1 Compact sets in the space of measure valued functions

On $\Pr\left(\mathbb{R}^d\right)$, the set of all probability measures on Borel sets of $\mathbb{R}^d$, we may introduce several metrics. There is a general kind of metric given by

$$\delta\left(\mu,\nu\right)=\sum_{i=1}^{\infty}2^{-i}\frac{\left|\langle\mu,\phi_i\rangle-\langle\nu,\phi_i\rangle\right|}{1+\left|\langle\mu,\phi_i\rangle-\langle\nu,\phi_i\rangle\right|}$$

(or $\delta\left(\mu,\nu\right)=\sum_{i=1}^{\infty}2^{-i}\left(\left|\langle\mu,\phi_i\rangle-\langle\nu,\phi_i\rangle\right|\wedge 1\right)$) where $\{\phi_i\}$ is a suitable sequences of bounded continuous functions. Another one, but on $\Pr_1\left(\mathbb{R}^d\right)$, the set of probability measures μ with finite first moment $\int_{\mathbb{R}^d}\left|x\right|\mu\left(dx\right)<\infty$, is the Wasserstein metric

$$\mathcal{W}_1\left(\mu,\nu\right)=\sup\left\{\left|\langle\mu,\phi\rangle-\langle\nu,\phi\rangle\right|:\left[\phi\right]_{\mathrm{Lip}}\leq 1\right\}$$

where $\left[\phi\right]_{\mathrm{Lip}}$ is the Lipschitz seminorm $\left[\phi\right]_{\mathrm{Lip}}=\sup_{s\neq t}\frac{\left|\phi(t)-\phi(s)\right|}{\left|t-s\right|}$ (in fact this is not the usual definition but a theorem, named Kantorovich-Rubinstein characterization). In both cases we have a complete separable metric space and the convergence is equivalent to the weak convergence of probability measures. As a consequence of the general version of Ascoli-Arzelà theorem in metric spaces, we have the following fact.

Proposition 1 *Let $\mathcal{E}=\Pr\left(\mathbb{R}^d\right)$ with the metrics $d=\delta$ above or $\mathcal{E}=\Pr_1\left(\mathbb{R}^d\right)$ with the Wasserstein metric $d=\mathcal{W}_1$. A family of measure-valued functions $F\subset C\left(\left[0,T\right];\mathcal{E}\right)$ is relatively compact in $C\left(\left[0,T\right];\mathcal{E}\right)$ if*

i) for every $t\in\left[0,T\right]$, for every $\epsilon>0$, there exists $r_{\epsilon,t}>0$ such that

$$\mu_t\left(B\left(0,r_{\epsilon,t}\right)\right)>1-\epsilon$$

for every $\mu_{\cdot}\in F$

ii) for every $\epsilon>0$, there exists $\delta>0$ such that $d\left(\mu_t,\mu_s\right)<\epsilon$ for every $\mu_{\cdot}\in F$ and $t,s\in\left[0,T\right]$ such that $\left|t-s\right|<\delta$.

To prove the second condition in examples, in the case of the Wasserstein metric, we may use the following fact.

Lemma 2 *If, for some $\alpha\in\left(0,1\right)$ and $p\geq 1$ with $\alpha p>1$, and a constant $C>0$ one has*

$$\int_0^T\int_0^T\frac{\mathcal{W}_1\left(\mu_t,\mu_s\right)^p}{\left|t-s\right|^{1+\alpha p}}dtds\leq C$$

for every $\mu_{\cdot}\in F$, then (ii) holds, with $\mathcal{E}=\Pr_1\left(\mathbb{R}^d\right)$.

Proof. Clearly, (ii) holds if there exists $\theta, C > 0$ such that $\sup_{s \neq t} \frac{\mathcal{W}_1(\mu_t, \mu_s)}{|t-s|^\theta} \leq C$, for every $\mu_. \in F$. Choose $\theta > 0$ such that $(\alpha - \theta) p > 1$. Then, for some constant $C > 0$,

$$|\langle \mu_t, \phi \rangle - \langle \mu_s, \phi \rangle| \leq C |t-s|^\theta [\langle \mu_t, \phi \rangle]_{W^{\alpha,p}}$$

and therefore

$$\mathcal{W}_1(\mu_t, \mu_s) \leq C |t-s|^\theta \sup \left\{ [\langle \mu_t, \phi \rangle]_{W^{\alpha,p}} : [\phi]_{\mathrm{Lip}} \leq 1 \right\}.$$

It follows that

$$\begin{aligned} \sup_{s \neq t} \frac{\mathcal{W}_1(\mu_t, \mu_s)}{|t-s|^\theta} &\leq C \sup_{[\phi]_{\mathrm{Lip}} \leq 1} \int_0^T \int_0^T \frac{|\langle \mu_t, \phi \rangle - \langle \mu_s, \phi \rangle|^p}{|t-s|^{1+\alpha p}} dt ds \\ &\leq C \int_0^T \int_0^T \frac{\left(\sup_{[\phi]_{\mathrm{Lip}} \leq 1} |\langle \mu_t, \phi \rangle - \langle \mu_s, \phi \rangle| \right)^p}{|t-s|^{1+\alpha p}} dt ds \\ &= C \int_0^T \int_0^T \frac{\mathcal{W}_1(\mu_t, \mu_s)^p}{|t-s|^{1+\alpha p}} dt ds. \end{aligned}$$

■

To prove the second condition in examples, for the metric δ, it is useful to have the following criterion.

Lemma 3 *Assume that, for every $\phi \in C_b(\mathbb{R}^d)$, one has the following property: for every $\epsilon > 0$, there exists $\delta > 0$ such that*

$$|\langle \mu_t, \phi \rangle - \langle \mu_s, \phi \rangle| < \epsilon$$

for every $\mu_. \in F$ and $t, s \in [0, T]$ such that $|t-s| < \delta$. Then condition (ii) above holds, with $\mathcal{E} = (\mathbb{R}^d)$. In particular, a sufficient condition is that there exist $\alpha \in (0,1)$ with the following property: for every $\phi \in C_b(\mathbb{R}^d)$ there is a constant $C_\phi > 0$ such that

$$[\langle \mu_., \phi \rangle]_{C^\alpha} \leq C_\phi$$

for every $\mu_. \in F$. Another sufficient condition is that there exist $\alpha \in (0,1)$ and $p \geq 1$ with $\alpha p > 1$ with the following property: for every $\phi \in C_b(\mathbb{R}^d)$ there is a constant $C_\phi > 0$ such that

$$[\langle \mu_., \phi \rangle]_{W^{\alpha,p}} \leq C_\phi$$

for every $\mu_. \in F$.

Proof. Given $\epsilon > 0$, let $k \in \mathbb{N}$ be such that $2^{-k} \leq \epsilon/2$. Notice that

$$\sum_{i=k+1}^{\infty} 2^{-i} \frac{|\langle \mu_t, \phi_i \rangle - \langle \mu_s, \phi_i \rangle|}{1 + |\langle \mu_t, \phi_i \rangle - \langle \mu_s, \phi_i \rangle|} \leq \sum_{i=k+1}^{\infty} 2^{-i} \leq 2^{-k} \leq \epsilon/2$$

hence

$$d\left(\mu_t,\mu_s\right)\leq\sum_{i=1}^{k}2^{-i}\frac{\left|\left\langle\mu_t,\phi_i\right\rangle-\left\langle\mu_s,\phi_i\right\rangle\right|}{1+\left|\left\langle\mu_t,\phi_i\right\rangle-\left\langle\mu_s,\phi_i\right\rangle\right|}+\epsilon/2.$$

Since k is finite, there exists $\delta>0$ such that

$$\left|\left\langle\mu_t,\phi_i\right\rangle-\left\langle\mu_s,\phi_i\right\rangle\right|<\epsilon/2$$

for every $\mu_.\in F$ and $t,s\in[0,T]$ such that $|t-s|<\delta$ and $i=1,...,k$. Then

$$d\left(\mu_t,\mu_s\right)\leq\frac{\epsilon}{2}\sum_{i=1}^{k}2^{-i}+\frac{\epsilon}{2}<\epsilon$$

for every $\mu_.\in F$ and $t,s\in[0,T]$ such that $|t-s|<\delta$. This is property (ii). ■

Remark 4 *The previous result is somewhat in the same spirit as Aubin-Lions lemma: the regularity in time necessary for tightness is sufficient when the space regularity is measured in a very weak sense. Compactness in time and in space are dealt with separately.*

Lemma 5 *Let $t\in[0,T]$ be given. If there exists a constant $C_t>0$ such that*

$$\int_{\mathbb{R}^d}|x|\,\mu_t\left(dx\right)\leq C_t$$

for every $\mu_.\in F$, then condition (i) is fulfilled, for that value of t.

Proof.

$$\mu_t\left(B\left(0,r\right)^c\right)\leq\frac{1}{r}\int_{\mathbb{R}^d}|x|\,\mu_t\left(dx\right)\leq\frac{C_t}{r}$$

hence, given $\epsilon>0$, we may choose $r>\frac{C_t}{\epsilon}$. ■

Remark 6 *We may replace $\int_{\mathbb{R}^d}|x|\,\mu_t\left(dx\right)$ in the hypothesis of the lemma by $\int_{\mathbb{R}^d}g\left(|x|\right)\mu_t\left(dx\right)$ with any increasing function g such that $\lim_{r\to+\infty}g\left(r\right)=+\infty$.*

Corollary 7 *Let $\alpha\in(0,1)$ and $p\geq1$ with $\alpha p>1$, be given. For any $M,R>0$ the set $\mathcal{K}_{M,R}\subset C\left([0,T];\mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$ defined as*

$$\mathcal{K}_{M,R}=\left\{\mu_.:\sup_{t\in[0,T]}\int_{\mathbb{R}^d}|x|\,\mu_t\left(dx\right)\leq M,\int_0^T\int_0^T\frac{\mathcal{W}_1\left(\mu_t,\mu_s\right)^p}{|t-s|^{1+\alpha p}}dtds\leq R\right\}$$

is relatively compact in $C\left([0,T];\mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$.

Corollary 8 *Let $\alpha \in (0,1)$ and $p \geq 1$ with $\alpha p > 1$, and g be an increasing function such that $\lim_{r\to+\infty} g(r) = +\infty$. For every $M > 0$ and every function $\phi \mapsto C_\phi$ from $C_b(\mathbb{R}^d)$ to $(0,\infty)$, the set $\mathcal{K}_{M,C.} \subset C([0,T]; \Pr(\mathbb{R}^d))$ defined as*

$$\mathcal{K}_{M,C.} = \left\{ \mu_. : \sup_{t\in[0,T]} \int_{\mathbb{R}^d} g(|x|)\, \mu_t(dx) \leq M, [\langle \mu_., \phi \rangle]_{W^{\alpha,p}} \leq C_\phi \text{ for every } \phi \in C_b\left(\mathbb{R}^d\right) \right\}$$

is relatively compact in $C([0,T]; \Pr(\mathbb{R}^d))$.

2 Mean field model. Compactness of the empirical measure

Consider the interacting particle system

$$dX_t^{i,N} = \frac{1}{N} \sum_{j=1}^{N} K\left(X_t^{i,N} - X_t^{j,N}\right) dt + \sigma dB_t^i$$

with $i = 1, ..., N$, $K : \mathbb{R}^d \to \mathbb{R}^d$ bounded Lipschitz, $\sigma > 0$. We assume to have a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ and independent Brownian motions B_t^i in $\mathbb{R}^d$.

Remark 9 *Each particle $X_t^{i,N}$ interacts with every other particle $X_t^{j,N}$ such that the distance $X_t^{i,N} - X_t^{j,N}$ is in the support of K. Even if K is compact support, in the limit when $N \to \infty$ we must expect that the number of particles $X_t^{j,N}$ interacting with $X_t^{i,N}$ goes to infinity since, as we shall see, the empirical measure will converge to a probability measure, so particles remain relatively concentrated. For this reason, this model is also called "long range". It is not strictly correct if we want to describe local interactions, like membrane interactions between cells.*

On the initial conditions $X_0^{i,N}$, we assume they are $\mathcal{F}_0$-measurable and:

a) $\sup_{i,N} E\left[\left|X_0^{i,N}\right|\right] < \infty$

b) there is $\mu_0 \in \Pr_1(\mathbb{R}^d)$ such that $\langle S_0^N, \phi \rangle \to \langle \mu_0, \phi \rangle$ in probability, for every $\phi \in C_b(\mathbb{R}^d)$.

This is true in particular when $X_0^{i,N} = X_0^i$, where $\{X_0^i\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. $\mathcal{F}_0$-measurable r.v.'s with law $\mu_0 \in \Pr_1(\mathbb{R}^d)$.

Denote by S_t^N the empirical measure and by Q^N the law of $S_.^N$ on $C([0,T]; \Pr_1(\mathbb{R}^d))$.

Theorem 10 *Under assumption (a), $\{Q^N\}_{N\in\mathbb{N}}$ is relatively compact on $C([0,T]; \Pr_1(\mathbb{R}^d))$.*

Proof. Step1. Given $\epsilon > 0$, we have to find a compact set $\mathcal{K}_\epsilon \subset \mathcal{E}$ such that

$$Q^N(\mathcal{K}_\epsilon) > 1 - \epsilon$$

for every $N \in \mathbb{N}$. We look for a set of the form $\mathcal{K}_{M,R}$ as described in Corollary 7. We have

$$\begin{aligned} Q^N\left(\mathcal{K}_{M,R}^c\right) &= P\left(S_\cdot^N \in \mathcal{K}_{M,R}^c\right) \\ &\leq P\left(\sup_{t\in[0,T]} \int_{\mathbb{R}^d} |x|\, S_t^N(dx) > M\right) + P\left(\int_0^T \int_0^T \frac{\mathcal{W}_1\left(S_t^N, S_s^N\right)^p}{|t-s|^{1+\alpha p}} dtds > R\right). \end{aligned}$$

We separately prove that both probabilities are smaller than $\epsilon/2$, for every $N \in \mathbb{N}$. In both cases we apply Chebishev inequality. We have

$$\int_{\mathbb{R}^d} |x|\, S_t^N(dx) = \frac{1}{N}\sum_{i=1}^N \left|X_t^{i,N}\right|$$

hence

$$\begin{aligned} P\left(\sup_{t\in[0,T]} \int_{\mathbb{R}^d} |x|\, S_t^N(dx) > M\right) &\leq \frac{1}{M} E\left[\sup_{t\in[0,T]} \int_{\mathbb{R}^d} |x|\, S_t^N(dx)\right] \\ &= \frac{1}{M} E\left[\sup_{t\in[0,T]} \frac{1}{N}\sum_{i=1}^N \left|X_t^{i,N}\right|\right] \\ &= \frac{1}{M}\frac{1}{N}\sum_{i=1}^N E\left[\sup_{t\in[0,T]} \left|X_t^{i,N}\right|\right] \leq \frac{C}{M} \end{aligned}$$

where the property $E\left[\sup_{t\in[0,T]} \left|X_t^{i,N}\right|\right] \leq C$ will be checked below in Step 2. Similarly, we have

$$\begin{aligned} P\left(\int_0^T \int_0^T \frac{\mathcal{W}_1\left(S_t^N, S_s^N\right)^p}{|t-s|^{1+\alpha p}} dtds > R\right) &\leq \frac{1}{R} E \int_0^T \int_0^T \frac{\mathcal{W}_1\left(S_t^N, S_s^N\right)^p}{|t-s|^{1+\alpha p}} dtds \\ &= \frac{1}{R}\int_0^T \int_0^T \frac{E\left[\mathcal{W}_1\left(S_t^N, S_s^N\right)^p\right]}{|t-s|^{1+\alpha p}} dtds. \end{aligned}$$

Thus we have to prove that

$$E\left[\mathcal{W}_1\left(S_t^N, S_s^N\right)^p\right] \leq C\,|t-s|^{1+\beta}$$

for some $p \geq 1$ and $\beta > 0$. We shall prove this below in Step 3. If so, choose $\alpha > 0$ so small that $\alpha p < \beta$ and get

$$P\left(\int_0^T \int_0^T \frac{\mathcal{W}_1\left(S_t^N, S_s^N\right)^p}{|t-s|^{1+\alpha p}} dtds > R\right) \leq \frac{C}{R}.$$

Now, given $\epsilon > 0$ above, we may find $M, R > 0$ such that $Q^N\left(\mathcal{K}_{M,R}^c\right) < \epsilon$.

Step2. We simply have

$$\begin{aligned}\left|X_t^{i,N}\right| &\leq \left|X_0^i\right| + \frac{1}{N}\sum_{j=1}^N \int_0^t \left|K\left(X_s^{i,N} - X_s^{j,N}\right)\right| ds + \sigma\left|B_t^i\right| \\ &\leq \left|X_0^i\right| + \|K\|_\infty T + \sigma\left|B_t^i\right|\end{aligned}$$

hence, recalling that $E\left[\left|X_0^i\right|\right] = \int |x|\,\mu_0(dx) < \infty$,

$$E\left[\sup_{t\in[0,T]}\left|X_t^{i,N}\right|\right] \leq C + \sigma E\left[\sup_{t\in[0,T]}\left|B_t^1\right|\right] \leq C.$$

Step3. In order to estimate $\mathcal{W}_1\left(S_t^N, S_s^N\right)$ we first estimate $\left|\left\langle S_t^N, \phi\right\rangle - \left\langle S_s^N, \phi\right\rangle\right|$ with $\mathrm{Lip}(\phi) \leq 1$. We have

$$\begin{aligned}\left|\left\langle S_t^N, \phi\right\rangle - \left\langle S_s^N, \phi\right\rangle\right| &= \left|\frac{1}{N}\sum_{i=1}^N\left(\phi\left(X_t^{i,N}\right) - \phi\left(X_s^{i,N}\right)\right)\right| \\ &\leq \frac{1}{N}\sum_{i=1}^N\left|\phi\left(X_t^{i,N}\right) - \phi\left(X_s^{i,N}\right)\right| \\ &\leq \frac{1}{N}\sum_{i=1}^N\left|X_t^{i,N} - X_s^{i,N}\right|.\end{aligned}$$

Hence

$$\mathcal{W}_1\left(S_t^N, S_s^N\right) \leq \frac{1}{N}\sum_{i=1}^N\left|X_t^{i,N} - X_s^{i,N}\right|.$$

Therefore (treating $\frac{1}{N}\sum_{i=1}^N$ as an integration w.r.t. a probability measure on the natural numbers $1, ..., N$, hence using Hölder inequality)

$$\mathcal{W}_1\left(S_t^N, S_s^N\right)^p \leq \frac{1}{N}\sum_{i=1}^N\left|X_t^{i,N} - X_s^{i,N}\right|^p$$

$$E\left[\mathcal{W}_1\left(S_t^N, S_s^N\right)^p\right] \leq \frac{1}{N}\sum_{i=1}^N E\left[\left|X_t^{i,N} - X_s^{i,N}\right|^p\right].$$

Thus we have to estimate $E\left[\left|X_t^{i,N} - X_s^{i,N}\right|^p\right]$. The system is reversible, hence we could reduce to $E\left[\left|X_t^{1,N} - X_s^{1,N}\right|^p\right]$, but it is unessential, just conceptually interesting (a paradigm is that the quantitative analysis of particle $i = 1$ is sufficient).

From the SDE we have

$$X_t^{i,N} - X_s^{i,N} = \frac{1}{N}\sum_{j=1}^{N}\int_s^t K\left(X_r^{i,N} - X_r^{j,N}\right)dr + \sigma\left(B_t^i - B_s^i\right)$$

hence

$$\begin{aligned}\left|X_t^{i,N} - X_s^{i,N}\right| &\leq \frac{1}{N}\sum_{j=1}^{N}\int_s^t \left|K\left(X_r^{i,N} - X_r^{j,N}\right)\right|dr + \sigma\left|B_t^i - B_s^i\right| \\ &\leq \|K\|_\infty |t-s| + \sigma\left|B_t^i - B_s^i\right|\end{aligned}$$

which implies

$$\begin{aligned}E\left[\left|X_t^{i,N} - X_s^{i,N}\right|^p\right] &\leq C|t-s|^p + CE\left[\left|B_t^i - B_s^i\right|^p\right] \\ &\leq C|t-s|^p + C|t-s|^{p/2} \\ &\leq C|t-s|^{p/2}\end{aligned}$$

renaming each time the constants. Choosing $p > 2$ and setting $\beta = \frac{p}{2} - 1 > 0$, the proof is now complete because we have proved

$$E\left[\mathcal{W}_1\left(S_t^N, S_s^N\right)^p\right] \leq C|t-s|^{1+\beta}.$$

■

It may be noticed that, from the proof of this theorem and the results of the previous section, we may extract a general criterion of compactness of a family of empirical measures S_t^N, independenly of the precise equations satisfied by the particles:

Proposition 11 *If, for some $p, \beta, C > 0$*

$$E\left[\sup_{t\in[0,T]}\int_{\mathbb{R}^d}|x|\,S_t^N(dx)\right] \leq C$$

$$Q\left[\mathcal{W}_1\left(S_t^N, S_s^N\right)^p\right] \leq C|t-s|^{1+\beta}$$

for every $N \in \mathbb{N}$, then the famyly of laws Q^N of $S_\cdot^N$ is relatively compact in $\Pr\left(C\left([0,T];\Pr_1\left(\mathbb{R}^d\right)\right)\right)$.

The fact that Q^N are laws of $S_\cdot^N$ is not essential. A general criterion is:

Proposition 12 *Set $\mathcal{C} = C\left([0,T];\Pr_1\left(\mathbb{R}^d\right)\right)$. Assume that $\mathcal{G} \subset \Pr(\mathcal{C})$ is a family of probability measures on $\mathcal{C}$. If, for some $p, \beta, C > 0$*

$$\int_{\mathcal{C}}\left(\sup_{t\in[0,T]}\int_{\mathbb{R}^d}|x|\,\mu_t(dx)\right)dQ(\mu_\cdot) \leq C$$

$$\int_{\mathcal{C}} \mathcal{W}_1 (\mu_t, \mu_s)^p \, dQ (\mu_.) \leq C |t-s|^{1+\beta}$$

for every $Q \in \mathcal{G}$, then $\mathcal{G}$ is relatively compact.

Here the notation $\mu_.$ stands for the generic element of $C\left([0,T]; \mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$.

3 Passage to the limit

Let us now examine the limit points of $\left\{Q^N\right\}_{N\in\mathbb{N}}$, the family of laws of the empirical process. The aim is to prove that, in the limit, we get solutions of the nonlinear Fokker-Planck equation

$$\frac{\partial \mu_t}{\partial t} = \frac{\sigma^2}{2} \Delta \mu_t - \mathrm{div} \left(\mu_t K * \mu_t\right) \tag{1}$$

with initial condition μ_0. However, these solutions are, at least a priori, only probability measures, depending on time (and a priori also on randomness). Thus we use the concept of measure-valued solutions, completely similar to the one given in a previous section in the linear case.

Let us first give the intuition about the result. Preliminarily, notice that

$$\frac{1}{N} \sum_{j=1}^{N} K\left(x - X_t^{j,N}\right) = \int K(x-y) \, S_t^N (dy) =: \left(K * S_t^N\right)(x)$$

hence the SDE of the interacting particles can be rewritten as

$$dX_t^{i,N} = \left(K * S_t^N\right)\left(X_t^{i,N}\right) dt + \sigma dB_t^i.$$

Let $\phi_t(x)$ be a test function of calls $C_b^{1.2}\left([0,T] \times \mathbb{R}^d\right)$. By Itô formula we have

$$\begin{aligned} d\phi_t\left(X_t^{i,N}\right) &= \frac{\partial \phi_t}{\partial t}\left(X_t^{i,N}\right) dt + \nabla \phi_t\left(X_t^{i,N}\right) \cdot \left(K * S_t^N\right)\left(X_t^{i,N}\right) dt \\ &\quad + \nabla \phi_t\left(X_t^{i,N}\right) \cdot \sigma dB_t^i + \frac{\sigma^2}{2} \Delta \phi_t\left(X_t^{i,N}\right) dt. \end{aligned}$$

Therefore, being $\left\langle S_t^N, \phi_t \right\rangle = \frac{1}{N}\sum_{i=1}^N \phi_t\left(X_t^{i,N}\right)$,

$$\begin{aligned} d\left\langle S_t^N, \phi_t \right\rangle &= \frac{1}{N}\sum_{i=1}^N \frac{\partial \phi_t}{\partial t}\left(X_t^{i,N}\right) dt \\ &+ \frac{1}{N}\sum_{i=1}^N \nabla\phi_t\left(X_t^{i,N}\right)\cdot\left(K * S_t^N\right)\left(X_t^{i,N}\right) dt \\ &+ \frac{1}{N}\sum_{i=1}^N \nabla\phi_t\left(X_t^{i,N}\right)\cdot \sigma dB_t^i \\ &+ \frac{1}{N}\sum_{i=1}^N \frac{\sigma^2}{2}\Delta\phi_t\left(X_t^{i,N}\right) dt \end{aligned}$$

$$= \left\langle S_t^N, \frac{\partial \phi_t}{\partial t}\right\rangle dt + \left\langle S_t^N, \nabla\phi_t\cdot\left(K * S_t^N\right)\right\rangle dt + dM_t^{\phi,N} + \frac{\sigma^2}{2}\left\langle S_t^N, \Delta\phi_t\right\rangle dt$$

where

$$M_t^{\phi,N} = \frac{1}{N}\sum_{i=1}^N \int_0^t \nabla\phi_s\left(X_s^{i,N}\right)\cdot\sigma dB_s^i.$$

Therefore we have proved:

Lemma 13 *The empirical measure S_t^N satisfies the identity*

$$d\left\langle S_t^N, \phi_t \right\rangle = \left\langle S_t^N, \frac{\partial \phi_t}{\partial t}\right\rangle dt + \left\langle S_t^N, \nabla\phi_t\cdot\left(K * S_t^N\right)\right\rangle dt + dM_t^{\phi,N} + \frac{\sigma^2}{2}\left\langle S_t^N, \Delta\phi_t\right\rangle dt.$$

Remark 14 *We cannot write $M_t^{\phi,N}$ in terms of S_t^N, at least directly. However, this is true for the quadratic variations:*

$$\left[M_\cdot^{\phi,N}, M_\cdot^{\phi,N}\right]_t = \frac{\sigma^2}{N^2}\sum_{i=1}^N \int_0^t \left|\nabla\phi_s\left(X_s^{i,N}\right)\right|^2 ds = \frac{\sigma^2}{N}\int_0^t \left\langle S_s^N, |\nabla\phi_s|^2\right\rangle ds.$$

By a representation theorem for stochastic integrals, there exists an auxiliary Brownian motion β_t such that

$$M_t^{\phi,N} = \frac{\sigma}{\sqrt{N}}\int_0^t \sqrt{\left\langle S_s^N, |\nabla\phi_s|^2\right\rangle} d\beta_s.$$

In this somewhat artificial way we may consider the identity satisfied by S_t^N as a closed equation.

From the identity of the lemma, if we assume true for a second that S_t^N weakly converges to a limit measure μ_t, we see (at least intuitively) that μ_t satisfies the nonlinear Fokker-Planck equation (1). To be rigorous, however, we have to deal with the weak convergence of Q^N, the laws of $S_\cdot^N$.

Theorem 15 *If Q is a weak limit point of a subsequence of $\left\{Q^N\right\}_{N\in\mathbb{N}}$, then Q is supported on the measure-valued solutions of the nonlinear Fokker-Planck equation*

$$\frac{\partial \mu_t}{\partial t} = \frac{\sigma^2}{2}\Delta\mu_t - \operatorname{div}\left(\mu_t K * \mu_t\right)$$

with initial condition μ_0.

Proof. For every $\phi \in C_b^{1.2}\left([0,T]\times\mathbb{R}^d\right)$, consider the functional

$$d\left\langle S_t^N, \phi_t\right\rangle = \left\langle S_t^N, \frac{\partial \phi_t}{\partial t}\right\rangle dt + \left\langle S_t^N, \nabla\phi_t \cdot \left(K * S_t^N\right)\right\rangle dt + dM_t^{\phi,N} + \frac{\sigma^2}{2}\left\langle S_t^N, \Delta\phi_t\right\rangle dt.$$

$$\Psi_\phi\left(\mu_\cdot\right) := \sup_{t\in[0,T]}\left|\left\langle\mu_t,\phi_t\right\rangle - \left\langle\mu_0,\phi_0\right\rangle - \int_0^t\left\langle\mu_s, \frac{\partial\phi_s}{\partial t} + \frac{\sigma^2}{2}\Delta\phi_s + \nabla\phi_s\cdot\left(K*\mu_s\right)\right\rangle ds\right|.$$

It is continuous on $C\left([0,T];\mathrm{Pr}_1\left(\mathbb{R}^d\right)\right)$. Hence, if Q^{N_k} is a subsequence which weakly converges to Q, by Portmanteau theorem we have

$$Q\left(\Psi_\phi\left(\mu_\cdot\right) > \delta\right) \le \liminf_{k\to\infty} Q^{N_k}\left(\Psi_\phi\left(\mu_\cdot\right) > \delta\right).$$

But

$$\begin{aligned} Q^{N_k}\left(\Psi_\phi\left(\mu_\cdot\right) > \delta\right) &= P\left(\Psi_\phi\left(S_\cdot^{N_k}\right) > \delta\right) \\ &= P\left(\sup_{t\in[0,T]}\left|\left\langle S_t^{N_k},\phi_t\right\rangle - \left\langle\mu_0,\phi_0\right\rangle - \int_0^t\left\langle S_s^{N_k}, \frac{\partial\phi_s}{\partial t} + \frac{\sigma^2}{2}\Delta\phi_s + \nabla\phi_s\cdot\left(K*S_s^{N_k}\right)\right\rangle ds\right| > \delta\right. \end{aligned}$$

From the identity of Lemma 13 we get

$$\begin{aligned} Q^{N_k}\left(\Psi_\phi\left(\mu_\cdot\right) > \delta\right) &= P\left(\sup_{t\in[0,T]}\left|\left\langle S_0^{N_k} - \mu_0, \phi_0\right\rangle - M_t^{\phi,N_k}\right| > \delta\right) \\ &\le P\left(\left|\left\langle S_0^{N_k} - \mu_0, \phi_0\right\rangle\right| + \sup_{t\in[0,T]}\left|M_t^{\phi,N_k}\right| > \delta\right) \\ &\le P\left(\left|\left\langle S_0^{N_k} - \mu_0, \phi_0\right\rangle\right| > \delta/2\right) + P\left(\sup_{t\in[0,T]}\left|M_t^{\phi,N_k}\right| > \delta/2\right). \end{aligned}$$

The first term goes to zero, as $k \to \infty$, by the assumption on S_0^N. For the second term we have

$$\begin{aligned} P\left(\sup_{t\in[0,T]}\left|M_t^{\phi,N_k}\right| > \delta/2\right) &\leq \frac{4}{\delta^2}E\left[\sup_{t\in[0,T]}\left|M_t^{\phi,N_k}\right|^2\right] \\ &\leq CE\left[\left|M_T^{\phi,N_k}\right|^2\right] \end{aligned}$$

because M_t^{ϕ,N_k} is a martingale. Now, it is well known that

$$E\left[\left(\int_0^t \nabla\phi_s\left(X_s^{i,N_k}\right)\cdot dB_s^i\right)\left(\int_0^t \nabla\phi_s\left(X_s^{i,N_k}\right)\cdot dB_s^j\right)\right] = 0$$

if $i \neq j$ because $B_\cdot^i$ and $B_\cdot^j$ are independent. Hence

$$E\left[\left|M_T^{\phi,N_k}\right|^2\right] = \frac{\sigma^2}{N_k^2}\sum_{i=1}^{N_k} E\left[\left(\int_0^t \nabla\phi_s\left(X_s^{i,N_k}\right)\cdot dB_s^i\right)^2\right].$$

By the isometry formula for Itô integrals, this is equal to

$$= \frac{\sigma^2}{N_k^2}\sum_{i=1}^{N_k}\int_0^t \left|\nabla\phi_s\left(X_s^{i,N_k}\right)\right|^2 ds$$

which is bounded by

$$\leq C\frac{\sigma^2}{N_k}\left\|\nabla\phi\right\|_\infty^2 T \to 0$$

as $k \to \infty$.

Therefore $Q\left(\Psi_\phi\left(\mu_\cdot\right) > \delta\right) = 0$. Since this holds true for every $\delta > 0$, we deduce

$$Q\left(\Psi_\phi\left(\mu_\cdot\right) = 0\right) = 1.$$

Hence, for every $\phi \in C_b^{1.2}\left([0,T]\times\mathbb{R}^d\right)$, Q is supported on the set of measure-valued functions $\mu_\cdot$ such that

$$\langle\mu_t,\phi_t\rangle = \langle\mu_0,\phi_0\rangle + \int_0^t \left\langle \mu_s, \frac{\partial\phi_s}{\partial t} + \frac{\sigma^2}{2}\Delta\phi_s + \nabla\phi_s\cdot(K*\mu_s)\right\rangle ds.$$

This implies that Q is supported on the set of measure-valued solutions of the nonlinear Fokker-Planck equation, by an argument of countable density of functions. ■

4 Uniqueness for the PDE and global limit of the empirical measure

Theorem 16 *The nonlinear (1), with initial condition* $\mu_0 \in \Pr\left(\mathbb{R}^d\right)$*, has one and only one measure-valued solution.*

Proof. Existence has been proved above: the support of any limit measure Q of the family $\left\{Q^N\right\}_{N\in\mathbb{N}}$ is non empty and it is made of such solutions. The proof of uniqueness is postponed. ■

Corollary 17 *The family* $\left\{Q^N\right\}_{N\in\mathbb{N}}$ *of laws of the empirical processes* S_t^N *weakly converge to* δ_μ *where* $\mu_. \in C\left([0,T]\,;\Pr_1\left(\mathbb{R}^d\right)\right)$ *is the unique solution of equation (1).*

Proof. From the compactness of $\left\{Q^N\right\}_{N\in\mathbb{N}}$ we know that from every subsequence we may extract a further subsequence which converges to some limit measure Q, supported on measure-valued solutions of (1). By the previous uniqueness theorem, $Q = \delta_\mu$ where $\mu_. \in C\left([0,T]\,;\Pr_1\left(\mathbb{R}^d\right)\right)$ is the unique solution of equation (1). Since this holds true for all limit points Q, and the weak convergence in $\Pr\left(C\left([0,T]\,;\Pr_1\left(\mathbb{R}^d\right)\right)\right)$ is a metric convergence, we deduce that the full sequence $\left\{Q^N\right\}_{N\in\mathbb{N}}$ converges, to δ_μ. ■