

Stochastics and Dynamics
© World Scientific Publishing Company

SPATIAL DYNAMICS IN INTERACTING SYSTEMS WITH DISCONTINUOUS COEFFICIENTS AND THEIR CONTINUUM LIMITS

GIOVANNI ZANCO

*Dipartimento di Economia e Finanza, Luiss University
Viale Romania 32, 00197 Roma, Italy
gzanco@luiss.it*

Received (Day Month Year)

Revised (Day Month Year)

We consider a discrete model in which particles are characterized by two quantities X and Y ; both quantities evolve in time according to stochastic dynamics and the equation that governs the evolution of Y is also influenced by mean-field interaction between particles. We allow for discontinuous coefficients and random initial condition and, under suitable assumptions, we prove that in the limit as the number of particles grows to infinity the dynamics of the system is described by the solution of a Fokker-Planck partial differential equation. We provide existence and uniqueness of a solution to the latter and show that such solution arises as the limit in probability of the empirical measures of the system.

Keywords: Mean-field interaction; Fokker-Planck equation; SDEs with discontinuous coefficients.

AMS Subject Classification: 60H10, 35Q84, 92C20, 62M45

1. Introduction

Interacting stochastic systems are now ubiquitous in many applications of mathematics to physics, biology, engineering, economics, human sciences; in particular the theory of mean field limits of interacting systems has been and is extensively used to model real life phenomena, and its investigation in the last years has provided significant advances in their mathematical comprehension. Research on applications to economics and neuroscience have motivated an analysis of interactions that depend at the same time on some measurable quantity characteristic of the system at hand and on the position in the physical space of the interacting particles (or agents in the economic interpretation).

In [1] one type of such interaction was considered, without spatial dynamics but with discontinuous coefficients. Here we show how the technique presented therein allows also to deal with a system of particles whose position evolves in time and whose interaction depends both on their position X and on some quantity Y , that could represent, for example, electric activity in a space-time evolving electrical network or available capital in a network of interacting economical agents. We model X and Y as variables in \mathbb{R}^d and \mathbb{R}^e , respectively, and we suppose that they evolve in time following stochastic dynamics with discontinuous drift. A particular type of discontinuity is allowed also in the interac-

2 *Giovanni Zanco*

tion term, see Section 2 for precise assumptions. We prove that as the number of particles goes to infinity the system is described by the solution of a Fokker-Planck partial differential equation with discontinuous coefficients, whose solution has however some regularity (actually the same as the law of the initial conditions). We not only show existence and uniqueness of a solution to the Fokker-Planck equation but we also rigorously prove that it arises as the continuum limit of the particle system. The crucial assumption is that the sets of discontinuities of the coefficients have Lebesgue measure zero; this ensures that they will also have zero measure with respect to the solution of the Fokker-Planck equation.

We will state our main result in Section 3 and we will prove it in Sections 4 and 5.

The classical references for the theory of Fokker-Planck equations (or their McKean-Vlasov interpretation) corresponding to interacting particle systems are [7] and [3]; we essentially follow the strategy outlined in the latter, using also some techniques derived from [5] for studying the empirical density.

2. Setting of the problem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and for $n \in \mathbb{N}$ and $i = 1, \dots, n$ consider the system in $\mathbb{R}^d \times \mathbb{R}^e$ given by

$$\left\{ \begin{array}{l} dX^{i,n}(t) = b_1(X^{i,n}(t), Y^{i,n}(t)) dt + \sigma_1(X^{i,n}(t)) dW_1^i(t) \\ X^{i,n}(0) = X_0^i \\ dY^{i,n}(t) = b_2(X^{i,n}(t), Y^{i,n}(t)) dt \\ \quad + \frac{1}{n} \sum_{j=1}^n g_2(X^{i,n}(t), Y^{i,n}(t); X^{j,n}(t), Y^{j,n}(t)) dt \\ \quad + \sigma_2(Y^{i,n}(t)) dW_2^i(t) \\ Y^{i,n}(0) = Y_0^i \end{array} \right. \quad (2.1)$$

where

- (i) $b_1: \mathbb{R}^d \times \mathbb{R}^e \rightarrow \mathbb{R}^d$ and $b_2: \mathbb{R}^d \times \mathbb{R}^e \rightarrow \mathbb{R}^e$ are bounded functions that are allowed to be discontinuous on some sets F_1 and F_2 , respectively, each with Lebesgue measure zero;
- (ii) $\sigma_1: \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ and $\sigma_2: \mathbb{R}^e \rightarrow \mathbb{R}^{e \times e}$ are C_b^1 positive functions such that both $\sigma_1 \sigma_1^*$ and $\sigma_2 \sigma_2^*$ are uniformly elliptic;
- (iii) $g_2: \mathbb{R}^d \times \mathbb{R}^e \times \mathbb{R}^d \times \mathbb{R}^e \rightarrow \mathbb{R}^e$ is the product of three functions $\theta: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ and $\eta_1, \eta_2: \mathbb{R}^e \times \mathbb{R}^e \rightarrow \mathbb{R}^e$, i.e.

$$g_2(x, y, \hat{x}, \hat{y}) = \theta(x, \hat{x}) \eta_1(y) \eta_2(\hat{y}),$$

with θ bounded and uniformly continuous and η_1, η_2 bounded and discontinuous each on a set of Lebesgue measure zero;

- (iv) $X_0^i, i \in \mathbb{N}$, are \mathbb{R}^d -valued i.i.d. random variables with a law λ_1 that is absolutely continuous with respect to the Lebesgue measure and has an L^2 density;

- (v) $Y_0^i, i \in \mathbb{N}$, are \mathbb{R}^e -valued i.i.d. random variables, independent of $\{X_0^i\}$, with a law λ_2 that is absolutely continuous with respect to the Lebesgue measure and has an L^2 density;
- (vi) $W_1^i, i \in \mathbb{N}$, are independent d -dimensional Brownian motions, independent also of $\{X_0^i\}$ and $\{Y_0^i\}$;
- (vii) $W_2^i, i \in \mathbb{N}$, are independent e -dimensional Brownian motions, independent also of $\{W_1^i\}, \{X_0^i\}$ and $\{Y_0^i\}$.

Under these assumptions it is well known that, for every fixed $N \in \mathbb{N}$, system (2.1) has a strong solution, that is unique in the sense of pathwise uniqueness (this can be proved for example using the techniques in [8] or [2]; see also [4]).

Remark 2.1. One can consider also different sets of assumptions for the system (2.1); for example the results herein apply with simple modifications also in the following cases:

- (a) $b_1 \equiv 0, \sigma_1 \equiv 0$; this case was treated in [1];
- (b) $\sigma_1 \equiv 0$ and b_1 Lipschitz with sublinear growth;
- (c) $g_2(x, y; \hat{x}, \hat{y}) = \theta(x, \hat{x})\eta(y, \hat{y})$ where η is such that for every y the set of discontinuity of $\eta(y, \cdot)$ has Lebesgue measure zero and for each \hat{y} the function $\eta(\cdot, \hat{y})$ is uniformly continuous.

Remark 2.2. Our results are also valid when system (2.1) is to be solved in $D \times E$ with D (resp. E) a d - (resp. e -) dimensional cube with periodic boundary conditions and the coefficients also share the required periodicity. In this case one has to choose properly the set of test functions and the family of mollifiers used in proposition 5.3.

For a function $f: \mathbb{R}^d \times \mathbb{R}^e \rightarrow \mathbb{R}$, $f(x, y) = f(x_1, \dots, x_d, y_1, \dots, y_e)$ we will denote by $\partial_x f$ the gradient of f with respect to the variable $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ and by $\partial_y f$ its gradient with respect to the variable $y = (y_1, \dots, y_e) \in \mathbb{R}^e$. Similarly $\text{div}_y f$ will denote the divergence of f with respect to y . The symbols ∇, ∇^2 and div will instead be used when differentiation is performed with respect to all variables. We will use the symbol $\langle \cdot, \cdot \rangle$ for the duality product between probability measures and test functions. Finally we will sometimes write $a \lesssim b$ if there exists $C \geq 0$ such that $a \leq Cb$.

3. Main result

If we denote by $S^{(n)}(t)$ the empirical measure of the solution $(X^{i,n}(t), Y^{i,n}(t))_{i=1, \dots, n}$ to system (2.1) at times t , namely

$$S^{(n)}(t) = \frac{1}{N} \sum_{i=1}^n \delta_{X^{i,n}(t), Y^{i,n}(t)},$$

by $W^i, i = 1, \dots, n$, the $d + e$ -dimensional Brownian motions

$$W^i(t) = (W_1^i(t), W_2^i(t))$$

4 *Giovanni Zanco*

and introduce, just for convenience of notation, the functions

$$\sigma(x, y) = \begin{pmatrix} \sigma_1(x) & 0 \\ 0 & \sigma_2(y) \end{pmatrix}$$

$$b(x, y) = (b_1(x, y), b_2(x, y))$$

we then have, for any test function $\varphi \in C_b^2(\mathbb{R}^d \times \mathbb{R}^e)$ with compact support,

$$\begin{aligned} d\langle S^{(n)}(t), \varphi \rangle &= \\ &= \frac{1}{n} \sum_{i=1}^n \nabla \varphi(X^{i,n}(t), Y^{i,n}(t)) \cdot b(X^{i,n}(t), Y^{i,n}(t)) dt \\ &\quad + \frac{1}{n} \sum_{i=1}^n \partial_y \varphi(X^{i,n}(t), Y^{i,n}(t)) \frac{1}{n} \sum_{j=1}^n g_2(X^{i,n}(t), Y^{i,n}(t), X^{j,n}(t), Y^{j,n}(t)) dt \\ &\quad + \frac{1}{n} \sum_{i=1}^n \nabla \varphi(X^{i,n}(t), Y^{i,n}(t)) \cdot \sigma(X^{i,n}(t), Y^{i,n}(t)) dW^i(t) \\ &\quad + \frac{1}{2n} \sum_{i=1}^n \text{Tr} \left[\sigma(X^{i,n}(t), Y^{i,n}(t)) \sigma(X^{i,n}(t), Y^{i,n}(t))^* \nabla^2 \varphi(X^{i,n}(t), Y^{i,n}(t)) \right] dt \\ &= \langle S^{(n)}(t), b \cdot \nabla \varphi \rangle dt + \langle S^{(n)}(t), \langle S^{(n)}(t), g_2(x, y, \cdot, \cdot) \rangle \partial_y \varphi \rangle dt \\ &\quad + \langle S^{(n)}(t), \frac{1}{2} \text{Tr} [\sigma \sigma^* \nabla^2 \varphi] \rangle dt + dM_\varphi^{(n)}(t); \end{aligned} \quad (3.1)$$

here we used the notation

$$\begin{aligned} &\langle S^{(n)}(t), \langle S^{(n)}(t), g_2(x, y, \cdot, \cdot) \rangle \partial_y \varphi \rangle \\ &= \int_{D \times E} \partial_y \varphi(x, y) S^{(n)}(t)(dx, dy) \int_{D \times E} g_2(x, \hat{x}, y, \hat{y}) S^{(n)}(t)(d\hat{x}, d\hat{y}). \end{aligned} \quad (3.2)$$

Thanks to the independence of the W^i 's and to the boundedness of σ and $\nabla \varphi$, it is not difficult to show that the term

$$M_\varphi^{(n)}(t) = \int_0^t \frac{1}{n} \sum_{i=1}^n \sigma(X^{i,n}(s), Y^{i,n}(s)) \nabla \varphi(X^{i,n}(s), Y^{i,n}(s)) dW^i(s) \quad (3.3)$$

is a martingale satisfying $\mathbb{E}[\sup_t |M_\varphi^{(n)}(t)|^2] \rightarrow 0$ as $n \rightarrow \infty$.

The above computation suggests that if the empirical measure $S^{(n)}(t)$ converges as n goes to infinity, its limit $\mu(t)$ should be a measure on $\mathbb{R}^d \times \mathbb{R}^e$ satisfying (in weak sense) the partial differential equation that is obtained formally integrating by parts Eq. (3.1), namely

$$\begin{aligned} \partial_t \mu(t) &= \frac{1}{2} \text{Tr} [\nabla^2 (\sigma \sigma^* \mu(t))] - \text{div}(b\mu(t)) \\ &\quad - \text{div}_y \left(\mu(t) \int g_2(\cdot, \cdot, \hat{x}, \hat{y}) \mu(t)(d\hat{x}, d\hat{y}) \right). \end{aligned} \quad (3.4)$$

The paper is devoted to the rigorous proof of this last statement, as formulated in Theorem 3.1 below.

To simplify further our notation we introduce the differential operator A and the measure-dependent differential operators B and L ; they act on twice differentiable test functions φ with all derivatives bounded as

$$A\varphi(x, y) = \frac{1}{2} \operatorname{Tr} [\sigma(x, y)\sigma^*(x, y)\nabla^2\varphi(xy)] , \quad (3.5)$$

$$B(\mu)\varphi(x, y) = b(x, y) \cdot \nabla\varphi(x, y) + \partial_y\varphi(x, y) \cdot \int_{\mathbb{R}^d \times \mathbb{R}^e} g_2(x, y, \hat{x}, \hat{y})\mu(\hat{x}, \hat{y}) , \quad (3.6)$$

$$L(\mu)\varphi(x, y) = B(\mu)\varphi(x, y) + A\varphi(x, y) . \quad (3.7)$$

Let us denote by $\operatorname{Pr}_1 = \operatorname{Pr}_1(\mathbb{R}^d \times \mathbb{R}^e)$ the set of Borel probability measures on $\mathbb{R}^d \times \mathbb{R}^e$ with bounded first moment, endowed with the 1-Wasserstein metric \mathcal{W}_1 .

We say that a function $\mu = \{\mu(t)\}_t \in C([0, T]; \operatorname{Pr}_1)$ is a weak solution of (3.4) if

$$\langle \mu(t), \varphi \rangle = \langle \mu(0), \varphi \rangle + \int_0^t \langle \mu(s), L(\mu(s))\varphi \rangle ds \quad (3.8)$$

for every $t \in [0, T]$ and for every test function $\varphi \in C_b^2(\mathbb{R}^d \times \mathbb{R}^e)$ with compact support. Our main result is:

Theorem 3.1. *Let (i)- (vii) hold and let μ_0 be the measure $\lambda_1 \times \lambda_2$ on $\mathbb{R}^d \times \mathbb{R}^e$. The partial differential equation (3.4) has a unique weak solution μ in the space $C([0, T]; \operatorname{Pr}_1)$ such that $\mu(0) = \mu_0$. The measure-valued function μ is the limit in probability of the empirical measures of the system (2.1) and it has an L^2 density.*

4. Uniqueness

Uniqueness can be proved using similar arguments as in [1]; we only sketch the strategy here, as details can be easily worked out.

The total variation distance between Borel probability measures ν and ξ on $\mathbb{R}^d \times \mathbb{R}^e$ is defined as

$$\operatorname{TV}(\nu, \xi) := \sup_{\substack{\varphi \in B_b \\ \|\varphi\|_\infty \leq 1}} |\langle \nu, \varphi \rangle - \langle \xi, \varphi \rangle| \quad (4.1)$$

where B_b denotes the space of bounded Borel functions on $\mathbb{R}^d \times \mathbb{R}^e$. For any $\mu_1, \mu_2 \in C([0, T]; \operatorname{Pr}_1)$ the map $t \mapsto \operatorname{TV}(\mu_1(t), \mu_2(t))$ is Borel and bounded on $[0, T]$.

The second order operator A defined in (3.5) generates a strongly continuous semigroup $P(t)$ of integral operators with C^2 densities $p(t)$, acting on bounded Borel functions on $\mathbb{R}^d \times \mathbb{R}^e$. As a consequence any weak solution μ of (3.4) is also a mild solution, that is, it satisfies, for every $t \in [0, T]$ and every test function φ , the equation

$$\langle \mu(t), \varphi \rangle = \langle \mu(0), P(t)\varphi \rangle + \int_0^t \langle \mu(s), B(\mu(s))P(t-s)\varphi \rangle ds . \quad (4.2)$$

6 *Giovanni Zanco*

Indeed if μ is a weak solution, differentiating the map $s \mapsto \langle \mu(s), P(t-s)\varphi \rangle$, using (3.8) and standard properties of the infinitesimal generator and then integrating on $[0, t]$ one finds exactly (4.2).

Thanks to the assumptions on σ , for every bounded Borel ψ

$$\|\nabla P(t)\psi\| \lesssim \frac{1}{\sqrt{t}} \|\psi\| ; \quad (4.3)$$

moreover $C_b^2(\mathbb{R}^d \times \mathbb{R}^e)$ is closed under the action of $P(t)$.

Let now $\mu_1, \mu_2 \in C([0, T]; \text{Pr}_1)$ be two weak (hence mild) solutions to Eq. (3.4). Thanks to the boundedness of b and g_2 and to (4.3) (recall (3.6)) we can bound the total variation distance between any two solutions of (3.4) as

$$\begin{aligned} \text{TV}(\mu_1(t), \mu_2(t)) &\leq \sup_{\substack{\varphi \in B_b \\ \|\varphi\|_\infty \leq 1}} \int_0^t |\langle \mu_1(s), [B(\mu_1(s)) - B(\mu_2(s))] P(t-s)\varphi \rangle| ds \\ &\quad + \sup_{\substack{\varphi \in B_b \\ \|\varphi\|_\infty \leq 1}} \int_0^1 |\langle \mu_1(s) - \mu_2(s), B(\mu_2(s)) P(t-s)\varphi \rangle| ds \\ &\lesssim \int_0^t \frac{1}{\sqrt{t-s}} \text{TV}(\mu_1(s), \mu_2(s)) ds ; \end{aligned}$$

uniqueness then follows immediately from Gronwall's lemma.

5. Existence

To prove existence we will follow a common strategy that uses tightness of the laws of the empirical measures and continuity of the partial differential equation above with respect to the weak convergence of probability measures (see for example [3]). However the fact that we allow for discontinuities requires some care; we will adapt some of the techniques already introduced in [1] for the case without dynamics in the variable X .

As the empirical measures are random processes, we can consider their law on the space $C([0, T]; \text{Pr}_1)$.

Proposition 5.1. *Let $\mathbb{Q}^{(n)}$ be the law of $S^{(n)}$ on $C([0, T]; \text{Pr}_1)$. The family $\{\mathbb{Q}^{(n)}\}_{n \in \mathbb{N}}$ is tight.*

Proof. The proof is standard. For every $M > 0$, $R > 0$, $(x_0, y_0) \in \mathbb{R}^d \times \mathbb{R}^e$, $\alpha \in (0, 1)$ and $q \geq 1$ such that $\alpha q > 1$, the set

$$\mathcal{K}_{M,R}^{(x_0, y_0)} = \left\{ \mu \in C([0, T]; \text{Pr}_1) : \sup_{t \in [0, T]} \int_{\mathbb{R}^d \times \mathbb{R}^e} |(x_0, y_0) - (x, y)| \mu(t) (dx, dy) \leq M, \int_0^T \int_0^T \frac{\mathcal{W}_1(\mu(t), \mu(s))^p}{|t-s|^{1+\alpha p}} dt ds \leq R \right\}$$

is relatively compact in $C([0, T]; \text{Pr}_1)$, by the Sobolev embedding and the characterization of the 1-Wasserstein distance using Lipschitz test functions. Thanks to the boundedness of all coefficients we have

$$\begin{aligned} & \mathbb{P} \left(\sup_{t \in [0, T]} \int_{\mathbb{R}^d \times \mathbb{R}^e} |(x_0, y_0) - (x, y)| S^{(n)}(t) (dx, dy) > M \right) \\ & \leq \frac{1}{M} \mathbb{E} \left[\left| \sup_{t \in [0, T]} \int_0^t |(x_0, y_0) - (x, y)| S^{(n)}(t) (dx, dy) \right| \right] \\ & \lesssim \frac{1}{M} \left(1 + \mathbb{E} \left[\sup_{t \in [0, T]} \left| \int_0^t \sigma(X^{i,n}(s), Y^{i,n}(s)) dW^i(s) \right| \right] \right) \\ & \lesssim \frac{1}{M}; \end{aligned}$$

moreover, for the same reasons

$$\begin{aligned} & \mathbb{P} \left(\int_0^T \int_0^T \frac{\mathcal{W}_1(S^{(n)}(t), S^{(n)}(s))^p}{|t-s|^{1+\alpha p}} dt ds > R \right) \\ & \leq \frac{1}{R} \int_0^T \int_0^T \frac{\mathbb{E} \left[\mathcal{W}_1(S^{(n)}(t), S^{(n)}(s))^p \right]}{|t-s|^{1+\alpha p}} dt ds \\ & \leq \frac{1}{nR} \sum_{i=1}^n \int_0^T \int_0^T \frac{\mathbb{E} \left[|(X^{i,n}(t), Y^{i,n}(t)) - (X^{i,n}(s), Y^{i,n}(s))|^p \right]}{|t-s|^{1+\alpha p}} dt ds \\ & \lesssim \frac{1}{R} \int_0^T \int_0^T \frac{|t-s|^{\frac{p}{2}}}{|t-s|^{\alpha p+1}} dt ds \\ & \lesssim \frac{1}{R} \end{aligned}$$

if $\alpha p < \frac{p}{2} - 1$. Therefore for every ε every $\mathbb{Q}^{(n)}$ gives probability less than ε to the complement of $\mathcal{K}_{M,R}^{(x_0, y_0)}$, for M and R big enough. \square

Let us call $\tilde{\mathcal{S}}$ the space in which we seek for the (unique) solution to the partial differential equation (3.4): we set

$$\tilde{\mathcal{S}} = \left\{ \mu \in C([0, T]; \text{Pr}_1) : \mu(t) \ll \mathcal{L}_{\mathbb{R}^d \times \mathbb{R}^e} \text{ with } \frac{d(\mu(t))}{d\mathcal{L}_{\mathbb{R}^d \times \mathbb{R}^e}} \in L^2(\mathbb{R}^d) \text{ for a.e. } t \in [0, T] \right\}. \quad (5.1)$$

Now for each measure $\mu_0 \in \text{Pr}(\mathbb{R}^d \times \mathbb{R}^e)$ and each test function φ we define the functional $\Xi_{\mu_0}^\varphi : C([0, T]; \text{Pr}_1) \rightarrow \mathbb{R}$ as

$$\Xi_{\mu_0}^\varphi(\mu) = \sup_{t \in [0, T]} \left| \langle \mu(t) - \mu_0, \varphi \rangle - \int_0^t L(\mu(s)) \varphi ds \right| \wedge 1.$$

Approximating the discontinuous part of L with smooth functions, it is easy to see (again thanks to boundedness) that $\Xi_{\mu_0}^\varphi$ is measurable.

8 *Giovanni Zanco*

Proposition 5.2. *For every choice of a test function φ and a probability measure μ_0 on $\mathbb{R}^d \times \mathbb{R}^e$ the functional $\Xi_{\mu_0}^\varphi$ is continuous on \tilde{S} .*

Proof. Let $\mu \in \tilde{S}$ and take a sequence $\{\mu^{(n)}\}_{n \in \mathbb{N}} \subset C([0, T]; \text{Pr}_1)$ such that $\mu^{(n)} \rightarrow \mu$ in $C([0, T]; \text{Pr}_1)$. Recalling the definition of L (see 3.7), we see that the convergence of the terms $\langle \mu^{(n)}(t) - \mu_0, \varphi \rangle$ and $\int_0^t \langle \mu^{(n)}(s), A\varphi \rangle ds$ is obvious; we bring therefore our attention on the part of L containing discontinuous coefficients.

Recall that F_1 and F_2 are the sets of discontinuity of the functions b_1 and b_2 , respectively, and let $F = F_1 \cup F_2$. Since $\mu \in \tilde{S}$ we have

$$\int_F \mu(s)(dx, dy) = 0$$

for almost every $s \in [0, T]$, so that the set F^c has probability 1. It follows that for almost every $s \in [0, T]$

$$\int_{\mathbb{R}^d \times \mathbb{R}^e} [b_1(x, y)\partial_x \varphi(x, y) + b_2(x, y)\partial_y \varphi(x, y)] \mu^{(n)}(s)(dx, dy)$$

converges to

$$\int_{\mathbb{R}^d \times \mathbb{R}^e} [b_1(x, y)\partial_x \varphi(x, y) + b_2(x, y)\partial_y \varphi(x, y)] \mu(s)(dx, dy)$$

and the convergence of the integral with respect to s follows then from the Dominated Convergence Theorem.

It remains to discuss the convergence of the term in which g_2 appears. For almost every s we have

$$\begin{aligned} & \left\langle \mu^{(n)}(s), \left\langle \mu^{(n)}(s), g_2(x, y, \cdot, \cdot) \right\rangle \partial_y \varphi \right\rangle - \left\langle \mu(s), \left\langle \mu(s), g_2(x, y, \cdot, \cdot) \right\rangle \partial_y \varphi \right\rangle \\ &= \left\langle \left(\mu^{(n)}(s) - \mu(s) \right), \left\langle \mu(s), g_2(x, y, \cdot, \cdot) \right\rangle \partial_y \varphi \right\rangle \\ & \quad + \left\langle \mu^{(n)}(s), \left\langle \left(\mu^{(n)}(s) - \mu(s) \right)_2(x, y, \cdot, \cdot) \right\rangle \partial_y \varphi \right\rangle \end{aligned}$$

and the first term on the right hand side can be handled exactly as before, for it is continuous on a set of measure 1 (the complement of the set of discontinuities of $\eta_1 \eta_2$). For the second term, we fix $\varepsilon > 0$ and we find, by tightness, a compact set K_ε such that $\mu^{(n)}(s)(K_\varepsilon^c) < \varepsilon$ for every n and $\mu(s)(K_\varepsilon^c) < \varepsilon$ as well. Therefore, recalling that $g_2(x, y, \hat{x}, \hat{y}) = \theta(x, \hat{x})\eta_1(y)\eta_2(\hat{y})$,

$$\begin{aligned} & \left\langle \mu(s)^{(n)}, \left\langle \left(\mu(s)^{(n)} - \mu(s) \right), g_2(x, y, \cdot, \cdot) \right\rangle \partial_y \varphi \right\rangle \\ & \leq 2\varepsilon \|\partial_y \varphi\| \|g_2\| + \|\partial_y \varphi\| \|\eta_1\| \int_{K_\varepsilon} \left| \left\langle \left(\mu^{(n)}(s) - \mu(s) \right), \theta(x, \hat{x})\eta_2(\hat{y}) \right\rangle \right| \mu^{(n)}(s)(d\hat{x}, d\hat{y}). \end{aligned}$$

The right hand side of this last inequality converges to 0 uniformly over K_ε . Indeed η_2 is discontinuous only on a set H_2 with Lebesgue measure zero; since the marginal of $\mu(s)$

with respect to y has an L^2 density, the same argument as above, applied to $(\mathbb{R}^d \times H_2)^{\mathbb{G}}$, yields the pointwise convergence

$$\int_{\mathbb{R}^d \times \mathbb{R}^e} \theta(x, \hat{x}) \eta_2(\hat{y}) \mu^{(n)}(s)(d\hat{x}, d\hat{y}) \rightarrow \int_{\mathbb{R}^d \times \mathbb{R}^e} \theta(x, \hat{x}) \eta_2(\hat{y}) \mu(s)(d\hat{x}, d\hat{y}).$$

The fact that the convergence is actually uniform is a consequence of the uniform continuity of θ , that ensures that the sequence of functions

$$\left\{ x \mapsto \int_{\mathbb{R}^d \times \mathbb{R}^e} \theta(x, \hat{x}) \eta_2(\hat{y}) \mu^{(n)}(s)(d\hat{x}, d\hat{y}) \right\}_{n \in \mathbb{N}}$$

is not only equi-bounded but also equi-continuous. \square

Remark 5.1. The last part of this proof is where uniform continuity in y is needed if one considers the alternative assumption (c) in Remark 2.1.

Proposition 5.3. *Take any weakly convergent subsequence of $\mathbb{Q}^{(n)}$ and let \mathbb{Q} be its limit point. Then $\mathbb{Q}(\tilde{S}) = 1$.*

Proof. Let $\gamma_1: \mathbb{R}^d \rightarrow \mathbb{R}$ and $\gamma_2: \mathbb{R}^e \rightarrow \mathbb{R}$ be smooth probability densities with compact support satisfying

$$\begin{aligned} |\partial_x \gamma_1(x) \cdot x| &\lesssim \gamma_1(x) \quad \forall x \in \mathbb{R}^d, \\ |\partial_y \gamma_2(y) \cdot y| &\lesssim \gamma_2(y) \quad \forall y \in \mathbb{R}^e, \end{aligned}$$

and set $\gamma(x, y) = \gamma_1(x) \gamma_2(y)$. Then

$$|\nabla \gamma(x, y) \cdot (x, y)| \lesssim \gamma(x, y) \quad \forall (x, y) \in \mathbb{R}^d \times \mathbb{R}^e.$$

Choose a sequence $\{\alpha^{(n)}\}_{n \in \mathbb{N}}$ of real numbers that converges to 0 and such that $(\alpha^{(n)})^{-3} \leq n$ for every $n \in \mathbb{N}$, and define mollifiers $\gamma^{(n)}(x, y)$ as

$$\gamma^{(n)}(x, y) = \frac{1}{\alpha^{(n)}} \gamma\left(\frac{x}{\alpha^{(n)}}, \frac{y}{\alpha^{(n)}}\right);$$

finally define the empirical densities

$$u^{(n)}(t)(x, y) = \left(S^{(n)}(t) * \gamma^{(n)} \right)(x, y)$$

and denote by $\mathbb{Q}^{u^{(n)}}$ their laws on $L^2(0, T; L^2(\mathbb{R}^d \times \mathbb{R}^e))$. Lemmata 5.1 and 5.2 in [1] can be applied to deduce that the family $\left\{ \mathbb{Q}^{u^{(n)}} \right\}_{n \in \mathbb{N}}$ belongs both to $L^2(0, T; W^{1,2}(\mathbb{R}^d \times \mathbb{R}^e))$ and to $W^{\beta,2}(0, T; H^{-2}(\mathbb{R}^d \times \mathbb{R}^e))$ for some $\beta > 0$, and thus is tight in $L^2(0, T; L^2(\mathbb{R}^d \times \mathbb{R}^e))$ thanks to a result by Simon [6].

Because of this fact and of Proposition 5.1 the laws

$$\xi^{(n)} = \mathcal{L}\left(S^{(n)}, u^{(n)}\right)$$

10 *Giovanni Zanco*

are tight in $C(0, T; \text{Pr}_1) \times L^2(0, T; L^2(\mathbb{R}^d \times \mathbb{R}^e))$; if \mathbb{Q} is as in the statement of the proposition, there must then exist a convergent subsequence $\xi^{(n_k)}$ whose limit ξ has marginal \mathbb{Q} on $C(0, T; \text{Pr}_1)$. By Skorohod's theorem we can find, on some probability space $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$, a sequence of random variables $(\hat{S}^{(n_k)}, \hat{u}^{(n_k)})$ with law $\xi^{(n_k)}$ that, with $\hat{\mathbb{P}}$ probability 1, converge strongly to some random variable (\hat{S}, \hat{u}) such that \hat{S} has law \mathbb{Q} . We want now to show that \hat{S} has \hat{u} as its density. To this end note that with $\hat{\mathbb{P}}$ probability 1 for almost every (t, x, y)

$$\hat{u}^{(n_k)}(t)(x, y) = \left(\gamma^{(n_k)} * \hat{S}^{(n_k)}(t) \right) (x, y).$$

Therefore for any $\varphi \in C(\mathbb{R}^d \times \mathbb{R}^e)$ we have, for $\hat{\mathbb{P}} \otimes \mathcal{L}_{[0, t]}$ almost every (ω, t) ,

$$\begin{aligned} \int_{\mathbb{R}^d \times \mathbb{R}^e} \varphi(x, y) \hat{u}(t)(dx, dy) &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d \times \mathbb{R}^e} \varphi(x, y) \hat{u}^{(n_k)}(t)(dx, dy) \\ &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d \times \mathbb{R}^e} \varphi(x, y) \int_{\mathbb{R}^d \times \mathbb{R}^e} \gamma^{(n_k)}(x - \hat{x}, y - \hat{y}) \hat{S}^{(n_k)}(t)(d\hat{x}, d\hat{y}) dx dy \\ &= \lim_{k \rightarrow \infty} \int_{\mathbb{R}^d \times \mathbb{R}^e} \left(\int_{\mathbb{R}^d \times \mathbb{R}^e} \gamma^{(n_k)}(x - \hat{x}, y - \hat{y}) \varphi(x, y) dx dy \right) \hat{S}^{(n_k)}(t)(d\hat{x}, d\hat{y}) \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^e} \varphi(\hat{x}, \hat{y}) \hat{S}(t)(d\hat{x}, d\hat{y}). \end{aligned}$$

Taking a countable dense set in $C(\mathbb{R}^d \times \mathbb{R}^e)$, we have that the above equality holds for any φ in such set. Therefore with $\hat{\mathbb{P}}$ probability 1 \hat{S} has an L^2 density, hence

$$\mathbb{Q}(\tilde{S}) = \hat{\mathbb{P}}(\hat{S} \in \tilde{S}) = 1. \quad \square$$

We can now give a proof of our main result, Theorem 3.1.

Consider a weakly convergent subsequence $\{\mathbb{Q}^{(n_k)}\}_{k \in \mathbb{N}}$ of $\{\mathbb{Q}^{(n)}\}_{n \in \mathbb{N}}$ (which exists thanks to Proposition 5.1) and let \mathbb{Q} be its limit; let moreover H be a countable dense subset of $C_c^2(\mathbb{R}^d \times \mathbb{R}^e)$. We have that, for $\varphi \in H$ and $\mu_0 \in \text{Pr}_1$ (writing simply C for $C([0, T]; \text{Pr}_1)$),

$$0 \leq \lim_{n \rightarrow \infty} \left(\int_C \Xi_{\mu_0}^{\varphi} d\mathbb{Q}^{(n)} \right)^2 \leq \lim_{n \rightarrow \infty} \mathbb{E} \left[\sup_{t \in [0, t]} \left| M_{\varphi}^{(n)}(t) \right|^2 \right] \lesssim \lim_{n \rightarrow \infty} \frac{1}{N} = 0,$$

where $M^{N, \varphi}(t)$ was defined in (3.3); the above inequality holds thanks to boundedness of φ and σ and to independence of the Brownian motions W^i .

Therefore, by the continuity of Ξ given in Proposition 5.2, we deduce that

$$\int_C \Xi_{\mu_0}^{\varphi} d\mathbb{Q} = 0;$$

hence, for $\varphi \in H$ and μ_0 as above, the set of measure-valued processes μ for which $\Xi_{\mu_0}^{\varphi}(\mu) = 0$ has \mathbb{Q} probability 1. Since H is countable and dense, this implies that, for given $\mu_0 \in \text{Pr}_1(\mathbb{R}^d \times \mathbb{R}^e)$, the set

$$\{\mu \in C([0, T]; \text{Pr}_1) : \Xi_{\mu_0}^{\varphi}(\mu) = 0 \forall \varphi \text{ test function}\}$$

has \mathbb{Q} probability 1. By uniqueness of solutions to (3.4) we also have that \mathbb{Q} must be a Dirac mass concentrated on the unique solution, thus a constant. This implies that the whole sequence $\{\mathbb{Q}^{(n)}\}_{n \in \mathbb{N}}$ converges in probability to \mathbb{Q} . Finally Proposition 5.3 ensures that the unique solution has a density in L^2 . Theorem 3.1 is proved.

6. Extensions

The existence of a L^2 density for the i.i.d. initial conditions X_0^i seems not to be fundamental, and is used here to obtain estimates that are uniform in (x, y) . One can show, repeating the arguments above with some refinements, that the solution of the Fokker-Planck equation (3.4) exists and is unique also when the initial data for the dynamics of $X^{i,n}$ do not have a density but just bounded second moment. In this case one should introduce the empirical density only with respect to y :

$$u^{(n)}(t)(x, y) = \int_{\mathbb{R}^d \times \mathbb{R}^e} \gamma_2(y - \hat{y}) S^{(n)}(t)(d\hat{x}, d\hat{y}),$$

and note that the estimates provided by Lemmata 5.1 and 5.2 in [1] still hold, because they are a consequence of boundedness of the coefficients and of an L^2 bound on $u^{(n)}(0)$ only. In this case, after passing to almost sure convergence by using Skorohod's theorem as in the proof of Proposition 5.3, one cannot expect that \hat{S} has a joint density. However one can repeat the argument for $\pi_y \hat{S}$, the marginal of \hat{S} with respect to y , and deduce that any limit point of any subsequence of the laws $\mathbb{Q}^{(n)}$ must give measure 1 to the set of measure-valued functions μ whose marginal with respect to y has an L^2 density. The rest of the proof remains unchanged.

If one uses the alternative assumption (b) in Remark 2.1, then the smoothing in the variable x given by the mollifiers $\gamma^{(n)}$ is useless. The details have to be carefully adjusted since there is no global boundedness estimate; however it should be sufficient to exploit the sublinear growth together with the L^2 estimate on the initial conditions. A rigorous study of this situation is left for future work. It is anyway important to notice that in this case one will not see a second order operator in the x variable in the limit Fokker-Planck equation, which will then take the form

$$\begin{aligned} \partial_t \mu_t = & \frac{1}{2} \operatorname{Tr} [\partial_y^2 (\sigma_2 \sigma_2^* \mu(t))] - \operatorname{div}(b\mu(t)) \\ & - \operatorname{div}_y \left(\mu(t) \int g_2(\cdot, \cdot, \hat{x}, \hat{y}) \mu(t)(d\hat{x}, d\hat{y}) \right). \end{aligned} \quad (6.1)$$

Acknowledgments

The author is deeply grateful to Franco Flandoli, among many other reasons, for having introduced him to the theory of interacting systems and having taught him many techniques that were very useful also in the realization of this article.

Bibliography

1. F. Flandoli, E. Priola, and G. Zanco. A mean-field model with discontinuous coefficients for neurons with spatial interaction. *Discrete & Continuous Dynamical Systems - A*, 39:3037, 2019.
2. I. Gyöngy and N. Krylov. Existence of strong solutions for Itô's stochastic equations via approximations. *Prob. Theory Relat. Fields*, 105:143–158, 1996.
3. S. Méléard. *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models*, pages 42–95. Springer Berlin Heidelberg, Berlin, Heidelberg, 1996.
4. Y. S. Mishura and A. Y. Veretennikov. Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *arXiv preprint*, 2020.
5. K. Oelschläger. On the derivation of reaction-diffusion equations as limit dynamics of systems of moderately interacting stochastic processes. *Probability Theory and Related Fields*, 4(82):565–586, 1989.
6. J. Simon. Compact sets in the space $l(0, t; b)$. *Ann. Mat. Pura. Appl.*, 4(146):65–96, 1987.
7. A.-S. Sznitman. *Topics in propagation of chaos, Ecole d'Eté de Probabilités de Saint-Flour XIX – 1989*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1991.
8. A. J. Veretennikov. Strong solutions and explicit formulas for solutions of stochastic integral equations. *Math. USSR-Sb. (N.S.)*, 39(3):487–403, 1981.