\langle \varphi, \Xi_n^{(m)} \rangle - \langle \varphi, \Xi_n \rangle| > \varepsilon\} = 0.$$

In fact, there exist  $\eta(\varepsilon, \|\varphi\|_\infty)$  and  $C(\varepsilon, \|\varphi\|_\infty)$  such that

$$\sup_n P\{|\langle \varphi, \Xi_n^{(m)} \rangle - \langle \varphi, \Xi_n \rangle| > \varepsilon\} \leq C(\varepsilon, \|\varphi\|_\infty) e^{-\eta(\varepsilon, \|\varphi\|_\infty)m}.$$

□

We are interested in applying the above lemma in the case  $S = D_E[0, \infty)$ , the Skorohod space of  $E$ -valued cadlag functions. Then, in addition to the  $\mathcal{P}(D_E[0, \infty))$ -valued random variables  $\Xi_n$ , it is natural to consider the  $\mathcal{P}(E)$ -valued processes

$$Z_n(t) = \frac{1}{N_n} \sum_{i=1}^{N_n} \delta_{X_i^n(t)},$$

(where  $N_n$  may be infinite) which will have sample paths in  $D_{\mathcal{P}(E)}[0, \infty)$ . Unlike  $\Xi_n$ , convergence of  $Z_n$  is not always assured.

*Example 4.3* a) For  $n = 1, 2, \dots$ , let  $\{\zeta_i^n\}$  be iid uniform  $[1, 1 + \frac{1}{n}]$  random variables and let  $X_i^n = \mathbf{1}_{[\zeta_i^n, \infty)}$ . Then the conditions of Lemma 4.2 are satisfied. Note, however, that  $X^n = (X_1^n, X_2^n, \dots)$  does not converge in  $D_{\mathbf{R}^\infty}[0, \infty)$ , and the  $\mathcal{P}(\mathbf{R})$ -valued marginal process  $Z^n$  defined by  $Z^n(t) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \delta_{X_i^n(t)}$  is

$$Z_n(t) = \begin{cases} \delta_0 & t < 1 \\ (n+1-nt)\delta_0 + (nt-n)\delta_1 & 1 \leq t < 1 + \frac{1}{n} \\ \delta_1 & 1 + \frac{1}{n} \leq t \end{cases}$$

does not converge in  $D_{\mathcal{P}(\mathbf{R})}[0, \infty)$ .

b) Let  $X_1^n, X_2^n, \dots$  be iid with

$$P\{X_i^n = \mathbf{1}_{[1+\frac{1}{n}, \infty)}\} = P\{X_i^n = \mathbf{1}_{[0,1)}\} = \frac{1}{2}.$$

Then  $X^n$  converges in distribution in  $D_{\mathbf{R}}[0, \infty)^\infty$ , but not in  $D_{\mathbf{R}^\infty}[0, \infty)$  and

$$\Xi_n = \frac{1}{2}(\delta_{\mathbf{1}_{[1+\frac{1}{n}, \infty)}} + \delta_{\mathbf{1}_{[0,1)}}) \rightarrow \Xi = \frac{1}{2}(\delta_{\mathbf{1}_{[1, \infty)}} + \delta_{\mathbf{1}_{[0,1)}})$$

in  $\mathcal{P}(D_{\mathbf{R}}[0, \infty))$ . The marginal process is

$$Z_n(t) = \begin{cases} \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1 & t \notin [1, 1 + \frac{1}{n}] \\ \delta_0 & t \in [1, 1 + \frac{1}{n}] \end{cases}$$

which does not converge in  $D_{\mathcal{P}(E)}[0, \infty)$ .

The examples above show that the convergence of an exchangeable family  $X^n$  in  $D_E[0, \infty)^\infty$  does not necessarily imply that the corresponding marginal process  $Z_n$  converges in  $D_{\mathcal{P}(E)}[0, \infty)$ ; however, if we strengthen the convergence of  $X^n$  to convergence in  $D_{E^\infty}[0, \infty)$ , then convergence of  $Z_n$  follows. Recall that the distinction between convergence in the product of Skorohod spaces and convergence in the Skorohod space for the product of the state spaces has to do with coalescence of discontinuities. For example, the sequence  $x_n = (\mathbf{1}_{[1+\frac{1}{n}, \infty)}, \mathbf{1}_{[1, \infty)})$  converges in  $D_{\mathbf{R}}[0, \infty) \times D_{\mathbf{R}}[0, \infty)$  but not in  $D_{\mathbf{R}^2}[0, \infty)$ .

**Lemma 4.4** *For  $n = 1, 2, \dots$ , let  $X^n = (X_1^n, \dots, X_{N_n}^n)$  be exchangeable families of  $D_E[0, \infty)$ -valued random variables such that  $N_n \rightarrow \infty$  and  $X^n \Rightarrow X$  in  $D_E[0, \infty)^\infty$ . Define*

$$\Xi_n = \frac{1}{N_n} \sum_{i=1}^{N_n} \delta_{X_i^n}, \quad \Xi = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m \delta_{X_i} \in \mathcal{P}(D_E[0, \infty))$$

and

$$Z_n(t) = \frac{1}{N_n} \sum_{i=1}^{N_n} \delta_{X_i^n(t)}, \quad Z(t) = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m \delta_{X_i(t)} \in \mathcal{P}(E).$$

a) Let  $D_\Xi = \{t : E[\Xi\{x : x(t) \neq x(t-)\}] > 0\}$ . Then for  $t_1, \dots, t_l \notin D_\Xi$ ,

$$(\Xi_n, Z_n(t_1), \dots, Z_n(t_l)) \Rightarrow (\Xi, Z(t_1), \dots, Z(t_l)).$$

b) If  $X^n \Rightarrow X$  in  $D_{E^\infty}[0, \infty)$ , then  $(X^n, Z_n) \Rightarrow (X, Z)$  in  $D_{E^\infty \times \mathcal{P}(E)}[0, \infty)$ . If  $X^n \rightarrow X$  in probability in  $D_{E^\infty}[0, \infty)$ , then  $(X^n, Z_n) \rightarrow (X, Z)$  in probability in  $D_{E^\infty \times \mathcal{P}(E)}[0, \infty)$ .

*Remark 4.5* a) The set  $D_\Xi$  is at most countable.

b) If for  $i \neq j$ , with probability one,  $X_i$  and  $X_j$  have no simultaneous discontinuities, then  $D_\Xi = \emptyset$  and convergence of  $X^n$  to  $X$  in  $D_E[0, \infty)^\infty$  implies convergence in  $D_{E^\infty}[0, \infty)$ . In particular, this conclusion holds if the  $X_i$  are continuous.

c) The convergence result immediately implies that if  $\{X^n\}$  is relatively compact in  $D_{E^\infty}[0, \infty)$ , then  $\{(X^n, Z_n)\}$  is relatively compact in  $D_{E^\infty \times \mathcal{P}(E)}[0, \infty)$ .

*Proof.* For  $x \in D_E[0, \infty)$ , let  $\pi_t(x) = x(t)$  and recall that the continuity set for  $\pi_t$  is  $\{x : x(t) = x(t-)\}$ . By Lemma 4.2,  $\Xi_n \Rightarrow \Xi$ . The continuity set of the mapping  $\mu \in \mathcal{P}(D_E[0, \infty)) \rightarrow \mu \pi_t^{-1} \in \mathcal{P}(E)$  is the collection of  $\mu$  such that the mapping  $s \rightarrow \mu \pi_s^{-1}$  is continuous at  $t$ . The first conclusion then follows by the continuous mapping theorem.

The compact containment condition (see Ethier and Kurtz (1986), Remark 3.7.3) for  $\{Z_n\}$  follows by Lemma A1.2 in the Appendix of Donnelly and Kurtz (1996). To complete the proof we use a version of Aldous's criterion (Ethier and Kurtz (1986), Theorem 3.8.6).

For  $s < t$ , define

$$\gamma_n(\varepsilon, s, t) = \Xi_n\{x : r(x(s), x(t)) \geq \varepsilon\},$$

and observe that if  $\rho_p$  is the Prohorov metric on  $\mathcal{P}(E)$  and  $\tau_n$  is  $\sigma(\Xi_n)$  measurable, then

$$\rho_p(Z_n(\tau_n), Z_n(\tau_n + u)) \leq \gamma_n(\varepsilon, \tau_n, \tau_n + u) + \varepsilon$$

and

$$\begin{aligned} & \rho_p(Z_n(\tau_n), Z_n(\tau_n + u)) \rho_p(Z_n(\tau_n - v), Z_n(\tau_n)) \\ & \leq \gamma_n(\varepsilon, \tau_n, \tau_n + u) \gamma_n(\varepsilon, \tau_n - v, \tau_n) + 2\varepsilon + \varepsilon. \end{aligned}$$

Note that

$$\begin{aligned} & \gamma_n(\varepsilon, \tau_n, \tau_n + u) \gamma_n(\varepsilon, \tau_n - v, \tau_n) \\ & = \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m \mathbf{1}_{\{r(X_i^n(\tau_n), X_i^n(\tau_n + u)) \geq \varepsilon\}} \frac{1}{m} \sum_{j=1}^m \mathbf{1}_{\{r(X_j^n(\tau_n - v), X_j^n(\tau_n)) \geq \varepsilon\}}, \end{aligned}$$

so

$$\begin{aligned} & P\{r(X_1^n(\tau_n), X_1^n(\tau_n + u)) \geq \varepsilon, r(X_2^n((\tau_n - v) \vee 0), X_2^n(\tau_n)) \geq \varepsilon\} \\ & \geq E[\gamma_n(\varepsilon, \tau_n, \tau_n + u) \gamma_n(\varepsilon, (\tau_n - v) \vee 0, \tau_n)]. \end{aligned}$$

Let  $\mathcal{F}_T^n$  be the collection of  $\sigma(\Xi_n)$ -measurable random variables satisfying  $0 \leq \tau_n \leq T$ . The relative compactness of  $\{(X_1^n, X_2^n)\}$  in  $D_{E^2}[0, \infty)$  implies that

$$\lim_{\delta \rightarrow 0} \sup_{0 \leq u, v \leq \delta} \sup_{\tau_n \in \mathcal{F}_T^n} P\{r(X_1^n(\tau_n), X_1^n(\tau_n + u)) \geq \varepsilon, r(X_2^n((\tau_n - v) \vee 0), X_2^n(\tau_n)) \geq \varepsilon\} = 0$$

and hence

$$\lim_{\delta \rightarrow 0} \sup_{0 \leq u, v \leq \delta} \sup_{\tau_n \in \mathcal{F}_T^n} E[\rho_p(Z_n(\tau_n), Z_n(\tau_n + u)) \wedge \rho_p(Z_n((\tau_n - v) \vee 0), Z_n(\tau_n))] = 0.$$

A similar argument gives

$$\lim_{\delta \rightarrow 0} E[\rho_p(Z_n(0), Z_n(\delta))] = 0,$$

and the relative compactness of  $\{Z_n\}$  follows. By Part (a), the finite dimensional distributions converge for finite subsets of a dense set of times, so  $Z_n \Rightarrow Z$  in  $D_{\mathcal{D}(E)}[0, \infty)$ .

If  $X^n \rightarrow X$  in probability in  $D_{E^\infty}[0, \infty)$ , then for all but countably many  $t$ ,  $X^n(t) \rightarrow X(t)$  in probability so by Lemma 4.2,  $Z_n(t) \rightarrow Z(t)$  in probability. By Lemma A2.1 of Donnelly and Kurtz (1996),  $Z_n \rightarrow Z$  in probability in  $D_{\mathcal{D}(E)}[0, \infty)$ .  $\square$

**Lemma 4.6** *If  $X = (X_1, X_2, \dots)$  is an exchangeable sequence in  $D_E[0, \infty)$ , then  $Z$  is continuous if and only if for  $i \neq j$ , with probability one  $X_i$  and  $X_j$  have no simultaneous discontinuity.*

*Proof.* Necessity is immediate. Suppose, with probability one,  $X_i$  and  $X_j$  have no simultaneous discontinuity. For  $T, \eta > 0$ , let

$$\tau = T \wedge \inf\{t : \rho_p(Z(t), Z(t-)) \geq \eta\}.$$

As in the proof of Lemma 4.4,

$$\rho_p(Z(\tau - v), Z(\tau))^2 \leq \gamma(\varepsilon, \tau - v, \tau)^2 + 2\varepsilon + \varepsilon,$$

where we simplify notation by assuming that  $Z(s) = Z(0)$  for  $s < 0$  and similarly for other processes. Then, by the exchangeability,

$$\begin{aligned} & \gamma(\varepsilon, \tau - v, \tau)^2 \\ &= \lim_{m \rightarrow \infty} \frac{1}{m} \sum_{i=1}^m \mathbf{1}_{\{r(X_{2i}(\tau-v), X_{2i}(\tau)) \geq \varepsilon\}} \frac{1}{m} \sum_{i=1}^m \mathbf{1}_{\{r(X_{2i+1}(\tau-v), X_{2i+1}(\tau)) \geq \varepsilon\}}, \end{aligned}$$

so, taking expectations and interchanging limits and expectations,

$$\begin{aligned} & P\{r(X_1(\tau - v), X_1(\tau)) \geq \varepsilon, r(X_2((\tau - v)), X_2(\tau)) \geq \varepsilon\} \\ & \geq E[\gamma(\varepsilon, \tau - v, \tau)^2]. \end{aligned}$$

It follows that

$$\begin{aligned} \eta^2 P\{\tau < T\} &\leq \lim_{v \rightarrow 0} E[\gamma(\varepsilon, \tau - v, \tau)^2] + 2\varepsilon + \varepsilon^2 \\ &\leq \lim_{v \rightarrow 0} P\{r(X_1(\tau - v), X_1(\tau)) \geq \varepsilon, r(X_2((\tau - v)), X_2(\tau)) \geq \varepsilon\} + 2\varepsilon + \varepsilon^2 \\ &= 2\varepsilon + \varepsilon^2, \end{aligned}$$

where the last equality follows by the assumption that  $X_1$  and  $X_2$  have no simultaneous discontinuities. Since  $\varepsilon$  is arbitrary, the lemma follows.  $\square$

#### 4.2 Conditions for noncollision of particles

The proof of the independence of the driving noises in the limiting particle model in Section 3 depends on showing that two particles cannot collide. If we let  $Y$  denote the difference of the two particle locations, then what we need to show is that  $Y$  does not hit zero. In general,  $Y$  can be written as

$$Y(t) = Y(0) + M(t) + \int_0^t B(s) ds, \quad (4.1)$$

where  $M$  is a  $d$ -dimensional, continuous martingale with tensor quadratic variation of the form

$$[M]_t \equiv (([M_i, M_j]_t)) = \int_0^t A(s) ds, \quad (4.2)$$

where under the nondegeneracy condition, Condition 2.3,

$$\inf_{t \leq T} \inf_{z \in \mathbf{R}^d, z \neq 0} \frac{z^T A(t) z}{|z|^2} > 0. \quad (4.3)$$

The proofs of the following lemmas are similar to the proofs for diffusion processes in Chapter 11 of Friedman (1976).

**Lemma 4.7** *Let  $d \geq 3$ . Suppose that  $Y$  satisfies (4.1) with  $M$  and  $A$  satisfying (4.2) and (4.3). Assume that  $A$  and  $B$  are cadlag and that  $|Y(0)| > 0$ . For  $\delta > 0$ , let  $\tau_\delta = \inf\{t > 0 : |Y(t)| \leq \delta\}$ . Then  $\lim_{\delta \rightarrow 0} \tau_\delta = \infty$  a.s.*

*Proof.* By a truncation argument, it is enough to prove the lemma under the assumption that there exist a constant  $c > 0$  such that  $|A| + |B| \leq c$ . If this inequality does not hold, let  $\tilde{W}$  be a standard Brownian motion independent of  $M$  and  $B$ . Define  $\zeta_c = \inf\{t : \sup_{s \leq t} (|A(s)| + |B(s)|) \geq c\}$ . Set  $B_c(t) = \mathbf{1}_{[0, \zeta_c)}(t)B(t)$ ,

$$M_c(t) = M(t \wedge \zeta_c) + \tilde{W}(t \vee \zeta_c) - \tilde{W}(\zeta_c),$$

and

$$Y_c(t) = Y(0) + M_c(t) + \int_0^t B_c(s) ds. \quad (4.4)$$

Then  $Y(t) = Y_c(t)$  for  $t \leq \zeta_c$  and  $[M_c]_t = \int_0^t A_c(s) ds$ , where

$$A_c(t) = \mathbf{1}_{[0, \zeta_c)}(t)A(t) + \mathbf{1}_{[\zeta_c, \infty)}(t)I.$$

Note that  $|A_c(t)| + |B_c(t)| \leq c \vee 1$ . Since  $\lim_{c \rightarrow \infty} \zeta_c = \infty$  a.s., if the conclusion of the lemma holds for  $Y_c$  for each  $c > 1$ , then it holds for  $Y$ .

Let  $\sigma(t)$  be the symmetric, positive definite square root of  $A(t)$ . Then  $\sigma(t)$  is nonsingular, and  $\sigma^{-1}(t)A(t)\sigma^{-1}(t) = I$ . For a matrix  $C$ , let  $\lambda_{\max}(C)$  denote the largest eigenvalue of  $C$ . Recursively, define  $\beta_0 = 0$  and

$$\beta_{k+1} = \inf\{t > \beta_k : \text{tr}(\sigma^{-1}(\beta_k)A(t)\sigma^{-1}(\beta_k)) \leq 2.5\lambda_{\max}(\sigma^{-1}(\beta_k)A(t)\sigma^{-1}(\beta_k))\}.$$

Note that for  $t = \beta_k$ ,  $\text{tr}(\sigma^{-1}(\beta_k)A(t)\sigma^{-1}(\beta_k)) = d$  and

$$\frac{z^T \sigma^{-1}(\beta_k)A(t)\sigma^{-1}(\beta_k)z}{|z|^2} \leq \lambda_{\max}(\sigma^{-1}(\beta_k)A(t)\sigma^{-1}(\beta_k)) = 1,$$

so that  $\beta_{k+1} > \beta_k$ , and the assumption that  $A$  is cadlag ensures that  $\lim_{k \rightarrow \infty} \beta_k = \infty$ .

Let  $T > 0$ . Proceeding inductively, suppose that  $\tau_0 \equiv \lim_{\delta \rightarrow 0} \tau_\delta > \beta_k \wedge T$  a.s., which is certainly true for  $k = 0$ . For  $t \geq \beta_k \wedge T$ , let

$$Z_k(t) = \sigma^{-1}(\beta_k \wedge T)Y(t).$$

Then for  $\beta_k \wedge T \leq t < \tau_0 \wedge \beta_{k+1} \wedge T$ , Itô's formula implies

$$\begin{aligned}
-\log |Z_k(t)| &= -\log |Z_k(\beta_k \wedge T)| - \int_{\beta_k \wedge T}^t |Z_k(s)|^{-2} Z_k(s)^T \sigma^{-1}(\beta_k \wedge T) dM(s) \\
&\quad - \int_{\beta_k \wedge T}^t |Z_k(s)|^{-2} Z_k(s)^T \sigma^{-1}(\beta_k \wedge t) B(s) ds \\
&\quad + \frac{1}{2} \int_{\beta_k \wedge T}^t |Z_k(s)|^{-2} \left( \frac{2Z_k(s)^T \sigma^{-1}(\beta_k \wedge T) A(s) \sigma^{-1}(\beta_k \wedge T) Z_k(s)}{|Z_k(s)|^2} \right. \\
&\quad \quad \quad \left. - \text{tr}(\sigma^{-1}(\beta_k \wedge T) A(s) \sigma^{-1}(\beta_k \wedge T)) \right) ds \\
&\leq -\log |Z_k(\beta_k \wedge T)| - \int_{\beta_k \wedge T}^t |Z_k(s)|^{-2} Z_k(s)^T \sigma^{-1}(\beta_k \wedge T) dM(s) \\
&\quad - \int_{\beta_k \wedge T}^t |Z_k(s)|^{-2} Z_k(s)^T \sigma^{-1}(\beta_k \wedge t) B(s) ds \\
&\quad - \frac{1}{4} \int_{\beta_k \wedge T}^t |Z_k(s)|^{-2} \lambda_{\max}(\sigma^{-1}(\beta_k \wedge T) A(s) \sigma^{-1}(\beta_k \wedge T)) ds \\
&\leq -\log |Z_k(\beta_k \wedge T)| - \int_{\beta_k \wedge T}^t |Z_k(s)|^{-2} Z_k(s)^T \sigma^{-1}(\beta_k \wedge T) dM(s) \\
&\quad + \int_{\beta_k \wedge T}^t (b|Z_k(s)|^{-1} - a|Z_k(s)|^{-2}) ds,
\end{aligned}$$

where  $b = \sup_{u,v} |\sigma^{-1}(u)| |B(v)| < \infty$  and  $a = \frac{1}{4} \inf_{u,v} \lambda_{\max}(\sigma^{-1}(u) A(v) \sigma^{-1}(u)) > 0$ . Since  $\sup_{z>0} (bz^{-1} - az^{-2}) = \frac{b^2}{4a}$ ,

$$\limsup_{\delta \rightarrow 0} E[-\log |Z_k(\beta_{k+1} \wedge T \wedge \tau_\delta)|] \leq E[-\log |Z_k(\beta_k \wedge T)|] + \frac{b^2 T}{4a},$$

and hence  $\lim_{\delta \rightarrow 0} \tau_\delta > \beta_{k+1} \wedge T$  a.s. Since  $T$  and  $k$  are arbitrary,  $\lim_{\delta \rightarrow 0} \tau_\delta = \infty$  a.s.  $\square$

Note that we used the assumption that  $d \geq 3$  in the previous lemma to ensure that  $\beta_{k+1} > \beta_k$  and  $\beta_k \rightarrow \infty$ . Similar results in the case  $d = 2$  require some regularity in  $A$  or equivalently  $\sigma^{-1}$ .

We define  $W(t) = \int_0^t \sigma^{-1}(s) dM(s)$ , and note that  $W$  is a two-dimensional, standard Brownian.

**Lemma 4.8** *Let  $d = 2$ . Suppose that  $Y$  satisfies (4.1) with  $M$  and  $A$  satisfying (4.2) and (4.3). Suppose that  $\gamma(t) \equiv \sigma^{-1}(t)$  satisfies*

$$\gamma(t) = \gamma(0) + \int_0^t \Gamma_1(s) dW_1(s) + \int_0^t \Gamma_2(s) dW_2(s) + \int_0^t C(s) ds, \quad (4.5)$$

where  $\Gamma_1, \Gamma_2, C$  are progressive,  $2 \times 2$ -matrix-valued processes, satisfying

$$\sup_{s \leq t} (|\Gamma_1(s)| + |\Gamma_2(s)| + |C(s)|) < \infty \quad a.s.$$

for each  $t > 0$ . Assume that  $|Y(0)| > 0$ . For  $\delta > 0$ , let  $\tau_\delta = \inf\{t > 0 : |Y(t)| \leq \delta\}$ . Then  $\lim_{\delta \rightarrow 0} \tau_\delta = \infty$  a.s.

*Proof.* As in the proof of Lemma 4.7, we can assume that  $|A(t)|$ ,  $|B(t)|$ ,  $|C(t)|$ ,  $|\Gamma_1(t)|$ , and  $|\Gamma_2(t)|$  are all bounded by a constant  $c$ . Define  $Z(t) = \gamma(t)Y(t)$ . Then setting  $\sigma(t) = (\sigma_1(t), \sigma_2(t))$  and

$$R(t) = (\Gamma_1(s)Y(s), \Gamma_2(s)Y(s)), \quad Q(t) = \Gamma_1(t)\sigma_1(t) + \Gamma_2(t)\sigma_2(t),$$

$$\begin{aligned} Z(t) &= Z(0) + \int_0^t \gamma(s)dY(s) + \int_0^t d\gamma(s)Y(s) + \begin{pmatrix} [\gamma_{11}, Y_1]_t + [\gamma_{12}, Y_2]_t \\ [\gamma_{21}, Y_1]_t + [\gamma_{22}, Y_2]_t \end{pmatrix} \\ &= Z(0) + W(t) + \int_0^t \gamma(s)B(s)ds + \int_0^t R(s)dW(s) \\ &\quad + \int_0^t C(s)Y(s)ds + \int_0^t Q(s)ds. \end{aligned}$$

As in Friedman (1976), page 253, let  $\varphi(r)$  satisfy  $\varphi'(r) = -r^{-1}e^{r^\theta/\theta}$  for some  $0 < \theta < 1$ . Then

$$\begin{aligned} \varphi(|Z(t)|) &= \varphi(|Z(0)|) - \int_0^t |Z(s)|^{-2}e^{\theta^{-1}|Z(s)|^\theta} Z(s)^T (I + R(s))dW(s) \\ &\quad - \int_0^t |Z(s)|^{-2}e^{\theta^{-1}|Z(s)|^\theta} Z(s)^T (\gamma(s)B(s) + C(s)Y(s) + Q(s))ds \\ &\quad - \int_0^t \frac{1}{2} |Z(s)|^{-2}e^{\theta^{-1}|Z(s)|^\theta} \left( \text{tr}(I + R(s))(I + R(s)^T) \right. \\ &\quad \left. - (2 - |Z(s)|^\theta)|Z(s)|^{-2}Z(s)^T (I + R(s))(I + R(s)^T)Z(s) \right) ds \\ &= \varphi(|Z(0)|) - \int_0^t |Z(s)|^{-2}e^{\theta^{-1}|Z(s)|^\theta} Z(s)^T (I + R(s))dW(s) \\ &\quad - \int_0^t |Z(s)|^{-2}e^{\theta^{-1}|Z(s)|^\theta} Z(s)^T (\gamma(s)B(s) + C(s)Y(s) + Q(s))ds \\ &\quad - \int_0^t \frac{1}{2} |Z(s)|^{-2}e^{\theta^{-1}|Z(s)|^\theta} \left( \text{tr}(R(s) + R(s)^T + R(s)R(s)^T) \right. \\ &\quad \left. - 2|Z(s)|^{-2}Z(s)^T (R(s) + R(s)^T + R(s)R(s)^T)Z(s) \right. \\ &\quad \left. + |Z(s)|^\theta |Z(s)|^{-2}Z(s)^T (I + R(s))(I + R(s)^T)Z(s) \right) ds \\ &\leq \varphi(|Z(0)|) - \int_0^t |Z(s)|^{-2}e^{\theta^{-1}|Z(s)|^\theta} Z(s)^T (I + R(s))dW(s) \\ &\quad + \int_0^t e^{\theta^{-1}|Z(s)|^\theta} (c_0 + c_1|Z(s)|^{-1} - c_2|Z(s)|^{-(2-\theta)})ds, \end{aligned}$$

and hence, setting  $\beta_a = \inf\{t : |Z(t)| \geq a\}$ , there exists a constant  $\alpha(a, c_0, c_1, c_2, \theta)$  such that

$$E[\varphi(|Z(\tau_\delta \wedge \beta_a \wedge T)|)] \leq E[\varphi(|Z(0)|)] + \alpha(a, c_0, c_1, c_2, \theta)T.$$

It follows that  $\lim_{\delta \rightarrow 0} \tau_\delta > \beta_a \wedge T$  a.s. for all  $a$  and  $T$ , and the lemma follows.  $\square$

### 4.3 Comments on previous work

Vaillancourt (1988) considers stochastic ordinary differential equations (SODEs) for the displacement of  $r^i$  of the following type<sup>6</sup>:

$$\begin{aligned} dr_N^i &= F(r_N^i(t), X_N(t), t)dt + \sum_{n=1}^N \sigma_n(r_N^i(t), N, X_N(t), t)d\beta_n(t) \\ r^i(s) &= q^i, i = 1, \dots, N, \quad X_N(t) := \sum_{i=1}^N \frac{1}{N} \delta_{r_N^i(t)}, \end{aligned} \quad (4.6)$$

where  $\sigma_n(r, \mu, t)$  are “nice”  $\mathcal{M}_{d \times d}$  valued functions. Adjusting the notation used in (1.7) to the setting of (4.6), the two-particle diffusion matrix is given by

$$\tilde{D}(r^i, r^j, N, \mu, t) := \sum_{n=1}^N \sigma_n(r^i, N, \mu, t) \sigma_n^T(r^j, N, \mu, t) \quad \forall i, j = 1, \dots, N. \quad (4.7)$$

As in (1.4), the quadratic variations does not converge to zero, and accordingly, the limit  $N \rightarrow \infty$  leads to the solution of a quasilinear SPDE. The main difference from the model (1.1)-(1.5) is that in Vaillancourt’s case the noise is changing with the number of particles. Hence, there is not a unique SPDE for all  $N$ , and an SPDE is only obtained for the limiting process as  $N \rightarrow \infty$ . See also Dawson and Vaillancourt (1995).

*Remark 4.9* The empirical process for independent branching Brownian motions is approximated by the measure valued Dawson-Watanabe process under suitable assumptions on the branching rate. This process can be represented as the solution of a “formal” SPDE in the space of measures (Dawson (1975)). Dawson’s work initiated the new research area of superprocesses (see, for example, Dawson (1993)). In dimension  $d = 1$  the formal SPDE for the Dawson-Watanabe process becomes a solvable SPDE for the density of the measure process. (See Konno and Shiga (1988).) The models considered by Vaillancourt (1988) and Kotelenez (1995) were motivated by Dawson’s work and ideas. Assuming spatial correlations for a different “physical” model, well-posed SPDEs for particle distributions were obtained in any dimension in these papers.

Borkar (1984) uses a Gaussian random field, called “Brownian medium”, as a driving term for stochastic ordinary differential equations (SODEs). The approach in Kunita (1990) (and the references therein) is similar, but he goes beyond Borkar’s work by considering flows of SODEs and bilinear SPDEs generated by those flows.

### 4.4 Kunita’s random field

Kunita (1990), Section 4.2, considers stochastic differential equations driven by a  $\mathbf{R}^k$ -valued, Gaussian, space-time random field  $S(r, t)$  that Kunita refers to as a  $C$ -valued

<sup>6</sup> For SODEs with coefficients  $F$  and  $\sigma_{ij}$ , independent of the empirical distribution and driven by finitely many Brownian motions. See, for example, Gihman and Skorohod (1974) as well as Ikeda and Watanabe (1989).

Brownian motion.  $S$  has mean and covariance of the form

$$E[S(r,t)] = \int_0^t b(r,u)du, \quad \text{Cov}(S(r^1,t), S(r^2,s)) = \int_0^{t \wedge s} a(r^1, r^2, u)du. \quad (4.8)$$

Under appropriate regularity assumptions, Kunita analyzes stochastic equations of the form

$$\psi_{s,t}(z) = z + \int_s^t S(\psi_{s,u}(z), du), \quad (4.9)$$

where the integral is defined by

$$\int_s^t S(\psi_{s,u}(z), du) = \lim_{\max |u_{i+1} - u_i| \rightarrow 0} \sum (S(\psi_{s,u_i}(z), u_{i+1}) - S(\psi_{s,u_i}(z), u_i)), \quad (4.10)$$

for  $s = u_0 < \dots < u_m = t$ . For example, if

$$S(r,t) = \int_0^t b(r,u)du + \sum_{n=1}^N \int_0^t \sigma_n(r,u) d\beta_n(u),$$

for independent, standard Brownian motions  $\beta_n$ , then

$$a(r^1, r^2, u) = \sum_{n=1}^N \sigma_n(r^1, u) \sigma_n^T(r^2, u)$$

and (4.9) becomes the Itô equation

$$\psi_{s,t}(z) = z + \int_s^t b(\psi_{s,u}(z), u)du + \sum_{n=1}^N \int_s^t \sigma_n(\psi_{s,u}(z), u) d\beta_n(u).$$

Similarly, if

$$S(r,t) := \int_0^t F(r,u)du + \int_0^t \int \mathcal{J}(r,q,u) w(dq, du), \quad (4.11)$$

then

$$b(r,u) = F(r,u), \quad a(r^1, r^2, u) = \int \mathcal{J}(r^1, q, u) \mathcal{J}^T(r^2, q, u) dq, \quad (4.12)$$

and (4.9) becomes

$$\psi_{s,t}(z) = z + \int_s^t F(\psi_{s,u}(z), u)du + \int_s^t \int \mathcal{J}(\psi_{s,u}(z), q, u) w(dq, du). \quad (4.13)$$

Assuming existence and uniqueness for all  $s$  and  $z$ , Kunita calls the solution of (4.9) a *forward Brownian flow*. Suitable Lipschitz and linear growth assumptions ensure the desired existence and uniqueness.<sup>7</sup> Following Kunita's notation and terminology, define the Lipschitz norms of the infinitesimal mean  $b(r,t)$  and infinitesimal

<sup>7</sup> Note that in Borkar's and Kunita's approaches the coefficients of the SODEs do not depend on the (empirical) distribution of the particles. Hence, their work represents a new approach to classical SODEs using a different and a more general Brownian noise. (See Kotelenetz (2007) and Proposition 4.10.)

covariance  $a(r^1, r^2, t)$ :

$$\|b(t)\|_{0,1} := \sup_{r \neq q} \frac{|b(r^1, t) - b(r^2, t)|}{|r^1 - r^2|} \quad (4.14)$$

$$\|a(t)\|_{0,1} := \sup_{r^1 \neq r^2, \tilde{r}^1 \neq \tilde{r}^2} \frac{|a(r^1, \tilde{r}^1, t) - a(r^2, \tilde{r}^1, t) - a(r^1, \tilde{r}^2, t) + a(r^2, \tilde{r}^2, t)|}{|r^1 - r^2| |\tilde{r}^1 - \tilde{r}^2|}.$$

Under the additional assumptions

$$|b(r, t)| \leq c(1 + |r|), \quad \|a(r^1, r^2, t)\| \leq c(1 + |r^1|)(1 + |r^2|) \quad (\text{linear growth}),$$

$$\sup_{t \geq 0} [\|b(t)\|_{0,1} + \|a(t)\|_{0,1}] \leq c \quad (\text{globally Lipschitz}), \quad (4.15)$$

Kunita (loc. cit. Section 4.2, Theorem 4.2.5) constructs the forward Brownian flow  $\psi_{s,t}(z)$  as the unique solution of the stochastic ordinary differential equation (4.9).

The linear growth condition, Condition 2.2, is similar to Kunita's Condition (4.14), except that we do not assume the coefficient to be continuous in  $t$ . Since we solve our SODEs on a finite time interval, our assumption of uniform boundedness in  $t$  is the same as in Kunita's Theorem 4.2.5, and it can easily be removed by working with localizing stopping times. Further, let us comment on the Lipschitz assumption in Kunita. Since this is an assumption on the components of the matrices involved we may without loss of generality assume that the coefficients are real valued. Condition 2.1 (for the Euclidean norm, uniformly in  $t$ ) then implies, in the notation of Kunita, that

$$|a(r^1, r^1, t) + a(r^2, r^2, t) - 2a(r^1, r^2, t)| \leq c|r^1 - r^2|^2, \quad (4.16)$$

which is apparently weaker than Kunita's continuity and Lipschitz assumption (4.15) on  $a(r^1, r^2, t)$ .

It remains to show that Kunita's  $C$ -valued Brownian motion  $S(r, t)$  can be represented as in (4.11). Since a Gaussian random field (random function) is uniquely determined by its mean and covariance, we have just to show that the class of infinitesimal means and covariances from (4.12) contains those considered by Kunita.

The case for the mean is trivial. Therefore, we will without loss of generality assume that  $b = 0$ . Fix  $t$ , and first consider the case of real-valued random fields. Set  $\mathbf{B} := C(\mathbf{R}^d; \mathbf{R})$  and endow  $\mathbf{B}$  with the topology of uniform convergence on bounded sets. Let  $\mathcal{B}$  be the Borel  $\sigma$ -algebra on  $\mathbf{B}$ . Given  $a(r, q, t)$  from (4.8), Kolmogorov's theorem implies the existence of a Gaussian random field with covariance  $a(r, q, t)$ . (See Gihman and Skorohod (1974), Chapter 3.1.) A multiparameter version of Kolmogorov's regularity theorem implies the existence of a regular version with sample paths in  $C(\mathbf{R}^{d+1}; \mathbf{R})$ . (See Kunita (1990), Section 1.4, Ibragimov (1984), and Lifshits (1995) Section 15, for general regularity results.) Hence, for fixed  $t$ , the Gaussian measure  $\mu$  on the space of functions from  $\mathbf{R}^d$  with values in  $\mathbf{R}$  is supported by  $\mathbf{B}$ , endowed with  $\mathcal{B}$ . (It is actually supported by an even smaller space of more regular functions. Recall the classical Wiener space and the regularity of the standard Brownian motion expressed by Levy's modulus of continuity.) Thus, we have the existence of a (continuous) Gaussian random field  $\xi(r, t)$  such that

$$a(r^1, r^2, t) = E(\xi(r^1, t)\xi(r^2, t)) = \int_{\mathbf{B}} \zeta(r^1)\zeta(r^2)\mu(d\zeta). \quad (4.17)$$

Note that  $\mathbf{B}$  is a Polish space. Hence, any measure on  $(\mathbf{B}, \mathcal{B})$  is a Radon measure. (See Bourbaki (1969), Ch. IX, Section 3.3.) Therefore, by Corollary 3.2.7 in Ch. 3.2 of Bogachev (1997),  $L_2(\mathbf{B}, \mathcal{B}, d\mu)$  is separable. Denoting Lebesgue measure by  $dp$ , the separability of  $L_2(\mathbf{B}, \mathcal{B}, d\mu)$  implies that

$$L_2(\mathbf{B}, \mathcal{B}, d\mu) \cong L_2(\mathbf{R}^d, \mathcal{B}^d, dp) \quad (\text{isometrically isomorphic}). \quad (4.18)$$

Note that for the special case of a separable Banach space  $\mathbf{B}$ , the separability of  $L_2(\mathbf{B}, \mathcal{B}, d\mu)$  follows from the Itô-Wiener chaos decomposition by construction of a countable complete orthonormal system of eigenfunctions for  $L_2(\mathbf{B}_N, \mathcal{B}_N, du_N)$ . (See Watanabe (1984), Section 1.2, or Gihman and Skorohod (1974), Ch. VIII.) The relation (4.18) immediately generalizes to  $\mathbf{R}^d$ -valued Gaussian random fields. Thus, there are Borel measurable,  $\mathbf{R}^d$ -valued functions  $\mathcal{J}(r, p, t)$ , whose one-dimensional components are square integrable in  $p$  and continuous in  $(r, t)$  such that<sup>8</sup>

$$\forall t, r^1, r^2 \quad a(r^1, r^2, t) = E(\xi(r^1, t)\xi^T(r^2, t)) = \int_{\mathbf{R}^d} \mathcal{J}(r^1, q, t) \mathcal{J}^T(r^2, q, t) dq. \quad (4.19)$$

Since the distribution of Gaussian fields are uniquely determined by their mean and their covariance operator we have the following proposition.

**Proposition 4.10** *S satisfying (4.8) can be represented as*

$$S(r, t) = \int_0^t F(r, u) du + \int_{\mathbf{R}^d} \mathcal{J}(r, q, u) w(dq, du). \quad (4.20)$$

Hence, (4.20) allows us to represent Kunita's Gaussian random fields directly as the sum of a deterministic integral and a stochastic integral, where the latter is driven by a standard Gaussian space-time white noise, as in our set-up of Section 1. Recalling Kunita's stronger regularity assumptions (4.15) on the mean and covariance, it follows that the Gaussian random fields which drive the stochastic differential equations of Kunita (1990) are essentially a special case of the fields given in (4.11).

*Remark 4.11* For Gaussian random fields with covariance as in the right side of (4.19), Gihman and Skorohod (1974) Chapter 4.5, Theorem 1 gives the representation (4.20).

**Acknowledgements** The final presentation of the paper has profited from careful refereeing.

## References

- de Acosta A (1982) Invariance principles in probability for triangular arrays of  $B$ -valued random vectors and some applications. *Ann Probab* 10(2):346–373  
 Asakura S, Oosawa F (1954) On interaction between two bodies immersed in a solution of macromolecules. *J Chem Phys* 22(7):1255–1256  
 Bogachev VI (1997) *Gaussovskie mery*. Fizmatlit “Nauka”, Moscow  
 Borkar VS (1984) Evolution of interacting particles in a Brownian medium. *Stochastics* 14(1):33–79

<sup>8</sup> (4.19) was conjectured by Dorogovtsev (2004), private communication).

- Bourbaki N (1969) *Éléments de mathématique*. Fasc. XXXV. Livre VI: Intégration. Chapitre IX: Intégration sur les espaces topologiques séparés. *Actualités Scientifiques et Industrielles*, No. 1343, Hermann, Paris
- Dawson D, Vaillancourt J (1995) Stochastic McKean-Vlasov equations. *NoDEA Nonlinear Differential Equations Appl* 2(2):199–229
- Dawson DA (1975) Stochastic evolution equations and related measure processes. *J Multivariate Anal* 5:1–52
- Dawson DA (1993) Measure-valued Markov processes. In: *École d'Été de Probabilités de Saint-Flour XXI—1991*, Lecture Notes in Math., vol 1541, Springer, Berlin, pp 1–260
- Donnelly P, Kurtz TG (1996) A countable representation of the Fleming-Viot measure-valued diffusion. *Ann Probab* 24(2):698–742
- Ethier SN, Kurtz TG (1986) *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons Inc., New York, characterization and convergence
- Friedman A (1976) *Stochastic differential equations and applications*. Vol. 2. Academic Press [Harcourt Brace Jovanovich Publishers], New York, probability and Mathematical Statistics, Vol. 28
- Gärtner J (1988) On the McKean-Vlasov limit for interacting diffusions. *Math Nachr* 137:197–248
- Gihman ĪĪ, Skorohod AV (1974) *The theory of stochastic processes*. I. Springer-Verlag, New York, translated from the Russian by S. Kotz, *Die Grundlehren der mathematischen Wissenschaften*, Band 210
- Götzelmann B, Evans R, Dietrich S (1998) Depletion forces in fluids. *Phys Rev E* 57(6):6785–6800, DOI 10.1103/PhysRevE.57.6785
- Haken H (1983) *Advanced Synergetics*. Springer-Verlag, Berlin
- Ibragimov IA (1984) On smoothness conditions for trajectories of random functions. *Theory of Probability and its Applications* 28(2):240–262, DOI 10.1137/1128023, URL <http://link.aip.org/link/?TPR/28/240/1>
- Ikeda N, Watanabe S (1989) *Stochastic differential equations and diffusion processes*, North-Holland Mathematical Library, vol 24, 2nd edn. North-Holland Publishing Co., Amsterdam
- van Kampen NG (1981) *Stochastic processes in physics and chemistry*. North-Holland Publishing Co., Amsterdam, *lecture Notes in Mathematics*, 888
- Konno N, Shiga T (1988) Stochastic partial differential equations for some measure-valued diffusions. *Probab Theory Related Fields* 79(2):201–225
- Kotelenez PM (1995) A class of quasilinear stochastic partial differential equations of McKean-Vlasov type with mass conservation. *Probab Theory Related Fields* 102(2):159–188
- Kotelenez PM (2005a) Correlated Brownian motions as an approximation to deterministic mean-field dynamics. *Ukrain Mat Zh* 57(6):757–769
- Kotelenez PM (2005b) From discrete deterministic dynamics to Brownian motions. *Stoch Dyn* 5(3):343–384
- Kotelenez PM (2007) *Stochastic Ordinary and Stochastic Partial Differential Equations*. Springer-Verlag, Berlin
- Kotelenez PM, Leitman MJ, Jr JAM (2007) On the depletion effect on colloids, submitted
- Kunita H (1990) *Stochastic flows and stochastic differential equations*, Cambridge Studies in Advanced Mathematics, vol 24. Cambridge University Press, Cambridge
- Kurtz TG, Protter P (1991) Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann Probab* 19(3):1035–1070
- Kurtz TG, Protter PE (1996) Weak convergence of stochastic integrals and differential equations. II. Infinite-dimensional case. In: *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, Lecture Notes in Math., vol 1627, Springer, Berlin, pp 197–285
- Kurtz TG, Xiong J (1999) Particle representations for a class of nonlinear SPDEs. *Stochastic Process Appl* 83(1):103–126
- Lifshits MA (1995) *Gaussian random functions*, Mathematics and its Applications, vol 322. Kluwer Academic Publishers, Dordrecht
- Oelschläger K (1984) A martingale approach to the law of large numbers for weakly interacting stochastic processes. *Ann Probab* 12(2):458–479
- Spohn H (1991) *Large scale dynamics of interacting particles*. Springer-Verlag, Berlin
- Vaillancourt J (1988) On the existence of random McKean-Vlasov limits for triangular arrays of exchangeable diffusions. *Stochastic Anal Appl* 6(4):431–446
- Walsh JB (1986) An introduction to stochastic partial differential equations. In: *École d'été de probabilités de Saint-Flour, XIV—1984*, Lecture Notes in Math., vol 1180, Springer, Berlin, pp 265–439

- 
- Wang H (1995) Interacting branching particle systems and superprocesses. PhD thesis, Carleton University, Ottawa, ON, Canada
- Watanabe S (1984) Lectures on stochastic differential equations and Malliavin calculus, Tata Institute of Fundamental Research Lectures on Mathematics and Physics, vol 73. Published for the Tata Institute of Fundamental Research, Bombay, notes by M. Gopalan Nair and B. Rajeev