White-noise driven conditional McKean-Vlasov limits for systems of particles with simultaneous and random jumps

Xavier Erny, Eva Löcherbach, Dasha Loukianova

White-noise driven conditional McKean-Vlasov limits for systems of particles with simultaneous and random jumps

Xavier Erny[*], **Eva Löcherbach**[†] **and Dasha Loukianova**[*]

[*] *Université Paris-Saclay, CNRS, Univ Evry, Laboratoire de Mathématiques et Modélisation d'Evry, 91037, Evry, France*

[†] *Statistique, Analyse et Modélisation Multidisciplinaire, Université Paris 1 Panthéon-Sorbonne, EA 4543 et FR FP2M 2036 CNRS*

Abstract: We study the convergence of $N-$particle systems described by SDEs driven by Brownian motion and Poisson random measure, where the coefficients depend on the empirical measure of the system. Every particle jumps with a jump rate depending on its position and on the empirical measure of the system. Jumps are simultaneous, that is, at each jump time, all particles of the system are affected by this jump and receive a random jump height that is centred and scaled in $N^{-1/2}$. This particular scaling implies that the limit of the empirical measures of the system is random, describing the conditional distribution of one particle in the limit system. We call such limits *conditional McKean-Vlasov limits*. The conditioning in the limit measure reflects the dependencies between coexisting particles in the limit system such that we are dealing with a *conditional propagation of chaos property*. As a consequence of the scaling in $N^{-1/2}$ and of the fact that the limit of the empirical measures is not deterministic the limit system turns out to be solution of a non-linear SDE, where not independent martingale measures and white noises appear having an intensity that depends on the conditional law of the process.

MSC 2010 subject classifications: 60K35, 60G09, 60H40, 60F05.
Keywords and phrases: Martingale measures, McKean-Vlasov equations, Mean field interaction, Interacting particle systems, Propagation of chaos, Exchangeability.

1. Introduction

McKean-Vlasov equations are stochastic differential equations where the coefficients depend on the distribution of the solution. Such equations typically arise as limits of mean field $N-$particle systems, where the coefficients depend on the empirical measure of the system. These kind of limits are referred to as *McKean-Vlasov limits* (see e.g. Gärtner (1988), Graham (1992) and Andreis, Dai Pra and Fischer (2018)).

In this paper we extend these limits to a rather general class of *conditional McKean-Vlasov limits*. More precisely, we consider the system of interacting particles with a diffusive term and jumps, given by

$$\begin{aligned} dX_t^{N,i} &= b(X_t^{N,i}, \mu_t^N)dt + \sigma(X_t^{N,i}, \mu_t^N)d\beta_t^i \\ &+ \frac{1}{\sqrt{N}} \sum_{k=1, k\neq i}^{N} \int_{\mathbb{R}_+ \times E} \Psi(X_{t-}^{N,k}, X_{t-}^{N,i}, \mu_{t-}^N, u^k, u^i) \mathbb{1}_{\{z \leqslant f(X_{t-}^{N,k}, \mu_{t-}^N)\}} d\pi^k(t,z,u), 1 \leqslant i \leqslant N, \end{aligned} \tag{1}$$

starting from the initial condition

$$(X_0^{N,i})_{1\leqslant i\leqslant N} \sim \nu_0^{\otimes N}.$$

Here $\mu_t^N = N^{-1}\sum_{j=1}^N \delta_{X_t^{N,j}}$ is the empirical measure of the system, β^i $(i \geqslant 1)$ are i.i.d. one-dimensional standard Brownian motions and π^k $(k \geqslant 1)$ i.i.d. Poisson measures on $\mathbb{R}_+ \times \mathbb{R}_+ \times E$, where $E = \mathbb{R}^{\mathbb{N}^*}$, $\mathbb{N}^* = \{1, 2, 3, \ldots\}$. Each π^k has intensity $ds \cdot dz \cdot \nu(du)$, with ν a product probability measure on E. The initial distribution ν_0 is a probability measure on $\mathbb{R}$ having a finite second moment. We assume that the Poisson measures, the Brownian motions and the initial conditions are independent. Since the jumps are scaled in $N^{-1/2}$, to prevent the jump term from exploding, we suppose that the height of the jump term is centred (see Assumption 3 below), that is, for all $x, y \in \mathbb{R}$, for all probability measures m on $\mathbb{R}$,

$$\int_E \Psi(x, y, m, u^1, u^2) d\nu(u) = 0.$$

Our model is close to the one considered in Andreis, Dai Pra and Fischer (2018). As there, any particle in position x jumps at a rate $f(x, m)$, whenever m is the common state of the system, that is, the current value of the empirical measure. Mainly motivated by applications coming from neuroscience (where jumps are spikes of the neurons leading to an increase of the potential of all other neurons, see e.g. De Masi et al. (2015) and Fournier and Löcherbach (2016), or Duarte, Ost and Rodríguez (2015) for a spatially structured model), jumps are simultaneous, that is, all particles in the system are affected by any of the jumps. More precisely, the random jump height depends both on the current position y of the particle receiving the jump and on the position x of the particle that causes the jump. Notice that contrarily to Andreis, Dai Pra and Fischer (2018) we do not include auto-interactions induced by jumps in the equation (1), i.e. terms of the type

$$\int_{\mathbb{R}_+\times E} \Theta(X_{t-}^{N,i}, \mu_{t-}^N, u^i) \mathbb{1}_{\{z\leqslant f(X_{t-}^{N,i}, \mu_{t-}^N)\}} d\pi^i(t, z, u).$$

Indeed, such terms would survive in the large population limit leading to discontinuous trajectories, and the presence of the indicator $\mathbb{1}_{\{z\leqslant f(X_{t-}^{N,i}, \mu_{t-}^N)\}}$ requires to work both in L^1 and in L^2 (see Graham (1992), see also Erny, Löcherbach and Loukianova (2020) where we dealt both with simultaneous small jumps and big ones). In the present paper, we decided to disregard these big jumps to focus on the very specific form of the limit process given in (3) below.

Coming back to Andreis, Dai Pra and Fischer (2018), the main difference to our work is that there the averaging regime is considered: the common contribution of all particles to the dynamic of a given particle, represented in (1) by the sum of the stochastic integrals with respect to the Poisson random measures, is scaled in N^{-1}. In this situation it was shown in Andreis, Dai Pra and Fischer (2018) that the *Propagation of chaos* phenomenon holds: the coordinates are i.i.d. in the limit. Moreover, the limit of the empirical measures is the distribution of any coordinate of the limit system, and the dynamic of one coordinate is described by a classical McKean-Vlasov equation.

The novelty of the present paper is that we consider (1) in a diffusive regime, where the common contribution of all particles to the dynamic of a given particle is scaled in $N^{-1/2}$. It has already been observed that this diffusive scaling gives rise to the *Conditional propagation of chaos* property (see Erny, Löcherbach and Loukianova (2020)): a common noise appears in the limit system, and the coordinates of the limit system are conditionally i.i.d given this common noise. Moreover the limit of the empirical measures is shown to be the conditional distribution of any coordinate of the

limit system, given the common noise. In Erny, Löcherbach and Loukianova (2020) the common noise is a Brownian motion created by the contribution of the jumps of all particles in the dynamic of a given particle, as a consequence of the scaling $1/\sqrt{N}$ and the central limit theorem.

It turns out that in the present work, to describe the precise dynamic of the limit, and in particular to identify the common noise, we need to rely on martingale measures and white noises (see Walsh (1986) and El Karoui and Méléard (1990)) as driving measures. More precisely, the limit system will be shown to be solution of a non-linear SDE driven by (white noise) martingale measures having an intensity that depends on the conditional law of the system itself. These martingale measures do only appear in the limit system as a consequence of the central limit theorem and the joint contribution of all small and centered jumps. The main reason for the appearance of the martingale measures instead of the Brownian motion is the spatial correlation of the finite system, i.e. the dependence on the positions both of the particle giving and the one receiving the input.

To the best of our knowledge, this is the first time that McKean-Vlasov limits are considered where the underlying driving martingale measures are only present in the limit system, but not at the level of the $N-$particle system. We refer however to Chevallier and Ost (2020) who work in the averaging regime and study the fluctuations of a stochastic system, associated to spatially structured Hawkes processes, around its mean field limit, and where particles in the mean field limit are still independent.

Processes driven by martingale measures having an intensity that depends on the law of the process itself have already appeared in the literature related to particle approximations of Boltzmann's equation, starting with the classical article by Tanaka (1978) that gave rise to a huge literature (to cite juste a few, see Graham and Méléard (1997), Meleard, Sylvie (1998), Fournier and Meleard (2002), Fournier and Mischler (2016)). In these papers, the underlying random measure is Poisson, and the dependence on the law arises at the level of the particle system that is designed to approximate Boltzmann's equation. In our work, the underlying random measure is white noise since jumps disappear in the limit, and the dependence on the (conditional) law of the process does only appear in the limit, as an effect of the conditional propagation of chaos.

Let us now describe the limit system associated to (1). To find its precise form, we mainly need to understand the limits of the martingales which are the jump terms of the system, given by

$$J_t^{N,i} = \frac{1}{\sqrt{N}} \sum_{k=1,k\neq i}^{N} \int_{[0,t]\times\mathbb{R}_+\times E} \Psi(X_{s-}^{N,k}, X_{s-}^{N,i}, \mu_{s-}^N, u^k, u^i) \mathbb{1}_{\{z\leqslant f(X_{s-}^{N,k}, \mu_{s-}^N)\}} d\pi^k(s,z,u).$$

In what follows we consider some concrete examples of Ψ and give the limits of the corresponding predictable quadratic covariations of $J_t^{N,i}$ to have a better understanding of its limit. We shall always assume that the jump rate function f is bounded. Let us begin with a situation close to that of Erny, Löcherbach and Loukianova (2020), where in the limit system each coordinate shares a common Brownian motion W.

Example 1. *Suppose that* $\Psi(x,y,m,u^1,u^2) = \Psi(u^1)$. *Then we have, for all* $1 \leqslant i,j \leqslant N$,

$$\langle J^{N,i}, J^{N,j}\rangle_t = \frac{1}{N} \sum_{k=1,k\neq i,j}^{N} \int_0^t \int_{\mathbb{R}} \Psi(u^k)^2 f(X_s^{N,k}, \mu_s^N) d\nu_1(u^k) ds$$
$$= \varsigma^2 \int_0^t \int_{\mathbb{R}} f(x, \mu_s^N) \mu_s^N(dx) ds + O(\frac{t}{N}),$$

since f is bounded, with $\varsigma^2 := \int_{\mathbb{R}} \Psi(u^1)^2 d\nu_1(u^1)$ and ν_1 the projection of ν on the first coordinate. Denote μ the limit of the empirical measures μ^N. Then the angle brackets process should converge as N goes to infinity to

$$\varsigma^2 \int_0^t \int_{\mathbb{R}} f(x, \mu_s) \mu_s(dx) ds.$$

As the limit quadratic covariations are non-null, in the limit system there will be a common Brownian motion W underlying each particle's motion. Thus, the limit system is given by

$$d\bar{X}_t^i = b(\bar{X}_t^i, \mu_t)dt + \sigma(\bar{X}_t^i, \mu_t)d\beta_t^i + \varsigma \sqrt{\int_{\mathbb{R}} f(x, \mu_t)\mu_t(dx)} dW_t,$$

where W is a standard one-dimensional Brownian motion. We will also show that $\mu = \mathcal{L}(\bar{X}^1|W)$, since μ is necessarily the directing measure of $(\bar{X}^i)_{i\geqslant 1}$. In particular, the conditioning in μ reflects the presence of some common noise, which is W here.

Now, let us consider an opposite situation where, in the limit system, each coordinate has its own Brownian motion W^i, and where these Brownian motions are independent.

Example 2. *In this example, we assume that $\Psi(x, y, m, u^1, u^2) = \Psi(u^2)$. As in the previous example, we begin by computing the angle brackets of the jump terms between particles i and j. Here we distinguish two cases: $i \neq j$ and $i = j$. If $i \neq j$, using the fact that ν is a product measure and that Ψ is centered (Assumption 3),*

$$\langle J^{N,i}, J^{N,j} \rangle_t = \frac{1}{N} \sum_{k=1, k\neq i,j}^{N} \int_0^t \int_E \Psi(u^i)\Psi(u^j) f(X_s^{N,k}, \mu_s^N) d\nu(u) ds = 0.$$

Moreover, if $i = j$,

$$\langle J^{N,i} \rangle_t = \frac{1}{N} \sum_{k=1, k\neq i}^{N} \int_0^t \int_{\mathbb{R}} \Psi(u^i)^2 f(X_s^{N,k}, \mu_s^N) d\nu_1(u^i) ds = \varsigma^2 \int_0^t \int_{\mathbb{R}} f(x, \mu_s^N)\mu_s^N(dx)ds + O(\frac{t}{N}).$$

As the quadratic covariations between different particles are null, there will be no common noise in the limit system. So, instead of having one common Brownian motion W as in the previous example, here, each particle is driven by its own Brownian motion. More precisely, in this example the limit system is

$$d\bar{X}_t^i = b(\bar{X}_t^i, \mu_t)dt + \sigma(\bar{X}_t^i, \mu_t)d\beta_t^i + \varsigma \sqrt{\int_{\mathbb{R}} f(x, \mu_t)\mu_t(dx)} dW_t^i,$$

where W^i $(i \geqslant 1)$ are independent standard one-dimensional Brownian motions, independent of β^i $(i \geqslant 1)$, and where $\mu = \mathcal{L}(\bar{X}^1)$ is deterministic in this particular case.

Finally let us show an example where, as in Example 1, each particle shares a common Brownian motion W, and, as in Example 2, each particle has also its own Brownian motion W^i, and where both W and W^i are produced by the common contribution of the small jumps.

Example 3. *Here we assume that* $\Psi(x,y,m,u^1,u^2)=\Psi(u^1,u^2)$. *The angle brackets of the jump terms of the particles* i *and* j *are, if* $i\neq j$,

$$\langle J^{N,i},J^{N,j}\rangle_t=\frac{1}{N}\sum_{k=1,k\neq i,j}^{N}\int_0^t\int_E\Psi(u^k,u^i)\Psi(u^k,u^j)f(X_s^{N,k},\mu_s^N)d\nu(u)ds$$
$$=\xi^2\int_0^t\int_{\mathbb{R}}f(x,\mu_s^N)\mu_s^N(dx)ds+O(\frac{t}{N}),$$

where we know that $\xi^2:=\int_E\Psi(u^1,u^2)\Psi(u^1,u^3)d\nu(u)\geqslant 0$ *since it is the covariance of the infinite exchangeable sequence* $(\Psi(U^1,U^k))_{k\geqslant 2}$, *where* $(U^k)_{k\geqslant 1}\sim\nu$. *And if* $i=j$,

$$\langle J^{N,i}\rangle_t=\frac{1}{N}\sum_{k=1,k\neq i}^{N}\int_0^t\int_E\Psi(u^i,u^k)^2f(X_s^{N,k},\mu_s^N)d\nu(u)ds=\varsigma^2\int_0^t\int_{\mathbb{R}}f(x,\mu_s^N)\mu_s^N(dx)ds+O(\frac{t}{N}),$$

where $\varsigma^2=\int_E\Psi(u^1,u^2)^2d\nu(u)$.

As in Example 1, there must be a common Brownian motion since the quadratic covariations between different particles are not zero. But, here $\langle J^{N,i},J^{N,j}\rangle_t\neq\langle J^{N,i}\rangle_t$ *if* $j\neq i$. *That is why there must be additional Brownian motions. Formally, the limit system in this example is*

$$\begin{aligned}d\bar{X}_t^i=&b(\bar{X}_t^i,\mu_t)dt+\sigma(\bar{X}_t^i,\mu_t)d\beta_t^i\\&+\xi\sqrt{\int_{\mathbb{R}}f(x,\mu_t)\mu_t(dx)}dW_t+\sqrt{(\varsigma^2-\xi^2)\int_{\mathbb{R}}f(x,\mu_t)\mu_t(dx)}dW_t^i,\end{aligned}\tag{2}$$

where we know that $\varsigma^2\geqslant\xi^2$ *by Cauchy-Schwarz's inequality. As before,* W,W^i $(i\geqslant 1)$ *are independent standard one-dimensional Brownian motions, and* $\mu=\mathcal{L}(\bar{X}^1|W)$ *is random in this case. Note that in the case where* $\Psi(x,y,m,u^1,u^2)=\Psi(u_1)$, *we have* $\xi^2=\varsigma^2$, *and in the case* $\Psi(x,y,m,u^1,u^2)=\Psi(u_2)$, *we have* $\xi^2=0$, *hence this example covers both Example (1) and Example (2).*

Before defining the limit system in the general case, let us explain the main difficulty that arises. If we apply the same reasoning as in Examples 1, 2 and 3 to the general model given in (1), we obtain for two different particles

$$\begin{aligned}&\langle J^{N,i},J^{N,j}\rangle_t=\\&\frac{1}{N}\sum_{k=1,k\neq i,j}^{N}\int_0^t f(X_s^{N,k},\mu_s^N)\int_E\Psi(X_s^{N,k},X_s^{N,i},\mu_s^N,u^k,u^i)\Psi(X_s^{N,k},X_s^{N,j},\mu_s^N,u^k,u^j)\nu(du)ds\\&=\int_0^t\int_{\mathbb{R}}f(x,\mu_s^N)\int_E\Psi(x,X_s^{N,i},\mu_s^N,u^1,u^2)\Psi(x,X_s^{N,j},\mu_s^N,u^1,u^3)\nu(du)\mu_s^N(dx)ds+O(\frac{t}{N}),\end{aligned}$$

under appropriate conditions on Ψ, see Assumption 3 below. And for the quadratic variation of the jump term of a single particle we get

$$\langle J^{N,i}\rangle_t=\int_0^t\int_{\mathbb{R}}f(x,\mu_s^N)\int_E\Psi(x,X^{N,i},\mu_s^N,u^1,u^2)^2\nu(du)\mu_s^N(dx)ds,$$

still up to an error term of order $1/N$.

Contrarily to the situation of the previous examples, the quadratic covariations depend on the positions of the particles i and j and can only be written as integrals of products where this integration involves, among others, the empirical measure of the process. This is the reason why we need to use martingale measures and white noises instead of Brownian motions, as introduced in Walsh (1986), confer also to El Karoui and Méléard (1990).

Let us briefly explain why using martingale measures is well adapted to our problem. If M is a martingale measure on $\mathbb{R}_+ \times F$ (with $(F, \mathcal{F})$ some measurable space), having intensity $dt \cdot m_t(dy)$, then for all $A, B \in \mathcal{F}$, $M_t(A) := M([0,t] \times A); t \geqslant 0$ is a square-integrable martingale and

$$\langle M.(A), M.(B) \rangle_t = \int_0^t \int_F \mathbb{1}_{A \cap B}(y) m_s(dy) ds.$$

Having this remark in mind, it is natural to write the limit system in a similar way as (2), but replacing the Brownian motions by martingale measures. More precisely, under appropriate conditions on the coefficients, the limit system $(\bar{X}^i)_{i \geqslant 1}$ of (1) will be shown to be of the form

$$\begin{aligned} d\bar{X}_t^i &= b(\bar{X}_t^i, \mu_t)dt + \sigma(\bar{X}_t^i, \mu_t)d\beta_t^i + \int_{\mathbb{R}} \int_{\mathbb{R}} \sqrt{f(x, \mu_t)} \tilde{\Psi}(x, \bar{X}_t^i, \mu_t, v) dM(t, x, v), \qquad (3) \\ &\quad + \int_{\mathbb{R}} \sqrt{f(x, \mu_t)} \kappa(x, \bar{X}_t^i, \mu_t) dM^i(t, x), \ i \geqslant 1, \\ (\bar{X}_0^i)_{i \geqslant 1} &\sim \nu_0^{\otimes \mathbb{N}^*}. \end{aligned}$$

In the above formula,

$$\mu_t := \mathcal{L}(\bar{X}_t^i | W), \qquad (4)$$

$$\tilde{\Psi}(x, y, m, v) := \int_{\mathbb{R}} \Psi(x, y, m, v, u^1) d\nu_1(u^1), \qquad (5)$$

$$\begin{aligned} \kappa(x, y, m)^2 &:= \int_E \Psi(x, y, m, u^1, u^2)^2 d\nu(u) - \int_{\mathbb{R}} \tilde{\Psi}(x, y, m, v)^2 d\nu_1(v) \\ &= \int_E \Psi(x, y, m, u^1, u^2)^2 d\nu(u) - \int_E \Psi(x, y, m, u^1, u^2) \Psi(x, y, m, u^1, u^3) d\nu(u). \qquad (6) \end{aligned}$$

Notice that the expression (6) is positive by Cauchy-Schwarz's inequality.

In the above equations, $M(dt, dx, dv)$ and $M^i(dt, dx)$ are orthogonal martingale measures on $\mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}$ ($\mathbb{R}_+ \times \mathbb{R}$ respectively) with respective intensities $dt \cdot \mu_t(dx) \cdot \nu_1(dv)$ and $dt \cdot \mu_t(dx)$, defined as

$$M_t^i(A) := \int_0^t \mathbb{1}_A(F_s^{-1}(p)) dW^i(s, p) \text{ and } M_t(A \times B) := \int_0^t \mathbb{1}_A(F_s^{-1}(p)) \mathbb{1}_B(v) dW(s, p, v), \qquad (7)$$

with W a white noise on $\mathbb{R}_+^2 \times \mathbb{R}$ with intensity $dt \cdot dp \cdot d\nu_1(v)$, and W^i ($i \geqslant 1$) independent white noises on $\mathbb{R}_+^2$, independent from W, with intensity $dt \cdot dp$. In the above formula, $F_s(x) := P\left(\bar{X}_s^i \leqslant x | W\right)$ is the conditional distribution function, conditionally on W, and F_s^{-1} is the generalized inverse of F_s. As in (1), we assume that the Brownian motions, the white noises and the initial conditions are independent.

In the case where $\Psi(x,y,m,u_1,u_2)=\Psi(u_1,u_2)$, we see that $\kappa^2=\varsigma^2-\xi^2$ is a constant and that $\tilde{\Psi}$ does only depend on v such that we can represent the two integrals with respect to the martingale measures in (3) as two integrals against Brownian motions, recovering all previous examples.

Let us give some comments on the above system of equations. We have already argued that, in general, μ_t is a random measure because of the scaling $N^{-1/2}$. We shall prove that μ_t is actually the law of $\bar{X}^1$ conditionally on the common noise of the system. This common noise is the white noise W underlying the martingale measure M. It is not immediately obvious that the definition of the martingale measures M and M^i in (7) and the limit system (3) are well-posed. In what follows, we shall give conditions ensuring that equation (3) admits a unique strong solution. This is the content of our first main theorem, Theorem 2.3. To prove this theorem, we propose a Picard iteration in which we construct a sequence of martingale measures whose intensities depend on the conditional law of the instance of the process within the preceding step. One main ingredient of the proof is the well-known fact that the Wasserstein-2−distance of the laws of two real-valued random variables is given by the L^2- distance of their inverse distribution functions - we apply this fact here to the conditional distribution functions.

Using arguments that are inspired by Erny, Löcherbach and Loukianova (2020), we then show in our second main theorem, Theorem 2.5, that the finite particle system converges to the limit system, that is, $(X^{N,i})_{1\leqslant i\leqslant N}$ converges to $(\bar{X}^i)_{i\geqslant 1}$ in distribution in $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}$. This convergence is the consequence of the well-posedness of an associated martingale problem. Contrarily to Erny, Löcherbach and Loukianova (2020) the finite system here depends on the empirical measure, and the conditions on the regularity of its coefficients are formulated in terms of the Wasserstein distance. To reconcile the convergence in distribution of the empirical measure with the Wasserstein-Lipschitz continuity of the coefficients gives an additional technical difficulty to the proof.

Organization of the paper. In Section 2, we state the assumptions and formulate the main results. Section 3 is devoted to the proof of Theorem 2.3. The proofs of Theorems 2.4 and 2.5 are gathered in Section 4. Finally, in Section 5 we discuss extensions of our results to the frame of multi-populations where the particles are organized within clusters.

General notation. Throughout this paper we shall use the following notation. Given any measurable space $(S,\mathcal{S})$, $\mathcal{P}(S)$ denotes the set of all probability measures on $(S,\mathcal{S})$, endowed with the topology of weak convergence. For $p\in\mathbb{N}^*$, $\mathcal{P}_p(\mathbb{R})$ denotes the set of probability measures on $\mathbb{R}$ that have a finite moment of order p. For two probability measures $\nu_1,\nu_2\in\mathcal{P}_p(\mathbb{R})$, the Wasserstein distance of order p between ν_1 and ν_2 is defined as

$$W_p(\nu_1,\nu_2)=\inf_{\pi\in\Pi(\nu_1,\nu_2)}\left(\int_S\int_S|x-y|)^p\pi(dx,dy)\right)^{1/p},$$

where π varies over the set $\Pi(\nu_1,\nu_2)$ of all probability measures on the product space $\mathbb{R}\times\mathbb{R}$ with marginals ν_1 and ν_2. Notice that the Wasserstein distance of order p between ν_1 and ν_2 can be rewritten as the infimum of $E\big[|X-Y|^p\big]^{1/p}$ over all possible couplings (X,Y) of the random elements X and Y distributed according to ν_1 and ν_2 respectively, i.e.

$$W_p(\nu_1,\nu_2)=\inf\left\{\mathbb{E}\left[|X-Y|^p\right]^{1/p}:\mathcal{L}(X)=\nu_1\text{ and }\mathcal{L}(Y)=\nu_2\right\}.$$

Moreover, $D(\mathbb{R}_+,\mathbb{R})$ (or just D for short) denotes the space of càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}$, endowed with the Skorokhod metric, and C and K denote arbitrary positive constants whose values can change from line to line in an equation. We write C_θ and K_θ if the constants depend on

some parameter θ. Finally, for any $n, p \in \mathbb{N}^*$, we note $C_b^n(\mathbb{R}^p)$ (resp. $C_b^n(\mathbb{R}^p, \mathbb{R}_+)$) the set of real-valued functions g (resp. non-negative functions g) defined on $\mathbb{R}^p$ which are n times continuously differentiable such that $g^{(k)}$ is bounded for each $0 \leqslant k \leqslant n$.

2. Assumptions and main results.

2.1. Assumptions

We start imposing a hypothesis under which equation (1) admits a unique strong solution and which grants a Lipschitz condition on the coefficients of the SDE.

Assumption 1.

i) For all $x, y \in \mathbb{R}, m, m' \in \mathcal{P}_1(\mathbb{R})$,

$$|b(x,m) - b(y,m')| + |\sigma(x,m) - \sigma(y,m')| \leqslant C(|x-y| + W_1(m,m')).$$

ii) f *is bounded and strictly positive, and* $\sqrt{f}$ *is Lipschitz, that is, for all* $x, y \in \mathbb{R}, m, m' \in \mathcal{P}_1(\mathbb{R})$,

$$|\sqrt{f}(x,m) - \sqrt{f}(y,m')| \leqslant C(|x-y| + W_1(m,m')).$$

iii) For all $x, x', y, y', u, v \in \mathbb{R}, m, m' \in \mathcal{P}_1(\mathbb{R})$,

$$|\Psi(x,y,m,u,v) - \Psi(x',y',m',u,v)| \leqslant M(u,v)(|x-x'| + |y-y'| + W_1(m,m')),$$

where $M : \mathbb{R}^2 \to \mathbb{R}_+$ *satisfies* $\int_E M(u^1,u^2)^2 d\nu(u) < \infty$.

iv)

$$\sup_{x,y,m} \int_E |\Psi(x,y,m,u^1,u^2)| d\nu(u) < \infty.$$

Notice that f bounded together with $\sqrt{f}$ Lipschitz implies that f is Lipschitz as well. As a consequence, relying on Theorem 2.1 of Graham (1992), Assumption 1 implies that equation (1) admits a unique strong solution.

In order to prove the well-posedness of the limit equation (3), we need additional assumptions. Recall that κ^2 has been introduced in (6) above.

Assumption 2.

i)

$$\inf_{x,y,m} \kappa(x,y,m) > 0,$$

ii)

$$\sup_{x,y,m} \int_E \Psi(x,y,m,u^1,u^2)^2 d\nu(u) < \infty.$$

Remark 2.1. *Using the third point of Assumption 1 we can prove that* κ^2 *is Lipschitz continuous with Lipschitz constant proportional to*

$$\left(\int_E M^2(u,v) d\nu(u,v) \times \sup_{x,y,m} \int_E \Psi^2(x,y,m,u,v) d\nu(u,v) \right)^{1/2}.$$

Assumption 2.i) allows then to prove that κ *is Lipschitz continuous. Assumption 2.ii) gives that* $\|\kappa\|_\infty := \sup_{x,y,m} \kappa(x,y,m) < \infty$.

To prove the convergence of the particle system $(X^{N,i})_{1\leqslant i\leqslant N}$ to the limit system, we need further assumptions on the function Ψ.

Assumption 3.

i) For all $x, y \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})$,

$$\int_E \Psi(x, y, m, u^1, u^2) d\nu(u) = 0,$$

ii)

$$\int_E \sup_{x,y,m} |\Psi(x, y, m, u^1, u^2)|^3 d\nu(u) < \infty.$$

iii) b *and* σ *are bounded.*

Remark 2.2. *We assume the functions* b *and* σ *to be bounded to simplify the proofs of Lemmas 6.1 and 6.2. However the results of these lemmas still hold true under the following weaker assumption: there exists* $C > 0$ *such that, for all* $x \in \mathbb{R}, m \in \mathcal{P}(\mathbb{R})$,

$$|b(x, m)| + |\sigma(x, m)| \leqslant C(1 + |x|).$$

In other words, b *and* σ *are bounded w.r.t. the measure variable and sublinear w.r.t. the space variable.*

Let us give an example of a function Ψ that satisfies all our assumptions and where the random quantity Ψ depends on the difference of the states of the jumping and the receiving particle as well as on the average state of the system as follows

$$\Psi(x, y, m, u, v) = uv \left(\varepsilon + \frac{\pi}{2} + \arctan(x - y + \int_{\mathbb{R}} z dm(z)) \right),$$

with $\nu = \mathcal{R}^{\otimes \mathbb{N}^*}$ and $\mathcal{R} = \frac{1}{2}(\delta_{-1} + \delta_{+1})$ the Rademacher distribution. In the formula above, the variables u and v can be seen as spins such that the receiving particle is excited if the orientation of its spin is the same as the spin of the sending particle, and inhibited otherwise. Ψ satisfies the Lipschitz condition of Assumption 1 because arctan is Lipschitz continuous. The hypothesis on the moments of Ψ are also satisfied since $\mathcal{N}(0, 1)$ has finite third moments and is centered. Finally, the first point of Assumption 2 holds true, because

$$\begin{aligned}
\kappa(x, y, m)^2 &= \int_E \Psi(x, y, m, u^1, u^2)^2 d\nu(u) - \int_E \Psi(x, y, m, u^1, u^2)\Psi(x, y, m, u^1, u^3) d\nu(u) \\
&= \left(\varepsilon + \frac{\pi}{2} + \arctan(x - y + \int_{\mathbb{R}} z dm(z)) \right)^2 \left(\int_E (u^1 u^2)^2 d\nu(u) - \int_E u^1 u^2 u^1 u^3 d\nu(u) \right) \\
&= \left(\varepsilon + \frac{\pi}{2} + \arctan(x - y + \int_{\mathbb{R}} z dm(z)) \right)^2.
\end{aligned}$$

2.2. Main results

Our first main result is the well-posedness of the limit equation.

Theorem 2.3. *Under Assumptions 1 and 2, equation (3) admits a unique strong solution $\bar{X}^i$ that possesses finite second moments. This solution also has finite fourth moments.*

Our second main result states the convergence of $(X^{N,i})_{1\leqslant i\leqslant N}$ to $(\bar{X}^i)_{i\geqslant 1}$.

Theorem 2.4. *Under Assumptions 1, 2 and 3, $(X^{N,i})_{1\leqslant i\leqslant N}$ converges to $(\bar{X}^i)_{i\geqslant 1}$ in distribution in $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}$.*

In the above statement, we implicitly define $X^{N,i}:=0$ if $i>N$.

As the systems $(X^{N,i})_{1\leqslant i\leqslant N}$ $(N\in\mathbb{N}^*)$ and $(\bar{X}^i)_{i\geqslant 1}$ are exchangeable, Theorem 2.4 is equivalent to

Theorem 2.5. *Under Assumptions 1, 2 and 3, the system $(\bar{X}^i)_{i\geqslant 1}$ is exchangeable with directing measure $\mu=\mathcal{L}(\bar{X}^1|W)$, where W is as in (7). Moreover, the sequence of empirical measures*

$$\mu^N:=N^{-1}\sum_{i=1}^N\delta_{X^{N,i}}$$

converges in law, as $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$- valued random variables, to μ.

Remark 2.6. *In Theorem 2.5, it is easy to prove that $\mathcal{L}(\bar{X}^1|W)$ is the directing measure of the system $(\bar{X}^i)_{i\geqslant 1}$. Indeed, it is sufficient to notice that conditionally on W, the variables $\bar{X}^i$ $(i\geqslant 1)$ are i.i.d. and to apply Lemma (2.12) of Aldous (1983).*

The proof of Theorem 2.5 is similar to the proof of Theorem 1.7 of Erny, Löcherbach and Loukianova (2020). It consists in showing that $(\mu^N)_N$ is tight on $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$, and that each converging subsequence converges to the same limit, using a convenient martingale problem. For this reason, in what follows we just give the proofs that substantially change compared to our previous paper. For the other proofs, we just give the main ideas and cite precisely the corresponding statement of Erny, Löcherbach and Loukianova (2020) that allows to conclude.

3. Proof of Theorem 2.3

The following Lemma shows that the definition (7) indeed defines the right martingale measures.

Lemma 3.1. *Let ν_1 be a probability measure on $\mathbb{R}$. Moreover, let $(\Omega,\mathcal{A},P)$ be a probability space, $(\mathcal{G}_t)_t$ be a filtration on it, $(\mathcal{F}_t)_t$ be a sub-filtration of $(\mathcal{G}_t)_t$ and W (resp. W^1) be a $(\mathcal{G}_t)_t-$white noise on $\mathbb{R}_+\times[0,1]\times\mathbb{R}$ (resp. $\mathbb{R}_+\times[0,1]$) of intensity $dt\cdot dp\cdot\nu_1(dv)$ (resp. $dt\cdot dp$). Let X be a continuous $\mathbb{R}-$valued process which is $(\mathcal{G}_t)_t$-adapted and $F_s(x):=\mathbb{P}(X_s\leqslant x|\mathcal{F}_s)$. Moreover, we suppose that for all $s>0$, $\mathbb{P}(X_s\leqslant x|\mathcal{F}_s)=\mathbb{P}(X_s\leqslant x|\mathcal{F}_\infty)$, where $\mathcal{F}_\infty=\sigma\{\mathcal{F}_t,\ t\geqslant 0\}$.*

Define for any $A,B\in\mathcal{B}(\mathbb{R})$,

$$M_t^1(A):=\int_0^t\int_0^1\mathbb{1}_A((F_s)^{-1}(p))dW^1(s,p),\quad M_t(A\times B):=\int_0^t\int_0^1\int_{\mathbb{R}}\mathbb{1}_A((F_s)^{-1}(p))\mathbb{1}_B(v)dW(s,p,v)$$

Then M^1 and M are martingale measures with respective intensities $dt\cdot\mu_t(dx)$ and $dt\cdot\mu_t(dx)\cdot\nu_1(dv)$, where $\mu_t:=\mathcal{L}(X_t|\mathcal{F}_t)$.

Proof. We only show the result for the martingale measure M^1. The main part of the proof consists in showing that the process $(\omega,s,p)\in\Omega\times\mathbb{R}_+\times[0,1]\mapsto(F_s)^{-1}(p)$ is $\mathcal{P}\otimes\mathcal{B}([0,1])-$measurable, with $\mathcal{P}$ the predictable sigma field related to the filtration $(\mathcal{G}_t)_t$.

To begin with, let us prove that $(\omega, s, x) \mapsto F_s(x)$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})-$measurable. We write

$$F_s(x) = P(X_s \leqslant x|\mathcal{F}_s) = \mathbb{E}\left[\varphi(x, X_s)|\mathcal{F}_s\right],$$

where $\varphi(x, y) := \mathbb{1}_{\{x \leqslant y\}}$. As φ is product measurable and bounded, it is the limit of functions of the form

$$\sum_{k=1}^{n} c_k \varphi_k(x)\psi_k(y),$$

where the functions φ_k, ψ_k $(1 \leqslant k \leqslant n)$ are Borel measurable and bounded. This limit can be taken to be increasing such that, by monotone convergence,

$$\mathbb{E}\left[\varphi(x, X_s)|\mathcal{F}_s\right] = \lim_n \sum_{k=1}^{n} c_k \varphi_k(x)\mathbb{E}\left[\psi_k(X_s)|\mathcal{F}_s\right].$$

Then, as ψ_k is a bounded and Borel function, it can be approximated by an increasing sequence of bounded and continuous functions $\psi_{k,m}$. Then, as for every n, m, the process

$$(\omega, s, x) \mapsto \sum_{k=1}^{n} c_k \varphi_k(x)\mathbb{E}\left[\psi_{k,m}(X_s)|\mathcal{F}_s\right] = \sum_{k=1}^{n} c_k \varphi_k(x)\mathbb{E}\left[\psi_{k,m}(X_s)|\mathcal{F}_\infty\right]$$

is continuous in s and $(\mathcal{F}_s)_s-$, whence $(\mathcal{G}_s)_s-$adapted, it is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})-$measurable.

Let $x \in \mathbb{R}$ be fixed. It is sufficient to show that $\{(\omega, s, p) : (F_s)^{-1}(p) \geqslant x\}$ is measurable. Let us write

$$\{(\omega, s, p) : (F_s)^{-1}(p) \geqslant x\} = \{(\omega, s, p) : F_s(x) \leqslant p\} = \{(\omega, s, p) : \varphi(F_s(x), p) > 0\},$$

where $\varphi(x, p) := \mathbb{1}_{\{x \leqslant p\}}$ is product measurable.

Then, the measurability of $\{(\omega, s, p) : (F_s)^{-1}(p) \geqslant x\}$ is a consequence of that of $(\omega, s, p) \mapsto \varphi(F_s(x), p)$ w.r.t. $\mathcal{P} \otimes \mathcal{B}([0, 1])$.

As a consequence, the process $(\omega, s, p) \in \Omega \times \mathbb{R}_+ \times [0, 1] \mapsto (F_s)^{-1}(p)$ is $\mathcal{P} \otimes \mathcal{B}([0, 1])-$measurable. The rest of the proof consists in writing

$$\langle M_\cdot^1(A)\rangle_t = \int_0^t \int_0^1 \mathbb{1}_A((F_s)^{-1}(p))dpds = \int_0^t \mu_s(A)ds.$$

The last inequality above is a classical property of the generalized inverse of the distribution function (see e.g. Fact 1 in Section 8 of Major (1978)). □

3.1. Construction of a strong solution of (3) - proof of Theorem 2.3

We construct a strong solution of (3) using a Picard iteration. Let $\beta^i, i \in \mathbb{N}^*$, be independent one-dimensional Brownian motions. Let $W, W^i,\ i \in \mathbb{N}^*$, be independent white noises on respectively $\mathbb{R}_+ \times [0, 1] \times \mathbb{R}$ and $\mathbb{R}_+ \times [0, 1]$ with respective intensities $dt \cdot dp \cdot \nu_1(dv)$ and $dt \cdot dp$, independent of the β^i. We suppose that all these processes are defined on the same probability space $(\Omega, \mathcal{F}, P)$ carrying also i.i.d. random variables X_0^i, $i \in \mathbb{N}^*$, which are independent of the β^i, W, W^i. Define for all $t \geqslant 0$,

$$\mathcal{G}_t := \sigma\{\, \beta_v^i;\, W(]u, v] \times A \times B);\, W^i(]u, v] \times A);\, 0 < u < v \leqslant t;\, A \in \mathcal{B}(\mathbb{R});\, B \in \mathcal{B}([0, 1]);\, i \in \mathbb{N}^* \,\};$$

$$\mathcal{W}_t := \sigma\{\, W(]u,v] \times A \times B); u < v \leqslant t;\ A \in \mathcal{B}(\mathbb{R}), B \in \mathcal{B}([0,1])\,\};$$
$$\mathcal{W} := \sigma\{W_t;\ t \geqslant 0\}.$$

Step 1. Fix an $i \in \mathbb{N}^*$, and introduce

$$\begin{aligned}
X_t^{i,[0]} :=& X_0^i,\\
\mu_t^{[0]} :=& \mathcal{L}(X_0^i),\ F_t^{[0]}(x) := P\left(X_0 \leqslant x | \mathcal{W}\right),\\
M^{[0]}([0,t] \times A \times B) :=& \int_0^t \int_0^1 \int_{\mathbb{R}} \mathbb{1}_A((F_s^{[0]})^{-1}(p)) \mathbb{1}_B(v) dW(s,p,v)\\
M^{i,[0]}([0,t] \times A) :=& \int_0^t \int_0^1 \mathbb{1}_A((F_s^{[0]})^{-1}(p)) dW^i(s,p).
\end{aligned}$$

Assuming everything is defined at order $n \in \mathbb{N}$, we introduce

$$\begin{aligned}
X_t^{i,[n+1]} :=& \int_0^t b(X_s^{i,[n]}, \mu_s^{[n]}) ds + \int_0^t \sigma(X_s^{i,[n]}, \mu_s^{[n]}) d\beta_s^i \\
&+ \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} \sqrt{f(x, \mu_s^{[n]})} \tilde{\Psi}(x, X_s^{i,[n]}, \mu_s^{[n]}, v) dM^{[n]}(s,x,v)\\
&+ \int_0^t \int_{\mathbb{R}} \sqrt{f(x, \mu_s^{[n]})} \kappa(x, X_s^{i,[n]}, \mu_s^{[n]}) dM^{i,[n]}(s,x),\\
\mu_s^{[n+1]} :=& \mathcal{L}(X_s^{i,[n+1]} | \mathcal{W}),\ F_s^{[n+1]}(x) := \mathbb{P}(X_s^{i,[n+1]} \leqslant x | \mathcal{W}),\\
M^{[n+1]}([0,t] \times A \times B) :=& \int_0^t \int_0^1 \int_{\mathbb{R}} \mathbb{1}_A((F_s^{[n+1]})^{-1}(p)) \mathbb{1}_B(v) dW(s,p,v),\\
M^{i,[n+1]}([0,t] \times A) :=& \int_0^t \int_0^1 \mathbb{1}_A((F_s^{[n+1]})^{-1}(p)) dW^i(s,p).
\end{aligned} \tag{8}$$

Note that $\forall t > 0$,

$$\mathcal{W} = \sigma(\mathcal{W}_t;\ \mathcal{W}_{]t;\infty[})$$

where

$$\mathcal{W}_{]t;\infty[} := \sigma\{]W(]u,v] \times A);\ , t < u < v;\ A \in \mathcal{B}(\mathbb{R})\}.$$

Remember also that white noises are processes with independent increments, more precisely, for all A, A' in $\mathcal{B}([0,1])$, for all B, B' in $\mathcal{B}(\mathbb{R})$, $W(]u,v] \times A \times B)$ and $W(]u',v'] \times A' \times B')$ are independent if $]u;v] \cap]u';v'] = \varnothing$.

Using this last remark, we see that by construction $X_t^{i,[n+1]}$ is independent from $\mathcal{W}_{]t;\infty[}$, and as a consequence, $\mathbb{P}(X_s^{i,[n+1]} \leqslant x | \mathcal{W}) = \mathbb{P}(X_s^{i,[n+1]} \leqslant x | \mathcal{W}_s)$. Taking $\mathcal{F}_t = \mathcal{W}_t$ and $\mathcal{F}_\infty = \mathcal{W}$ we see that all assumptions of Lemma (3.1) are satisfied. Hence, for each $n \in \mathbb{N}$; $i \in \mathbb{N}^*$ the martingale measures $M^{[n]}$ and $M^{i,[n]}$ are well defined and have respectively the intensities $dt \cdot \mu_t^{[n]}(dx) \cdot \nu_1(dv)$ and $dt \cdot \mu_t^{[n]}(dx)$, where $\mu_t^{[n]} := \mathcal{L}(X_t^{i,[n]} | \mathcal{W}) = \mathcal{L}(X_t^{i,[n]} | \mathcal{W}_t)$.

In what follows, we shall consider $u_t^{[n]} := \mathbb{E}\left[\left(X_t^{i,[n+1]} - X_t^{i,[n]}\right)^2\right]$. Let us introduce

$$h(x,y,m,v) := \sqrt{f(x,m)} \tilde{\Psi}(x,y,m,v) \text{ and } g(x,y,m) := \sqrt{f(x,m)} \kappa(x,y,m).$$

Note that the assumptions of the theorem guarantee that h and g are Lipschitz continuous. Indeed, using Assumption 1 (ii) and (iii), for all $x, y, x', y', v \in \mathbb{R}, m, m' \in \mathcal{P}_1(\mathbb{R})$,

$$|h(x,y,m,v) - h(x',y',m',v)| \leqslant C(v)(|x-x'| + |y-y'| + W_1(m,m')),$$

where

$$C(v) := \int_E M(u,v)\nu_1(du) + \int_E |\Psi(x,y,m,v,u)|\nu_1(du).$$

Using Jensen's inequality together with Assumptions 1 (iii) and 2 (ii) we see that C satisfies $\int_{\mathbb{R}} C(v)^2 dv < \infty$. Moreover, using Assumption 1 (ii) together with Remark 2.1, for all $x, x', y, y' \in \mathbb{R}, m, m' \in \mathcal{P}_1(\mathbb{R})$,

$$|g(x,y,m) - g(x',y',m')| \leqslant K(|x-x'| + |y-y'| + W_1(m,m')),$$

where $K \leqslant C(\|\kappa\|_\infty + \sqrt{\|f\|_\infty})$.

Step 2. We now prove that our Picard scheme converges. Classical arguments imply the existence of a constant $C > 0$ such that

$$\begin{aligned}
&\frac{1}{C}\left(X_t^{i,[n+1]} - X_t^{i,[n]}\right)^2 \leqslant \\
&\left(\int_0^t (b(X_s^{i,[n]}, \mu_s^{[n]}) - b(X_s^{i,[n-1]}, \mu_s^{[n-1]}))ds\right)^2 + \left(\int_0^t (\sigma(X_s^{i,[n]}, \mu_s^{[n]}) - \sigma(X_s^{i,[n-1]}, \mu_s^{[n-1]}))d\beta_s^i\right)^2 \\
&+ \left(\int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} h(x, X_s^{i,[n]}, \mu_s^{[n]}, v) dM^{[n]}(s,x,v) - \int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} h(x, X_s^{i,[n-1]}, \mu_s^{[n-1]}, v) dM^{[n-1]}(s,x,v)\right)^2 \\
&\qquad + \left(\int_0^t \int_{\mathbb{R}} g(x, X_s^{i,[n]}, \mu_s^{[n]}) dM^{i,[n]}(s,x) - \int_0^t \int_{\mathbb{R}} g(x, X_s^{i,[n-1]}, \mu_s^{[n-1]}) dM^{i,[n-1]}(s,x)\right)^2 \\
&\leqslant t\int_0^t \left(b(X_s^{i,[n]}, \mu_s^{[n]}) - b(X_s^{i,[n-1]}, \mu_s^{[n-1]})\right)^2 ds + \left(\int_0^t (\sigma(X_s^{i,[n]}, \mu_s^{[n]}) - \sigma(X_s^{i,[n-1]}, \mu_s^{[n-1]}))d\beta_s^i\right)^2 \\
&+ \left(\int_0^t \int_0^1 \int_{\mathbb{R}} \left[h((F_s^{[n]})^{-1}(p), \bar{X}_s^{i,[n]}, \mu_s^{[n]}, v) - h((F_s^{[n-1]})^{-1}(p), \bar{X}_s^{i,[n-1]}, \mu_s^{[n-1]}, v)\right] dW(s,p,v)\right)^2 \\
&\qquad + \left(\int_0^t \int_0^1 \left[g((F_s^{[n]})^{-1}(p), \bar{X}_s^{i,[n]}, \mu_s^{[n]}) - g((F_s^{[n-1]})^{-1}(p), \bar{X}_s^{i,[n-1]}, \mu_s^{[n-1]})\right] dW^i(s,p)\right)^2. \quad (9)
\end{aligned}$$

Using Burkholder-Davis-Gundy's inequality to control the expectation of the stochastic integrals above, and using the fact that for all $\mu, \nu \in \mathcal{P}_2(\mathbb{R})$,

$$W_1(\mu,\nu) \leqslant W_2(\mu,\nu),$$

we have that

$$u_t^{[n]} \leqslant C(1+t)\int_0^t \mathbb{E}\left[(X_s^{i,[n]} - X_s^{i,[n-1]})^2\right] ds + C(1+t)\int_0^t \mathbb{E}\left[W_2(\mu_s^{[n]}, \mu_s^{[n-1]})^2\right] ds$$

$$+ C \int_0^t \mathbb{E}\left[\int_0^1 ((F_s^{[n]})^{-1}(p) - (F_s^{[n-1]})^{-1}(p))^2 dp\right] ds. \quad (10)$$

A classical result (see e.g. Theorem 8.1 of Major (1978)) states that, if F, G are two distribution functions with associated probability measure μ and ν, respectively, then

$$\int_0^1 (F^{-1}(p) - G^{-1}(p))^2 dp = \inf_{X\sim\mu, Y\sim\nu} \mathbb{E}\left[(X-Y)^2\right] = W_2(\mu,\nu),$$

where the infimum is taken over all possible couplings (X, Y) of μ and ν.

This implies that

$$\int_0^1 ((F_s^{[n]})^{-1}(p) - (F_s^{[n-1]})^{-1}(p))^2 dp = W_2(\mu_s^{[n]}, \mu_s^{[n-1]})^2.$$

Since $X^{i,[n]}$ and $X^{i,[n-1]}$, conditionally on $\mathcal{W}$, are respectively realizations of $F^{i,[n]}$ and $F^{i,[n-1]}$, we have that, for every $s \geqslant 0$, almost surely,

$$W_2(\mu_s^{[n]}, \mu_s^{[n-1]})^2 \leqslant \mathbb{E}\left[(X_s^{i,[n]} - X_s^{i,[n-1]})^2 | \mathcal{W}\right].$$

Integrating with respect to W implies that

$$\mathbb{E}\left[\int_0^1 ((F_s^{[n]})^{-1}(p) - (F_s^{[n-1]})^{-1}(p))^2 dp\right] \leqslant \mathbb{E}\left[(X_s^{i,[n]} - X_s^{i,[n-1]})^2\right].$$

Consequently, we have shown that there exists some constant $C > 0$ such that, for all $t \geqslant 0$,

$$u_t^{[n]} \leqslant C(1+t)\int_0^t u_s^{[n-1]} ds. \quad (11)$$

Classical computations then give

$$u_t^{[n]} \leqslant C^n (1+t)^n \frac{t^n}{n!}.$$

Now, introducing $v_t^{[n]} := 2^n u_t^{[n]}$, we have that

$$\sum_{n\geqslant 0} v_t^{[n]} < \infty.$$

Hence, using that for all $x \in \mathbb{R}, \varepsilon > 0$, $|x| \leqslant \max(\varepsilon, x^2/\varepsilon) \leqslant \varepsilon + x^2/\varepsilon$, and applying this with $\varepsilon = 1/2^n$ and $x = X_t^{i,[n+1]} - X_t^{i,[n]}$, we have

$$\sum_{n\geqslant 0} \mathbb{E}\left[|X_t^{i,[n+1]} - X_t^{i,[n]}|\right] \leqslant \sum_{n\geqslant 0} \frac{1}{2^n} + \sum_{n\geqslant 0} v_t^{[n]} < \infty.$$

As a consequence, we can define, almost surely,

$$\bar{X}_t^i := X_0^i + \sum_{n\geqslant 0} (X_t^{i,[n+1]} - X_t^{i,[n]}) < +\infty,$$

and we know that $\mathbb{E}\left[|\bar{X}_t^i - X_t^{i,[n]}|\right]$ vanishes as n goes to infinity, and that $X_t^{i,[n]}$ converges almost surely to $\bar{X}_t^i$.

Step 3. Let us prove that $\bar{X}^i$ has finite fourth moments. Let $w_t^{[n]} := \mathbb{E}\left[(X_t^{i,[n]})^4\right]$.

By equation (8), we have

$$\begin{aligned}\frac{1}{C}\left(X_t^{i,[n]}\right)^4 \leqslant & \left(\int_0^t b(X_s^{i,[n-1]}, \mu_s^{[n-1]})ds\right)^4 + \left(\int_0^t \sigma(X_s^{i,[n-1]}, \mu_s^{[n-1]})d\beta_s^i\right)^4 \\ & + \left(\int_0^t \int_{\mathbb{R}} \int_{\mathbb{R}} h(x, X_s^{i,[n-1]}, \mu_s^{[n-1]}, v)dM^{[n-1]}(s,x,v)\right)^4 \\ & + \left(\int_0^t \int_{\mathbb{R}} g(x, X_s^{i,[n-1]}, \mu_s^{[n-1]})dM^{i,[n-1]}(s,x)\right)^4 \\ & =: A_1 + A_2 + A_3 + A_4. \end{aligned} \tag{12}$$

First of all, let us note that our Lipschitz assumptions allow to consider the following control: for any $x \in \mathbb{R}, m \in \mathcal{P}_1(\mathbb{R})$,

$$|b(x,m)| \leqslant |b(x,m) - b(0,\delta_0)| + |b(0,\delta_0)| \leqslant C(1 + |x| + W_1(\mu, \delta_0)) = C\left(1 + |x| + \int_{\mathbb{R}} |y| dm(y)\right),$$

and similar controls for the functions σ, h, g. Using this control and Jensen's inequality, we have

$$\begin{aligned}\mathbb{E}\left[A_1\right] \leqslant & t^3 \int_0^t \mathbb{E}\left[b(X_s^{i,[n-1]}, \mu_s^{[n-1]})^4\right] ds \\ \leqslant & Ct^3 \int_0^t \left(1 + w_s^{[n-1]} + \mathbb{E}\left[\left(\int_{\mathbb{R}} |y| \mu_s^{[n-1]}(dy)\right)^4\right]\right) ds \\ \leqslant & Ct^3 \int_0^t \left(1 + w_s^{[n-1]}\right) ds. \end{aligned}$$

We can obtain a similar control for the expressions A_2, A_3 and A_4 using Burkholder-Davis-Gundy's inequality noticing that the stochastic integrals involved are local martingales. We just give the details for A_2.

$$\begin{aligned}\mathbb{E}\left[A_2\right] \leqslant & \mathbb{E}\left[\left(\int_0^t \sigma(X_s^{i,[n-1]}, \mu_s^{[n-1]})^2 ds\right)^2\right] \\ \leqslant & \mathbb{E}\left[t \int_0^t \sigma(X_s^{i,[n-1]}, \mu_s^{[n-1]})^4 ds\right] \leqslant Ct \int_0^t \left(1 + w_s^{[n-1]}\right) ds, \end{aligned}$$

where the last inequality can be obtained with the same reasoning as the one used to control $\mathbb{E}\left[A_1\right]$. With the same reasoning, we have the following controls for A_3 and A_4 :

$$\mathbb{E}\left[A_3\right] + \mathbb{E}\left[A_4\right] \leqslant Ct \int_0^t \left(1 + w_s^{[n-1]}\right) ds.$$

Using the previous control in the inequality (12), we have that for all $t \geqslant 0$,

$$w_t^{[n+1]} \leqslant C(1+t^3) + C(1+t^3)\int_0^t w_s^{[n]} ds,$$

whence

$$w_t^{[n]} \leqslant \sum_{k=1}^{n} \frac{t^{k-1}}{(k-1)!} C^k (1+t^3)^k \leqslant C(1+t^3)e^{Ct(1+t^3)}.$$

Consequently

$$\sup_{n\in\mathbb{N}} \sup_{0\leqslant s\leqslant t} \mathbb{E}\left[(X_s^{i,[n]})^4\right] < \infty, \tag{13}$$

for some constant $C > 0$. Then Fatou's lemma implies the result: for all $t \geqslant 0$,

$$\sup_{0\leqslant s\leqslant t} \mathbb{E}\left[(\bar{X}_s^i)^4\right] < \infty. \tag{14}$$

Step 4. Finally, we conclude the proof showing that $\bar{X}^i$ is solution to the limit equation. Roughly speaking, this step consists in letting n tend to infinity in (8).

We want to prove that, for all $t \geqslant 0$,

$$\bar{X}_t^i = G_t^i(\bar{X}^i, \mu), \tag{15}$$

where

$$\begin{aligned} G_t^i(\bar{X}^i, \mu) := &\int_0^t b(\bar{X}_s^i, \mu_s)ds + \int_0^t \sigma(\bar{X}_s^i, \mu_s)d\beta_s^i \\ &+ \int_0^t \int_0^1 \int_{\mathbb{R}} h(F_s^{-1}(p), \bar{X}_s^i, \mu_s, v)dW(s,p,v) \\ &+ \int_0^t \int_0^1 g(F_s^{-1}(p), \bar{X}_s^i, \mu_s)dW^i(s,p), \end{aligned}$$

where the functions h and g have been introduced in *Step 1*, $\mu_t := \mathcal{L}(\bar{X}_t^i|\mathcal{W})$, and F_t^{-1} is the generalized inverse of

$$F_t(x) := \mathbb{P}\left(\bar{X}_t^i \leqslant x \middle| \mathcal{W}\right).$$

Let us note that G_t^i has to be understood as a notation, we do not use its functional properties.

By construction, we have

$$X_t^{i,[n+1]} = G_t^i(X^{i,[n]}, \mu^{[n]}). \tag{16}$$

We have proved in *Step 2* that $X_t^{[n+1]}$ converges to $\bar{X}_t^i$ in L^1. In other words, the LHS of (16) converges to the LHS of (15) in L^1. Now, it is sufficient to prove that the RHS converges in L^2. This will prove that the equation (15) holds true.

With the same computations as the ones used to obtain (11) (recalling that this inequality relies on (9) and (10)), we have

$$\mathbb{E}\left[\left(G_t^i(\bar{X}^i, \mu) - G_t^i(X^{i,[n]}, \mu^{[n]})\right)^2\right] \leqslant C(1+t)\int_0^t \mathbb{E}\left[\left(\bar{X}_s^i - X_s^{i,[n]}\right)^2\right] ds.$$

This proves that $G_t^i(X^{[n],i}, \mu^{[n]})$ converges to $G_t^i(\bar{X}^i, \mu)$ in L^2 by dominated convergence: indeed, we know that for all $s \leqslant t$, $X_s^{[n],i}$ converges to $\bar{X}_s^i$ almost surely thanks to *Step 2.*, and (13) and (14) give the uniform integrability.

3.2. Trajectorial uniqueness

We continue the proof of Theorem 2.3 by proving the uniqueness of the solution. For that sake, let $\hat{X}^i$ and $\check{X}^i$ be two strong solutions defined with respect to the same initial condition X_0^i and the same white noises W and W^i. Let

$$u_t := \mathbb{E}\left[(\hat{X}_t^i - \check{X}_t^i)^2\right].$$

According to the computation of the previous subsection, there exists a constant $C > 0$ such that for all $t \geqslant 0$,

$$u_t \leqslant C(1+t)\int_0^t u_s ds.$$

Then Grönwall's lemma implies that $u_t = 0$ for all $t \geqslant 0$, implying the uniqueness.

4. Proof of Theorems 2.4 and 2.5

The proof of Theorem 2.5 follows the same steps as the proof of Theorem 1.7 of Erny, Löcherbach and Loukianova (2020): in a first time, we prove that the sequence $(\mu^N)_N$ is tight on $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$, and then we prove that each converging subsequence of $(\mu^N)_N$ converges to the same limit, by proving that the limits of such subsequences are solutions to some martingale problem which is well-posed.

4.1. Tightness of $(\mu^N)_N$

The proof that the sequence $(\mu^N)_N$ is tight on $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$, is almost the same as that of Proposition 2.1 in Erny, Löcherbach and Loukianova (2020). As the systems $(X^{N,i})_{1\leqslant i\leqslant N}$ $(N \in \mathbb{N}^*)$ are exchangeable, it is equivalent to the tightness of the sequence $(X^{N,1})_N$ on $D(\mathbb{R}_+,\mathbb{R})$ (see Proposition 2.2-(ii) of Sznitman (1989)).

The tightness of $(X^{N,1})_N$ is straightforward using Aldous' criterion (see Theorem 4.5 of Jacod and Shiryaev (2003)), observing that, under our conditions, $\sup_N \mathbb{E}\left[\sup_{s\leqslant t}|X_s^{N,1}|\right] < \infty$ (see Lemma 6.1).

4.2. Martingale problem

To identify the structure of any possible limit of the sequence $(\mu^N)_N$, we introduce a convenient martingale problem. We have already used such a kind of martingale problem in a similar context in Section 2.2 of Erny, Löcherbach and Loukianova (2020). Since our limit system is necessarily an infinite exchangeable system, the martingale problem is constructed such that it reflects the correlations between the particles. It is therefore stated in terms of couples of particles.

Consider a probability measure $Q \in \mathcal{P}(\mathcal{P}(D(\mathbb{R}_+,\mathbb{R})))$. In what follows the role of Q will be to be the law of any possible limit μ of μ^N. Our martingale problem is stated on the canonical space

$$\Omega = \mathcal{P}(D(\mathbb{R}_+,\mathbb{R})) \times D(\mathbb{R}_+,\mathbb{R})^2.$$

We endow Ω with the product of the associated Borel sigma-fields and with the probability measure defined for all $A \in \mathcal{B}(\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))),\ B \in \mathcal{B}(D(\mathbb{R}_+,\mathbb{R})^2)$ by

$$P_Q(A\times B) := \int_{\mathcal{P}_1(D(\mathbb{R}_+,\mathbb{R}))} \mathbb{1}_A(m) m\otimes m(B) Q(dm). \tag{17}$$

We write an atomic event $\omega \in \Omega$ as $\omega = (m,y)$ with $y = (y_t)_{t\geqslant 0} = (y_t^1, y_t^2)_{t\geqslant 0}$. We write μ and $Y = (Y^1, Y^2)$ for the random variables $\mu(\omega) = m$ and $Y(\omega) = y$. The definition (17) implies that Q is the distribution of μ, and that conditionally on μ, Y^1 and Y^2 are i.i.d. with distribution μ. More precisely, $\mu\otimes\mu$ is a regular conditional distribution of (Y^1,Y^2) given μ.

For $t \geqslant 0$ we write m_t for the t-th marginal of m : $m_t(C) = m(Y \in D^2(\mathbb{R}_+,\mathbb{R});\ y_t \in C)$; $\quad C \in \mathcal{B}(\mathbb{R})$. We denote μ_t the r.v. $\mu_t(\omega) = m_t$ and consider the filtration $(\mathcal{G}_t)_{t\geqslant 0}$ given by

$$\mathcal{G}_t = \sigma(Y_s, s\leqslant t) \vee \sigma(\mu_s(A);\ A\in\mathcal{B}(\mathbb{R}),\ s\leqslant t).$$

For all $g \in C_b^2(\mathbb{R}^2)$, define

$$\begin{aligned} Lg(y,m,x,v) :=& b(y^1,m)\partial_{y^1} g(y) + b(y^2,m)\partial_{y^2} g(y) \\ &+ \frac{1}{2}\sigma(y^1,m)^2\partial^2_{y^1} g(y) + \frac{1}{2}\sigma(y^2,m)^2\partial^2_{y^2} g(y) \\ &+ \frac{1}{2} f(x,m)\kappa(x,y^1,m)^2\partial^2_{y^1} g(y) + \frac{1}{2} f(x,m)\kappa(x,y^2,m)^2\partial^2_{y^2} g(y) \\ &+ \frac{1}{2} f(x,m)\sum_{i,j=1}^{2} \tilde\Psi(x,y^i,m,v)\tilde\Psi(x,y^j,m,v)\partial^2_{y^i y^j} g(y), \end{aligned} \tag{18}$$

and put for all $t \geqslant 0$

$$M_t^g = M_t^g(\mu, Y) := g(Y_t) - g(Y_0) - \int_0^t \int_{\mathbb{R}}\int_{\mathbb{R}} Lg(Y_s,\mu_s,x,v)\nu_1(dv)\mu_s(dx)ds. \tag{19}$$

Definition 4.1. *Q is solution to the martingale problem $(\mathcal{M})$ if*

- *$Q-$almost surely, $\mu_0 = \nu_0$,*
- *for all $g \in C_b^2(\mathbb{R}^2)$, $(M_t^g)_{t\geqslant 0}$ is a $(P_Q, (\mathcal{G}_t)_t)-$martingale.*

Let us state a first result that allows us to partially recover our limit equation (3) from the martingale problem $(\mathcal{M})$. It is the equivalent of Lemma 2.4 of Erny, Löcherbach and Loukianova (2020) in the framework of white noises.

Lemma 4.2. *Grant Assumptions 1, 2 and 3. Let Q be a solution of $(\mathcal{M})$. Using the notation above, there exist on an extension of $(\Omega,\mathcal{F},(\mathcal{G}_t)_{t\geqslant 0}, P_Q)$ two Brownian motions β^1,β^2 and three white noises W^1,W^2,W; defined on $(\mathbb{R}_+\times[0,1],\ \mathcal{B}(\mathbb{R}_+\times[0,1]),\ dt\otimes dp)$ for W^i, $i=1,2$; and on $(\mathbb{R}_+\times[0,1]\times\mathbb{R},\ \mathcal{B}(\mathbb{R}_+\times[0,1])\times\mathbb{R}),\ dt\otimes dp\otimes\nu_1(dv))$ for W, such that $\beta^1,\beta^2,W^1,W^2,W$ are all independent and such that (Y_t) admits the representation*

$$\begin{aligned} dY_t^1 =& b(Y_t^1,\mu_t)dt + \sigma(Y_t^1,\mu_t)d\beta_t^1 \\ &+ \int_0^1\int_{\mathbb{R}} \sqrt{f(F_s^{-1}(p),\mu_t)}\tilde\Psi(F_s^{-1}(p),Y_t^1,\mu_t,v)dW(t,p,v) \end{aligned}$$

$$
\begin{aligned}
&+ \int_0^1 \sqrt{f(F_s^{-1}(p),\mu_t)}\kappa(F_s^{-1}(p),Y_t^1,\mu_t)dW^1(t,p),\\
dY_t^2 =& b(Y_t^2,\mu_t)dt + \sigma(Y_t^2,\mu_t)d\beta_t^2\\
&+ \int_0^1\int_{\mathbb{R}} \sqrt{f(F_s^{-1}(p),\mu_t)}\tilde{\Psi}(F_s^{-1}(p),Y_t^2,\mu_t,v)dW(t,p,v)\\
&+ \int_{\mathbb{R}} \sqrt{f(F_s^{-1}(p),\mu_t)}\kappa(F_s^{-1}(p),Y_t^2,\mu_t)dW^2(t,p),
\end{aligned}
$$

where F_s is the distribution function related to μ_s, and F_s^{-1} its generalized inverse.

Proof. Theorem II.2.42 of Jacod and Shiryaev (2003) implies that Y is a continuous semimartingale with characteristics (B,C) given by

$$
\begin{aligned}
B_t^i =& \int_0^t b(Y_s^i,\mu_s)ds, \quad 1\leqslant i\leqslant 2,\\
C_t^{i,i} =& \int_0^t \sigma(Y_s^i,\mu_s)^2 ds + \int_0^t\int_{\mathbb{R}} f(x,\mu_s)\kappa(x,Y_s^i,\mu_s)^2\mu_s(dx)ds\\
&+ \int_0^t\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,\mu_s)\tilde{\Psi}(x,Y_s^i,\mu_s,v)^2\nu_1(dv)\mu_s(dx)ds, \quad 1\leqslant i\leqslant 2,\\
C_t^{1,2} =& \int_0^t\int_{\mathbb{R}}\int_{\mathbb{R}} f(x,\mu_s)\tilde{\Psi}(x,Y_s^1,\mu_s,v)\tilde{\Psi}(x,Y_s^2,\mu_s,v)\nu_1(dv)\mu_s(dx)ds.
\end{aligned}
$$

As we are interested in finding five white noises, we need to have five local martingales. This is why we introduce artificially $Y_t^i := 0$ for $3\leqslant i\leqslant 5$. The rest of the proof is then an immediate consequence of Theorem III-10 and Theorem III-6 of El Karoui and Méléard (1990). More precisely, Theorem III-10 implies the existence of five orthogonal martingale measures M^i $(1\leqslant i\leqslant 5)$ on $\mathbb{R}_+\times\mathbb{R}\times\mathbb{R}$ with intensity $dt\cdot\mu_t(dx)\cdot\nu_1(dv)$ such that

$$
\begin{aligned}
dY_t^1 =& b(Y_t^1,\mu_t)dt + \int_{\mathbb{R}}\int_{\mathbb{R}} \sigma(Y_t^1,\mu_t)dM^1(t,x,v)\\
&+ \int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{f(x,\mu_t)}\tilde{\Psi}(x,Y_t^1,\mu_t,v)dM^5(t,x,v)\\
&+ \int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{f(x,\mu_t)}\kappa(x,Y_t^1,\mu_t)dM^3(t,p,v),\\
dY_t^2 =& b(Y_t^2,\mu_t)dt + \int_{\mathbb{R}}\int_{\mathbb{R}} \sigma(Y_t^2,\mu_t)dM^2(t,x,v)\\
&+ \int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{f(x,\mu_t)}\tilde{\Psi}(x,Y_t^2,\mu_t,v)dM^5(t,x,v)\\
&+ \int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{f(x,\mu_t)}\kappa(x,Y_t^2,\mu_t)dM^4(t,x,v).
\end{aligned}
$$

Then, $\beta_t^i := M^i([0,t]\times\mathbb{R}\times\mathbb{R})$ $(i=1,2)$ are standard one-dimensional Brownian motions, and Theorem III-6 of El Karoui and Méléard (1990) allows us to write M^i $(3\leqslant i\leqslant 5)$ as

$$
M^5([0,t]\times A\times B) = \int_0^t \mathbb{1}_A(F_s^{-1}(p))\mathbb{1}_B(v)dW(s,p,v)
$$

$$M^3([0,t]\times A\times B)=\int_0^t \mathbb{1}_A(F_s^{-1}(p))dW^1(s,p)$$
$$M^4([0,t]\times A\times B)=\int_0^t \mathbb{1}_A(F_s^{-1}(p))dW^2(s,p),$$

where W^1, W^2, W are white noises with respective intensities $dt\cdot dp$, $dt\cdot dp$ and $dt\cdot dp\cdot \nu_1(dv)$. □

We now prove a key result for the proof of our main results. Recall that the martingale problem $(\mathcal{M})$ is given by definition (4.1).

Theorem 4.3. *Grant Assumptions 1, 2 and 3. Then the law of every limit in distribution of the sequence (μ^N) is solution of the martingale problem $(\mathcal{M})$.*

Proof. Let μ be the limit in distribution of some subsequence of (μ^N) and let $Q=Q_\mu$ be its law. In the following, we still note this subsequence (μ^N). Firstly, we clearly have that $\mu^0=\nu^0$. For $0\leqslant s_1\leqslant ...\leqslant s_k\leqslant s$; $\psi_1,...,\psi_k\in C_b(\mathcal{P}(\mathbb{R}))$; $\varphi_1,...,\varphi_k\in C_b(\mathbb{R}^2)$; $\varphi\in C_b^3(\mathbb{R}^2)$ define the following functional on $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$:

$$F(\mu):=\psi_1(\mu_{s_1})...\psi_k(\mu_{s_k})\int_{D(\mathbb{R}_+,\mathbb{R})^2}\mu\otimes\mu(d\gamma)\varphi_1(\gamma_{s_1})\dots\varphi_k(\gamma_{s_k})\left[M_t^\varphi(\mu,\gamma)-M_s^\varphi(\mu,\gamma)\right],$$

where $(M_t^\varphi(\mu,\gamma))_t$ is given by (19). To show that $(M_t^\varphi(\mu,\gamma))_{t\geqslant 0}$ is a $(P_{Q_\mu},(\mathcal{G}_t)_t)$ martingale, we have to show that

$$\int_{\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))}F(m)Q_\mu(dm)=\mathbb{E}\left[F(\mu)\right]=0,$$

where the expectation $\mathbb{E}\left[\cdot\right]$ is taken with respect to $P_{Q_\mu}=P_Q$. Note that the first equality is just the transfer formula, and hence the expectation is taken on the probability space where μ is defined.

Step 1. We show that

$$\mathbb{E}\left[F(\mu^N)\right]\to\mathbb{E}\left[F(\mu)\right] \tag{20}$$

as $N\to\infty$. This statement is not immediately clear since the functional F is not continuous. The main difficulty comes from the fact that μ^N converges to μ in distribution for the topology of the weak convergence, but the terms appearing in the function F require the convergence of μ_t^N to μ_t for the topology of the metric W_1.

The proof (20) is actually quite technical, therefore we postpone it to the Appendix in Lemma 6.2.

Step 2. In this step we show that $\mathbb{E}\left[F(\mu)\right]$ is equal to 0. Applying F to μ^N gives

$$\begin{aligned}F(\mu^N)=\psi_1(\mu^N_{s_1})...\psi_k(\mu^N_{s_k})\frac{1}{N^2}\sum_{i,j=1}^N\varphi_1(X^{N,i}_{s_1},X^{N,j}_{s_1})...\varphi_k(X^{N,i}_{s_k},X^{N,j}_{s_k})\\ \Big[\varphi(X^{N,i}_t,X^{N,j}_t)-\varphi(X^{N,i}_s,X^{N,j}_s)\\ -\int_s^t b(X^{N,i}_r,\mu^N_r)\partial_{x^1}\varphi(X^{N,i}_r,X^{N,j}_r)dr-\int_s^t b(X^{N,j}_r,\mu^N_r)\partial_{x^2}\varphi(X^{N,i}_r,X^{N,j}_r)dr\\ -\frac{1}{2}\int_s^t\sigma(X^{N,i}_r,\mu^N_r)^2\partial^2_{x^1}\varphi(X^{N,i}_r,X^{N,j}_r)dr-\frac{1}{2}\int_s^t\sigma(X^{N,j}_r,\mu^N_r)^2\partial^2_{x^2}\varphi(X^{N,i}_r,X^{N,j}_r)dr\end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{2}\int_s^t \frac{1}{N}\sum_{k=1}^N f(X_r^{N,k},\mu_r^N)\kappa(X_r^{N,k},X_r^{N,i},\mu_r^N)^2\partial^2_{x^1}\varphi(X_r^{N,i},X_r^{N,j})dr \\
&- \frac{1}{2}\int_s^t \frac{1}{N}\sum_{k=1}^N f(X_r^{N,k},\mu_r^N)\kappa(X_r^{N,k},X_r^{N,j},\mu_r^N)^2\partial^2_{x^2}\varphi(X_r^{N,i},X_r^{N,j})dr \\
&- \frac{1}{2}\int_s^t\int_{\mathbb{R}}\sum_{k=1}^N f(X_r^{N,k},\mu_r^N)\frac{1}{N}\sum_{h,l=1}^2 \tilde{\Psi}(X_r^{N,k},X_r^{N,i_h},\mu_r^N,v)\tilde{\Psi}(X_r^{N,k},X_r^{N,i_l},\mu_r^N,v)\cdot \\
&\qquad\qquad \partial^2_{x^h x^l}\varphi(X_r^{N,i},X_r^{N,j})\nu_1(dv)dr\Big], \quad (21)
\end{aligned}
$$

with $i_1 = i$ and $i_2 = j$. Denote $\tilde{\pi}^k(dr,dz,du) := \pi(dr,dz,du) - dr\cdot dz\cdot\nu(du)$ the compensated version of π^k. For $(i,j) \in [\![1,\dots,N]\!]^2,\ s<t$ let us define

$$
M_{s,t}^{N,i,j} := \int_s^t \sigma(X_r^{N,i},\mu_r^N)\partial_{x^1}\varphi(X_r^{N,i},X_r^{N,j})d\beta_r^i + \int_s^t \sigma(X_r^{N,j},\mu_r^N)\partial_{x^2}\varphi(X_r^{N,i},X_r^{N,j})d\beta_r^j; \quad (22)
$$

$$
\begin{aligned}
W_{s,t}^{N,i,j} := &\sum_{k=1,k\neq i,j}^N \int_{]s,t]\times\mathbb{R}_+\times E} \mathbb{1}_{\{z\leqslant f(X_{r-}^{N,k},\mu_{r-}^N)\}} \\
&\Big[\varphi(X_{r-}^{N,i}+\frac{1}{\sqrt{N}}\Psi(X_{r-}^{N,k},X_{r-}^{N,i},\mu_{r-}^N,u^k,u^i),\ X_{r-}^{N,j}+\frac{1}{\sqrt{N}}\Psi(X_{r-}^{N,k},X_{r-}^{N,j},\mu_{r-}^N,u^k,u^j)) \\
&\qquad -\varphi(X_{r-}^{N,i},X_{r-}^{N,j})\Big]\tilde{\pi}^k(dr,dz,du); \quad (23)
\end{aligned}
$$

$$
\begin{aligned}
\Delta_{s,t}^{N,i,j} := &\sum_{k=1,k\neq i,j}^N \int_s^t\int_E f(X_r^{N,k},\mu_{r-}^N) \\
&\Big[\varphi(X_{r-}^{N,i}+\frac{1}{\sqrt{N}}\Psi(X_{r-}^{N,k},X_{r-}^{N,i},\mu_{r-}^N,u^k,u^i),\ X_{r-}^{N,j}+\frac{1}{\sqrt{N}}\Psi(X_{r-}^{N,k},X_{r-}^{N,j},\mu_{r-}^N,u^k,u^j)) \\
&\qquad -\varphi(X_{r-}^{N,i},X_{r-}^{N,j})\Big]\nu(du)dr \quad (24)
\end{aligned}
$$

and

$$
\begin{aligned}
\Gamma_{s,t}^{N,i,j} &= \Delta_{s,t}^{N,i,j} \\
&- \sum_{l=1}^2\sum_{k=1,k\neq i,j}^N \int_s^t\int_E f(X_r^{N,k},\mu_r^N)\frac{1}{\sqrt{N}}\Psi(X_{r-}^{N,k},X_{r-}^{N,i_l},\mu_{r-}^N,u^k,u^{i_l})\partial_{x^l}\varphi(X_r^{N,i},X_r^{N,j})\nu(du)dr \\
&- \sum_{h,l=1}^2 \int_s^t\int_E \frac{1}{N}\sum_{k=1,k\neq i,j}^N f(X_r^{N,k},\mu_r^N)\Psi(X_r^{N,k},X_r^{N,i_h},\mu_r^N,u^k,u^{i_h})\Psi(X_r^{N,k},X_r^{N,i_l},\mu_r^N,u^k,u^{i_l}) \\
&\qquad \partial^2_{x^h x^l}\varphi(X_r^{N,i},X_r^{N,j})\nu(du)dr, \quad (25)
\end{aligned}
$$

with again $i_1 = i$ and $i_2 = j$.

Applying Ito's formula, we have

$$\varphi(X_t^{N,i}, X_t^{N,j}) = \varphi(X_s^{N,i}, X_s^{N,j}) + M_{s,t}^{N,i,j} + W_{s,t}^{N,i,j} + \Delta_{s,t}^{N,i,j} + \int_s^t b(X_r^{N,i}, \mu_r^N)\partial_{x^1}\varphi(X_r^{N,i}, X_r^{N,j})dr + \int_s^t b(X_r^{N,j}, \mu_r^N)\partial_{x^2}\varphi(X_r^{N,i}, X_r^{N,j})dr + \frac{1}{2}\int_s^t \sigma(X_r^{N,i}, \mu_r^N)^2\partial^2_{x^1}\varphi(X_r^{N,i}, X_r^{N,j})dr + \frac{1}{2}\int_s^t \sigma(X_r^{N,j}, \mu_r^N)^2\partial^2_{x^2}\varphi(X_r^{N,i}, X_r^{N,j})dr. \quad (26)$$

Note that thanks to Assumption (3) i), the term in the second line of (25) is zero. Again, if $i \neq j$, using the definitions (5) and (6)

$$\int_E \Psi(X_r^{N,k}, X_r^{N,i}, \mu_r^N, u^k, u^i)\Psi(X_r^{N,k}, X_r^{N,j}, \mu_r^N, u^k, u^j)\nu(du) = \int_{\mathbb{R}} \tilde{\Psi}(X_r^{N,k}, X_r^{N,i}, \mu_r^N, v)\tilde{\Psi}(X^{N,r}, X^{N,j}, \mu_r^N, v)\nu_1(dv),$$

and

$$\int_E \Psi(X_r^{N,k}, X_r^{N,i}, \mu_r^N, u^k, u^i)^2\nu(du) = \kappa(X_r^{N,k}, X_r^{N,i}, \mu_r^N)^2 + \int_{\mathbb{R}} \tilde{\Psi}(X_r^{N,k}, X_r^{N,i}, \mu_r^N, v)^2\nu_1(dv).$$

This implies that the last line of (25) is equal to the sum of three last lines of (21), up to an error term which is of order $O(t/N)$ if we include the terms $k = i, j$ in (25).

As a consequence, plugging (26) in (21), using the definitions (23), (22), (25) and the previous remark we obtain

$$F(\mu^N) = \psi_1(\mu_{s_1}^N)...\psi_k(\mu_{s_k}^N)\frac{1}{N^2}\sum_{i,j=1}^N \varphi_1(X_{s_1}^{N,i}, X_{s_1}^{N,j})...\varphi_k(X_{s_k}^{N,i}, X_{s_k}^{N,j})\left[M_{s,t}^{N,i,j} + W_{s,t}^{N,i,j} + \Gamma_{s,t}^{N,i,j}\right].$$

Using (26) we see that $(M_{s,t}^{N,i,j} + W_{s,t}^{N,i,j})$, $t \geqslant s$, is a martingale with respect to the filtration $(\mathcal{F}_t^{X^N})_{t\geqslant 0}$ with $\mathcal{F}_t^{X^N} := \sigma(X_u^{N,i}; X_u^{N,j};\ s \leqslant u \leqslant t)$ on the space where X^N is defined. Hence, using that φ_{s_k} and ψ_{s_k} are bounded, and that μ^N is $(\mathcal{F}_t^{X^N})_{t\geqslant 0}$ adapted,

$$\mathbb{E}\left[F(\mu^N)\right] = \mathbb{E}\left[\mathbb{E}\left[F(\mu^N)|\mathcal{F}_s^{X^N}\right]\right] = \mathbb{E}\left[\psi_1(\mu_{s_1}^N)...\psi_k(\mu_{s_k}^N)\frac{1}{N^2}\sum_{i,j=1}^N \varphi_1(X_{s_1}^{N,i}, X_{s_1}^{N,j})...\varphi_k(X_{s_k}^{N,i}, X_{s_k}^{N,j})\mathbb{E}\left[\Gamma_{s,t}^{N,i,j}|\mathcal{F}_s^{X,\mu}\right]\right] \leqslant \frac{C}{N^2}\sum_{i,j=1}^N \mathbb{E}\left[|\Gamma_{s,t}^{N,i,j}|\right], \quad (27)$$

implying that

$$|\mathbb{E}\left[F(\mu^N)\right]| \leqslant C\mathbb{E}\left[|\Gamma_{s,t}^{N,1,2}| + \frac{|\Gamma_{s,t}^{N,1,1}|}{N}\right].$$

Taylor-Lagrange's inequality gives for all $i \neq j$,

$$\begin{aligned}
&\mathbb{E}\left[|\Gamma_{s,t}^{N,i,j}|\right] \leqslant \\
&C\frac{1}{N\sqrt{N}} \sum_{k=1,k\neq i,j}^{N} \sum_{n=0}^{3} \int_s^t \int_E \mathbb{E}\left[\Psi(X_r^{N,k}, X_r^{N,i}, \mu_r^N, u^k, u^i)^n \Psi(X_r^{N,k}, X_r^{N,j}, \mu_r^N, u^k, u^j)^{3-n}\right] \nu(du)dr \\
&\leqslant C\frac{1}{N\sqrt{N}} \sum_{k=1,k\neq i,j}^{N} \int_s^t \int_E \mathbb{E}\left[|\Psi(X_r^{N,k}, X_r^{N,i}, \mu_r^N, u^k, u^i)|^3 + |\Psi(X_r^{N,k}, X_r^{N,j}, \mu_r^N, u^k, u^j)|^3\right] \nu(du)dr \\
&\leqslant C\frac{1}{\sqrt{N}},
\end{aligned}$$

and a similar result holds for $\Gamma_{s,t}^{N,1,1}$. Consequently,

$$|\mathbb{E}\left[F(\mu^N)\right]| \leqslant CN^{-1/2},$$

implying together with (20) that

$$\mathbb{E}\left[F(\mu)\right] = \lim_N \mathbb{E}\left[F(\mu^N)\right] = 0.$$

□

4.3. *Proof of Theorem 2.5*

The beginning of the proof is similar to the proof of Theorem 2.6 of Erny, Löcherbach and Loukianova (2020). It mainly consists in applying Lemma 4.2 and Theorem 4.3.

Let μ be the limit in distribution of some converging subsequence of (μ^N) (that we still note (μ^N)). Then, by Proposition (7.20) of Aldous, μ is the directing measure of an exchangeable system $(\bar{Y}^i)_{i\geqslant 1}$, and $(X^{N,i})_{1\leqslant i\leqslant N}$ converges in distribution to $(\bar{Y}^i)_{i\geqslant 1}$.

According to Theorem 4.3 and Lemma 4.2 for every $i \neq j$, there exist, on an extension, Brownian motions $\beta^{i,j,1}, \beta^{i,j,2}$ and white noises $W^{i,j,1}, W^{i,j,2}, W^{i,j}$ with respective intensities $dt \cdot dp$, $dt \cdot dp$ and $dt \cdot dp \cdot \nu_1(dv)$ all independent such that

$$\begin{aligned}
d\bar{Y}_t^i =& b(\bar{Y}_t^i, \mu_t)dt + \sigma(\bar{Y}_t^i, \mu_t)d\beta_t^{i,j,1} \\
&+ \int_0^1 \int_{\mathbb{R}} \sqrt{f(F_s^{-1}(p), \mu_t)}\tilde{\Psi}(F_s^{-1}(p), \bar{Y}_t^i, \mu_t, v)dW^{i,j}(t,p,v) \\
&+ \int_0^1 \sqrt{f(F_s^{-1}(p), \mu_t)}\kappa(F_s^{-1}(p), \bar{Y}_t^i, \mu_t)dW^{i,j,1}(t,p), \\
d\bar{Y}_t^j =& b(\bar{Y}_t^j, \mu_t)dt + \sigma(\bar{Y}_t^j, \mu_t)d\beta_t^{i,j,2} \\
&+ \int_0^p \int_{\mathbb{R}} \sqrt{f(F_s^{-1}(p), \mu_t)}\tilde{\Psi}(F_s^{-1}(p), \bar{Y}_t^j, \mu_t, v)dW^{i,j}(t,p,v) \\
&+ \int_{\mathbb{R}} \sqrt{f(F_s^{-1}(p), \mu_t)}\kappa(F_s^{-1}(p), \bar{Y}_t^j, \mu_t)dW^{i,j,2}(t,p),
\end{aligned}$$

where F_s is the distribution function related to μ_s, and F_s^{-1} its generalized inverse. We can construct this extension in a global way such that it works for all couples (i,j) simultaneously.

As the system $(\bar{Y}^i)_{i\geqslant 1}$ is exchangeable, we know that $W^{i,j,1} = W^i$, $W^{i,j,2} =: W^j$, $W^{i,j} =: W$, $\beta^{i,j,1} =: \beta^i$ and $\beta^{i,j,2} =: \beta^j$.

It remains to identify the structure of μ_t appearing above as the one prescribed in (4) as conditional law. This is what we are going to do now. To do so, we introduce the following auxiliary system

$$\begin{aligned} dZ_t^{N,i} =& b(\tilde{Z}_t^{N,i}, \mu_t^{Z,N})dt + \sigma(Z_t^{N,i}, \mu_t^{Z,N})d\beta_t^i \\ &+ \int_0^1 \int_{\mathbb{R}} \sqrt{f((F_s^{Z,N})^{-1}(p), \mu_t^{Z,N})}\tilde{\Psi}((F_s^{Z,N})^{-1}(p), Z_t^{N,i}, \mu_t^{Z,N}, v)dW(t,p,v) \\ &+ \int_0^1 \sqrt{f((F_s^{Z,N})^{-1}(p), \mu_t^{Z,N})}\kappa((F_s^{Z,N})^{-1}(p), Z_t^{N,i}, \mu_t^{Z,N})dW^i(t,p), \end{aligned}$$

where $\mu^{Z,N} := N^{-1}\sum_{i=1}^N \delta_{Z^{N,i}}$, $F^{Z,N}$ is the distribution function related to $\mu^{Z,N}$. In addition, in the rest of the proof, let $\mu^X := \mathcal{L}(\bar{X}^1|\mathcal{W})$ and μ^Y be the directing measure of $(\bar{Y}^i)_{i\geqslant 1}$ that was denoted by μ so far in this proof.

With similar computations as in Theorem 2.3, one can prove that, for all $t \geqslant 0$,

$$\mathbb{E}\left[(Z_t^{N,i} - \bar{X}_t^i)^2\right] \leqslant C(1+t)\int_0^t \mathbb{E}\left[(Z_s^{N,i} - \bar{X}_s^i)^2\right]ds + C(1+t)\int_0^t \mathbb{E}\left[W_1(\mu_s^X, \mu_s^{Z,N})^2\right]ds.$$

Besides, introducing $\mu^{X,N} := N^{-1}\sum_{i=1}^N \delta_{\bar{X}^i}$, we have

$$\begin{aligned} \mathbb{E}\left[W_1(\mu_s^X, \mu_s^{Z,N})^2\right] \leqslant& 2\mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right] + 2\mathbb{E}\left[W_1(\mu_s^{X,N}, \mu_s^{Z,N})^2\right] \\ \leqslant& 2\mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right] + 2\mathbb{E}\left[|Z_s^{N,i} - \bar{X}_s^i|\right]^2 \\ \leqslant& 2\mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right] + 2\mathbb{E}\left[(Z_s^{N,i} - \bar{X}_s^i)^2\right]. \end{aligned}$$

As a consequence, we obtain the following inequality, for all $t \geqslant 0$,

$$\mathbb{E}\left[(Z_t^{N,i} - \bar{X}_t^i)^2\right] \leqslant C(1+t)\int_0^t \mathbb{E}\left[(Z_s^{N,i} - \bar{X}_s^i)^2\right]ds + C_t\int_0^t \mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right]ds,$$

with C_t a constant depending on t. Grönwall's lemma now implies that, for all $t \geqslant 0$,

$$\mathbb{E}\left[(Z_t^{N,i} - \bar{X}_t^i)^2\right] \leqslant C_t\int_0^t \mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right]ds. \tag{28}$$

Now, let us prove that the expression above vanishes by dominated convergence. Note that $\mu_s^{X,N}$ converges weakly to μ_s^X a.s. by Glivenko-Cantelli's theorem (applying the theorem conditionally on $\mathcal{W}$) recalling that the variables $\bar{X}_s^j$ are conditionally i.i.d. given $\mathcal{W}$. And $\mu_s^{X,N}(|x|)$ converges to $\mu_s^X(|x|)$ a.s. as a consequence of the strong law of large numbers applied conditionally on $\mathcal{W}$ (once again because of the conditional independence property). Hence the characterization (i) of Definition 6.8 and Theorem 6.9 of Villani (2008) implies that $W_1(\mu_s^{X,N}, \mu_s^X)$ vanishes a.s. as N goes to infinity for every $s \geqslant 0$.

We also have, by Jensen's inequality and the definition of W^1, for all $t \geqslant 0$,

$$\mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^4\right] \leqslant C\mathbb{E}\left[\mu_s^X(x^4)\right] + C\mathbb{E}\left[\mu_s^{X,N}(x^4)\right] = C\mathbb{E}\left[(\bar{X}_s^1)^4\right] + \mathbb{E}\left[\frac{1}{N}\sum_{j=1}^{N}(\bar{X}_s^j)^4\right] \leqslant C\mathbb{E}\left[(\bar{X}_s^1)^4\right].$$

Consequently

$$\sup_{0\leqslant s\leqslant t}\mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^4\right] \leqslant C\sup_{0\leqslant s\leqslant t}\mathbb{E}\left[(\bar{X}_s^1)^4\right] < \infty,$$

recalling (14).

So we have just proven that the sequence of variables $W_1(\mu_s^X, \mu_s^{X,N})^2$ vanishes a.s. and for every $s \geqslant 0$, and that this sequence is uniformly integrable on $\Omega \times [0,t]$ for any $t \geqslant 0$. This implies, by dominated convergence, that

$$\int_0^t \mathbb{E}\left[W_1(\mu_s^X, \mu_s^{X,N})^2\right] ds \underset{N\to\infty}{\longrightarrow} 0.$$

Recalling (28), we have that

$$\mathbb{E}\left[(Z_t^{N,i} - \bar{X}_t^i)^2\right] \underset{N\to\infty}{\longrightarrow} 0. \tag{29}$$

With the same reasoning, we can prove that for all $t \geqslant 0$,

$$\mathbb{E}\left[(Z_t^{N,i} - \bar{Y}_t^i)^2\right] \underset{N\to\infty}{\longrightarrow} 0. \tag{30}$$

Finally, using (29) and (30), we have that, for all $t \geqslant 0$,

$$\mathbb{E}\left[(\bar{Y}_t^i - \bar{X}_t^i)^2\right] = 0.$$

This proves that the systems $(\bar{X}^i)_{i\geqslant 1}$ and $(\bar{Y}^i)_{i\geqslant 1}$ are equal. Hence, recalling that $\mu =: \mu^Y$ is a limit of μ^N and also the directing measure of $(Y^i)_{i\geqslant 1}$, it is also the directing measure of $(\bar{X}^i)_{i\geqslant 1}$, which is $\mu^X := \mathcal{L}(\bar{X}^1|\mathcal{W})$.

4.4. *Proof of Theorem 2.4*

The proof of Theorem 2.4 is now a direct consequence of Theorem 2.5 and Proposition (7.20) of Aldous (1983).

5. Model of interacting populations

The aim of this section is to generalize the previous model, considering n populations instead of one. Within each population, the particles interact as in the previous model, and, in addition, there are interactions at the level of the populations. The chaoticity properties of this kind of model have been studied by Graham (2008). Two examples of such systems are given in Figure 1.

If we consider a number of N particles, we note, for each $1 \leqslant k \leqslant n$, N_k the number of particles of the $k-$th population. In particular $N = N_1 + ... + N_n$. We assume that, for all $1 \leqslant k \leqslant n$, N_k/N converges to some positive number, such that each population survives in the limit system.

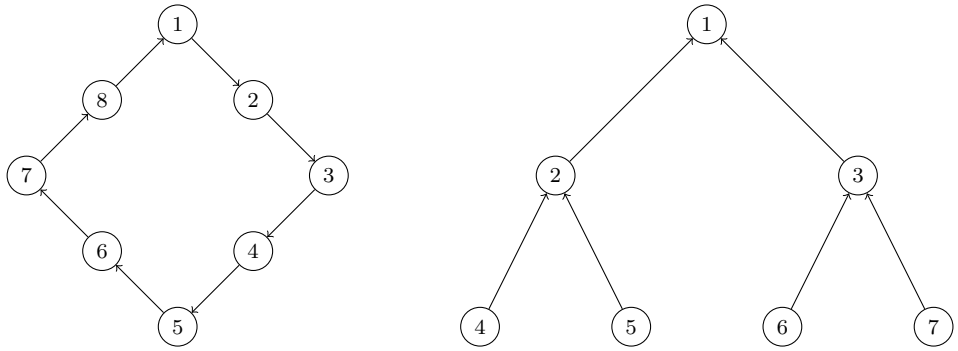

FIGURE 1. *Two examples of interacting populations.*

For all $1 \leqslant k \leqslant n$, let $I(k) \subseteq [\![1, n]\!]$ be the set of populations that are "inputs" of the population k, that is, such that particles within these populations have a direct influence on those in population k. The dynamic of the $N-$particle system $(X^{N,k,i})_{\substack{1\leqslant k\leqslant n \\ 1\leqslant i\leqslant N_k}}$ is governed by the following SDEs.

$$dX^{N,k,i} = b^k\left(X_t^{N,k,i}, \mu_t^{N,k}\right)dt + \sigma^k\left(X_t^{N,k,i}, \mu_t^{N,k}\right)d\beta_t^{k,i}$$
$$+ \sum_{l\in I(k)} \frac{1}{\sqrt{N_l}} \sum_{\substack{j=1 \\ (l,j)\neq(k,i)}}^{N_l} \int_{\mathbb{R}_+\times E^n} \Psi^{lk}(X_{t-}^{N,l,j}, X_{t-}^{N,k,i}, \mu_{t-}^{N,l}, \mu_{t-}^{N,k}, u^{l,j}, u^{k,i})\mathbb{1}_{\left\{z\leqslant f^l(X_{t-}^{N,l,j}, \mu_{t-}^{N,l})\right\}} d\pi^{l,j}(t,z,u).$$

In the above equation,

$$\mu_t^{N,k} = N_k^{-1}\sum_{j=1}^{N_k} \delta_{X_t^{N,k,j}},$$

$\pi^{l,j}$ $(1 \leqslant l \leqslant n, j \geqslant 1)$ are independent Poisson measures of intensity $dtdz\nu(du)$, where ν is a probability measure on $(\mathbb{R}^{\mathbb{N}^*})^n$ which is of the form

$$\nu = (\nu^{1,1})^{\otimes\mathbb{N}^*} \otimes (\nu^{2,1})^{\otimes\mathbb{N}^*} \otimes ... \otimes (\nu^{n,1})^{\otimes\mathbb{N}^*}.$$

The associated limit system $(\bar{X}^{k,i})_{\substack{1\leqslant k\leqslant n \\ i\geqslant 1}}$ is given by

$$\begin{aligned} d\bar{X}^{k,i} =& b^k\left(\bar{X}_t^{k,i}, \bar{\mu}_t^k\right)dt + \sigma^k\left(\bar{X}_t^{k,i}, \bar{\mu}_t^k\right)d\beta_t^{k,i} \\ &+ \sum_{l\in I(k)} \int_{\mathbb{R}}\int_{\mathbb{R}} \sqrt{f^l(x,\bar{\mu}_t^l)}\tilde{\Psi}^{lk}(x, \bar{X}_t^{k,i}, \mu_t^l, \mu_t^k, v)dM^l(t,x,v) \\ &+ \sum_{l\in I(k)} \int_{\mathbb{R}} \sqrt{f^l(x,\bar{\mu}_t^l)}\kappa^{lk}(x, \bar{X}_t^{k,i}, \mu_t^l, \mu_t^k)dM^{l,k,i}(t,x), \end{aligned}$$

with

$$\tilde{\Psi}^{lk}(x,y,m_1,m_2,v) := \int_{\mathbb{R}} \Psi^{lk}(x,y,m_1,m_2,v,w)d\nu^{k,1}(w),$$

$$\kappa^{lk}(x,y,m_1,m_2)^2 := \int_{E^n} \Psi^{lk}(x,y,m_1,m_2,u^{l,1},u^{k,2})^2 d\nu(u) - \int_{\mathbb{R}} \tilde{\Psi}^{lk}(x,y,m,v)^2 d\nu^{l,1}(v)$$
$$= \int_{E^n} \Psi^{lk}(x,y,m_1,m_2,u^{l,1},u^{k,2})^2 d\nu(u) - \int_{E^n} \Psi^{lk}(x,y,m,u^{l,1},u^{k,2})\Psi^{lk}(x,y,m,u^{l,1},u^{k,3}) d\nu(u),$$

and

$$M_t^{l,k,i}(A) = \int_0^t \mathbb{1}_A((F_s^l)^{-1}(p)) dW^{l,k,i}(s,p) \text{ and } M_t^l(A\times B) = \int_0^t \mathbb{1}_A((F_s^l)^{-1}(p))\mathbb{1}_B(v) dW^l(s,p,v).$$

In the above formulas, $\mu_t^k := \mathcal{L}\left(\bar{X}_t^{k,1} | \sigma\left(\bigcup_{l\in I(k)} \mathcal{W}^l\right)\right)$ and $(F_s^l)^{-1}$ is the generalized inverse of the function $F_s^l(x) := P(\bar{X}_s^{l,1} \leqslant x)$. Finally, $W^{l,k,i}$ and W^l $(1 \leqslant l \leqslant n, k \in I(l), i \geqslant 1)$ are independent white noises of respective intensities $dsdp$ and $dsdp\nu^{l,1}(dv)$, and

$$\mathcal{W}_t^l := \sigma\{\, W^l(]u,v] \times A \times B); u < v \leqslant t;\ A \in \mathcal{B}([0,1]), B \in \mathcal{B}(\mathbb{R})\,\};$$
$$\mathcal{W}^l := \sigma\{W_t^l;\ t \geqslant 0\}.$$

Both previous systems are "multi-exchangeable" in the sense that each population is "internally exchangeable" as in Corollary (3.9) of Aldous (1983). With the same reasoning as in the proof of Theorem 2.3, we can prove the existence of unique strong solutions $(X^{N,k,i})_{\substack{1\leqslant k\leqslant n\\ 1\leqslant i\leqslant N_k}}$ and $(\bar{X}^{k,i})_{\substack{1\leqslant k\leqslant n\\ i\geqslant 1}}$ as well as the convergence of the $N-$particle system to the limit system:

Theorem 5.1. *The following convergence in distribution in $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))^n$ holds true:*

$$\left(\mu^{N,1},\mu^{N,2},...,\mu^{N,n}\right) \longrightarrow \left(\bar{\mu}^1,\bar{\mu}^2,...,\bar{\mu}^n\right),$$

as $N\to\infty$.

Before giving a sketch of the proof of Theorem 5.1, we quickly state the following result.

Proposition 5.2. *Let $r \leqslant n$ and $1 \leqslant k_1 < ... < k_r \leqslant n$. If the sets $I(k_i)$ $(1 \leqslant i \leqslant r)$ are disjoint, then the random variables μ^{k_i} $(1 \leqslant i \leqslant r)$ are independent.*

Proof. For any $1 \leqslant k \leqslant n$, the system $(\bar{X}^{k,i})_{i\geqslant 1}$ is conditionally i.i.d. given $\sigma\left(\bigcup_{l\in I(k)} \mathcal{W}^l\right)$. So, by Lemma (2.12) of Aldous (1983), μ^k is $\sigma\left(\bigcup_{l\in I(k)} \mathcal{W}^l\right)-$measurable. □

Remark 5.3. *In the two examples of Figure 1, all the variables μ^k $(1 \leqslant k \leqslant n)$ are independent.*

Coming back to Theorem 5.1, its proof is similar to the proof of Theorem 2.5. The main argument relies on a generalization of the martingale problem discussed in Section 4.2. Let us formulate it. Consider

$$\Omega' := \mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))^n \times \left(D(\mathbb{R}_+,\mathbb{R})^2\right)^n,$$

and write any atomic event $\omega' \in \Omega'$ as

$$\omega' = \left(\mu^1,\mu^2,...,\mu^n,Y^{1,1},Y^{1,2},Y^{2,1},Y^{2,2},...,Y^{n,1},Y^{n,2}\right) = (\mu,Y),$$

$\mu = (\mu^1,\ldots,\mu^n), Y = (Y^1,\ldots,Y^n)$.

For $Q \in \mathcal{P}(\mathcal{P}_1(D(\mathbb{R}_+,\mathbb{R}))^n)$, consider the law P' on Ω' defined by

$$P'(A \times B_1 \times ... \times B_n) = \int_{\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))^n} \mathbb{1}_A(m) m^1 \otimes m^1(B_1)...m^n \otimes m^n(B_n) Q(dm),$$

with A a Borel set of $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))^n$ and $B_1, ..., B_n$ Borel sets of $D(\mathbb{R}_+,\mathbb{R})^2$.

Then, we say that Q is solution to our martingale problem if, for all $g \in C_b^2((\mathbb{R}^2)^n)$,

$$g(Y_t) - g(Y_0) - \int_0^t \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} Lg(Y_s, \mu_s, x, v) \mu_s^1 \otimes ... \otimes \mu_s^n(dx) \nu^{1,1} \otimes ... \otimes \nu^{n,1}(dv)$$

is a martingale, where

$$\begin{aligned} Lg(y,m,x,u) = &\sum_{k=1}^n \sum_{i=1}^2 b^k(y^{k,i}, m^k) \partial_{y^{k,i}} g(y) + \frac{1}{2} \sum_{k=1}^n \sum_{i=1}^2 \sigma^k(y^{k,i}, m^k)^2 \partial^2_{y^{k,i}} g(y) \\ &+ \frac{1}{2} \sum_{k=1}^n \sum_{l \in I(k)} \sum_{i=1}^2 f^l(x^l, m^l) \kappa^{lk}(x^l, y^{k,i}, m^l, m^k)^2 \partial^2_{y^{k,i}} g(y) \\ &+ \frac{1}{2} \sum_{k_1,k_2=1}^n \sum_{i_1,i_2=1}^2 \sum_{l \in I(k_1) \cap I(k_2)} f^l(x^l, m^l) \tilde{\Psi}^{lk_1}(x^l, y^{k_1,i_1}, m^l, m^{k_1}, u^l) \cdot \\ &\qquad \tilde{\Psi}^{lk_2}(x^l, y^{k_2,i_2}, m^l, m^{k_2}, u^l) \partial^2_{y^{k_1,i_1} y^{k_2,i_2}} g(y). \end{aligned}$$

Sketch of proof of Theorem 5.1. To prove the convergence in distribution of $(\mu^{N,1}, ..., \mu^{N,n})_N$, we begin by proving its tightness. Following the same reasoning as in Section 4.1, we can prove that, for each $1 \leqslant k \leqslant n$, the sequence $(\mu^{N,k})_N$ is tight on $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$. Hence, the sequence $(\mu^{N,1}, ..., \mu^{N,n})_N$ on $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))^n$.

Then a generalization of Lemma 4.2 allows to prove that the distribution of $(\bar{\mu}^1, ..., \bar{\mu}^n)$ is the unique solution of the martingale problem defined above.

Finally, we can conclude the proof showing that the law of any limit of a converging subsequence of $(\mu^{N,1}, ..., \mu^{N,n})_N$ is solution to the martingale problem using similar computations as the one in the proof of Theorem 4.3. □

6. Appendix

6.1. A priori estimates

Lemma 6.1. *Grant Assumptions 1, 2 and 3. For all $T > 0$,*

$$\sup_{n \in \mathbb{N}*} \mathbb{E}\left[\sup_{t \leqslant T} \left|X_t^{N,1}\right|^2\right] < \infty.$$

Proof. Notice that

$$\sup_{0 \leqslant s \leqslant t} |X_s^{N,1}| \leqslant (X_0^{N,1}) + ||b||_\infty t + \sup_{0 \leqslant s \leqslant t} \left|\int_0^s \sigma(X_r^{N,1}, \mu_r^N) d\beta_r^1\right| + \frac{1}{\sqrt{N}} \sup_{0 \leqslant s \leqslant t} |M_s^N|,$$

where M^N is the local martingale

$$M_t^N := \sum_{k=2}^{N} \int_{[0,t]\times\mathbb{R}_+\times E} \Psi(X_{s-}^{N,k}, X_{s-}^{N,1}, \mu_{s-}^N, u^k, u^1) \mathbb{1}_{\{z\leqslant f(X_{s-}^{N,k}, \mu_{s-}^N)\}} d\pi^k(s,z,u).$$

Consequently, by Burkholder-Davis-Gundy's inequality and Assumption 3,

$$\mathbb{E}\left[\sup_{0\leqslant s\leqslant t} |X_s^{N,1}|^2\right] \leqslant C + C||b||_\infty^2 t^2 + ||\sigma||_\infty^2 t + t||f||_\infty \frac{N-1}{N} \int_E \sup_{x,y,m} \Psi(x,y,m,u^1,u^2)^2 d\nu(u).$$

This proves the result. □

6.2. *Proof of* (20)

Lemma 6.2. *Grant Assumptions 1, 2 and 3. With the notation introduced in the proof of Theorem 4.3, we have*

$$\mathbb{E}\left[F(\mu^N)\right] \underset{N\to\infty}{\longrightarrow} \mathbb{E}\left[F(\mu)\right].$$

Proof. Let us recall that μ^N denotes the empirical measure of $(X^{N,i})_{1\leqslant i\leqslant N}$ and that μ is the limit in distribution of (a subsequence of) μ^N.

Step 1. We first show that almost surely, μ is supported by continuous trajectories. For that sake, we start showing that $P^N := \mathbb{E}\left[\mu^N\right] = \mathcal{L}(X^{N,1})$ is $C-$tight. This follows from Prop VI. 3.26 in Jacod and Shiryaev (2003), observing that

$$\lim_{N\to\infty} \mathbb{E}\left[\sup_{s\leqslant T} |\Delta X_s^{N,1}|^3\right] = 0,$$

which follows from our conditions on ψ. Indeed, writing $\psi^*(u^1,u^2) := \sup_{x,y,m} \psi(x,y,m,u^1,u^2)$, we can stochastically upper bound

$$\sup_{s\leqslant T} |\Delta X_s^{N,1}|^3 \leqslant \sup_{k\leqslant K} |\psi^*(U^{k,1}, U^{k,2})|^3 / N^{3/2},$$

where $K \sim Poiss(NT\|f\|_\infty)$ is Poisson distributed with parameter $NT\|f\|_\infty$, and where $(U^{k,1}, U^{k,2})_k$ is an i.i.d. sequence of $\nu_1\otimes\nu_1-$distributed random variables, independent of K. The conclusion then follows from the fact that due to our Assumption 3, $\mathbb{E}\left[|\psi^*(U^{k,1},U^{k,2})|^3\right] < \infty$ such that we can upper bound

$$\begin{aligned}\mathbb{E}\left[\sup_{k\leqslant K} |\psi^*(U^{k,1},U^{k,2})|^3/N^{3/2}\right] &\leqslant \mathbb{E}\left[\frac{1}{N^{3/2}} \sum_{k=1}^{K} |\psi^*(U^{k,1},U^{k,2})|^3\right] \\ &\leqslant \frac{\mathbb{E}\left[|\psi^*(U^{k,1},U^{k,2})|^3\right]}{N^{3/2}} \mathbb{E}\left[K\right] = \frac{\mathbb{E}\left[|\psi^*(U^{k,1},U^{k,2})|^3\right]}{N^{3/2}} NT\|f\|_\infty \to 0\end{aligned}$$

as $N\to 0$.

As a consequence of the above arguments, we know that $\mathbb{E}\left[\mu(\cdot)\right]$ is supported by continuous trajectories. In particular, almost surely, μ is also supported by continuous trajectories. Indeed,

$\mu(C(\mathbb{R}_+,\mathbb{R}))$ is a r.v. taking values in $[0,1]$, and its expectation equals one. Thus $\mu(C(\mathbb{R}_+,\mathbb{R}))$ equals one a.s.

We now turn to the heart of this proof and show that $\mathbb{E}\left[F(\mu^N)\right] \to \mathbb{E}\left[F(\mu\right]$. The latter expression contains terms like

$$\int_s^t b(Y_r^1,\mu_r)\partial_{x^1}\varphi(Y_r^1,Y_r^2)dr$$

for some bounded smooth function φ. However, by our assumptions, the continuity of $m \mapsto b(x,m)$ is expressed with respect to the Wasserstein $1-$distance. Yet, we only have information on the convergence of μ_r^N to μ_r for the topology of the weak convergence.

In what follows we make use of Skorokhod's representation theorem and realize all random measures μ^N and μ on an appropriate probability space such that we have almost sure convergence of these realizations (we do not change notation), that is, we know that almost surely,

$$\mu^N \to \mu$$

as $N \to \infty$. (Recall that we have already chosen a subsequence in the beginning of the proof of Theorem 4.3). Since μ is almost surely supported by continuous trajectories, we also know that almost surely, $\mu_t^N \to \mu_t$ weakly for all t (this is a consequence of Theorem 12.5.(i) of Billingsley (1999)).

Step 2. In a first time, let us prove that, a.s., for all r, μ_r^N converges to μ_r for the metric W_1. Thus we need to show additionally that almost surely, for all $t \geqslant 0$, $\int |x| d\mu_t^N(x) \to \int |x| d\mu_t(x)$.

To prove this last fact, it will be helpful to consider rather the convergence of the triplets $(\mu^N, X^{N,1}, \mu^N(|x|))$. Since the sequence of laws of these triplets is tight as well (the tightness of $(\mu^N)_N$ and $(X^{N,1})_N$ have been stated in Section 4.1, and the tightness of $(\mu^N(|x|)_N)$ is classical from Aldous' criterion since $\mu_t^N(|x|) = N^{-1}\sum_{k=1}^N |X_t^{N,k}|$), we may assume that, after having chosen another subsequence and then a convenient realization of this subsequence, we dispose of a sequence of random triplets such that almost surely, as $N \to \infty$,

$$(\mu^N, X^{N,1}, \mu^N(|x|)) \to (\mu, Y, A),$$

where $A = (A_t)_t$ is some process having càdlàg trajectories. In addition, it can be proven that the sequence $(\mu^N(|x|))_N$ is $C-$tight (for similar reasons as $(X^{N,1})_N$), hence A has continuous trajectories.

Taking a bounded and continuous function $\Phi : D(\mathbb{R}_+,\mathbb{R}) \to \mathbb{R}$, we observe that, as $N \to \infty$,

$$\mathbb{E}\left[\int_{D(\mathbb{R}_+,\mathbb{R})} \Phi d\mu\right] \leftarrow \mathbb{E}\left[\int_{D(\mathbb{R}_+,\mathbb{R})} \Phi d\mu^N\right] = \mathbb{E}\left[\Phi(X^{N,1})\right] \to \mathbb{E}\left[\Phi(Y)\right],$$

such that $\mathbb{E}\left[\mu\right] = \mathcal{L}(Y)$.

Notice that from the above follows that Y is necessarily a continuous process, since $\mathbb{E}\left[\mu\right]$ is supported by continuous trajectories. Notice also that for the moment we do not know if $A = \mu(|x|)$.

Using that $\sup_N \mathbb{E}\left[\sup_{t\leqslant T}|X_t^{N,1}|^2\right] < \infty$ (see our a priori estimates Lemma 6.1), we deduce that the sequence $(\sup_{t\leqslant T}|X_t^{N,1}|^{3/2})_N$ is uniformly integrable. Therefore, $\mathbb{E}\left[\sup_{t\leqslant T}|X_t^{N,1}|^{3/2}\right] \to \mathbb{E}\left[\sup_{t\leqslant T}|Y_t|^{3/2}\right] < \infty$. In particular, we also have that

$$\mathbb{E}\left[\sup_{t\leqslant T}\mu_t(|x|^{3/2})\right] < \infty \quad \text{and thus} \quad \sup_{t\leqslant T}\mu_t(|x|^{3/2}) < \infty \text{ almost surely,}$$

for all T, since

$$\mathbb{E}\left[\sup_{t\leqslant T}\mu_t(|x|^{3/2})\right] = \mathbb{E}\left[\sup_{t\leqslant T}\int_{D(\mathbb{R}_+,\mathbb{R})}|\gamma_t|^{3/2}\mu(d\gamma)\right] \leqslant \mathbb{E}\left[\int_{D(\mathbb{R}_+,\mathbb{R})}\sup_{t\leqslant T}|\gamma_t|^{3/2}\mu(d\gamma)\right]$$
$$= \mathbb{E}\left[\sup_{t\leqslant T}|Y_t|^{3/2}\right] < \infty.$$

We know that, a.s., μ^N converges weakly to μ and $\mu(C(\mathbb{R}_+,\mathbb{R})) = 1$. Let us fix some $\omega \in \Omega$ for which the two previous properties hold. In the following, we omit this ω in the notation. Let $\varepsilon > 0$, $t \leqslant T$ and choose M such that $\int |x| \wedge M d\mu_t \geqslant \int |x| d\mu_t - \varepsilon$. Then, as $N \to \infty$, almost surely,

$$\int |x| d\mu_t^N \geqslant \int |x| \wedge M d\mu_t^N \to \int |x| \wedge M d\mu_t.$$

Thus

$$\liminf_N \int |x|\ d\mu_t^N \geqslant \int |x| d\mu_t - \varepsilon,$$

such that

$$\liminf_N \int |x|\ d\mu_t^N \geqslant \int |x| d\mu_t. \tag{31}$$

Fatou's lemma implies that

$$\mathbb{E}\left[\liminf_N \int |x|\ d\mu_t^N\right] \leqslant \liminf_N \mathbb{E}\left[\int |x|\ d\mu_t^N\right] = \liminf_N \mathbb{E}\left[|X_t^{N,1}|\right] = \mathbb{E}\left[|Y_t|\right] = \mathbb{E}\left[\int |x| d\mu_t\right].$$

Together with (31) this implies that, almost surely,

$$\liminf_N \int |x|\ d\mu_t^N = \int |x| d\mu_t.$$

Finally, since $\int |x| d\mu^N \to A$ and since A is continuous, for all t,

$$\liminf_N \int |x|\ d\mu_t^N = \limsup_N \int |x|\ d\mu_t^N = \int |x| d\mu_t.$$

This implies that almost surely, for all $t \geqslant 0$, $\int |x| d\mu_t^N(x) \to \int |x| d\mu_t(x) = A_t < \infty$. In particular, almost surely, for all $t \geqslant 0$,

$$W_1(\mu_t^N, \mu_t) \to 0$$

(see e.g. Theorem 6.9 of Villani (2008)).

Step 3. Now we prove that $\mathbb{E}\left[F(\mu^N)\right]$ converges to $\mathbb{E}\left[F(\mu)\right]$, where we recall that

$$F(\mu) = \psi_1(\mu_{s_1}) \cdot \ldots \cdot \psi_k(\mu_{s_k}) \int_{D(\mathbb{R}_+,\mathbb{R})^2} \mu \otimes \mu(d\gamma)\varphi_1(\gamma_{s_1}) \ldots \varphi_k(\gamma_{s_k})$$
$$\left[\varphi(\gamma_t) - \varphi(\gamma_s) - \int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v)\nu_1(dv)\mu_r(dx)dr\right],$$

where $\psi_i \in C_b(\mathcal{P}(\mathbb{R})), \varphi_i \in C_b(\mathbb{R}^2)$ $(1 \leqslant i \leqslant k)$ and $\varphi \in C_b^3(\mathbb{R}^2)$. Let us recall some facts: by the boundedness of the functions ψ_i $(1 \leqslant i \leqslant k)$ and our boundedness Assumption 3, it is sufficient to prove the two following convergence:

$$\mathbb{E}\left[|\psi_1(\mu^N_{s_1}) \cdot \ldots \cdot \psi_k(\mu^N_{s_k}) - \psi_1(\mu_{s_1}) \cdot \ldots \cdot \psi_k(\mu_{s_k})|\right] \underset{N\to\infty}{\longrightarrow} 0, \tag{32}$$

$$\mathbb{E}\left[|G(\mu^N) - G(\mu)|\right] \underset{N\to\infty}{\longrightarrow} 0, \tag{33}$$

with

$$G(\mu) := \int_{D(\mathbb{R}_+,\mathbb{R})^2} \mu \otimes \mu(d\gamma)\varphi_1(\gamma_{s_1}) \ldots \varphi_k(\gamma_{s_k}) \left[\varphi(\gamma_t) - \varphi(\gamma_s) - \int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v)\nu_1(dv)\mu_r(dx)dr\right].$$

Indeed, since the functions ψ_i $(1 \leqslant i \leqslant k)$ and G are bounded, we have

$$\mathbb{E}\left[|F(\mu^N) - F(\mu)|\right] \leqslant C\mathbb{E}\left[|\psi_1(\mu^N_{s_1}) \cdot \ldots \cdot \psi_k(\mu^N_{s_k}) - \psi_1(\mu_{s_1}) \cdot \ldots \cdot \psi_k(\mu_{s_k})|\right] + C\mathbb{E}\left[|G(\mu^N) - G(\mu)|\right].$$

The convergence (32) follows from dominated convergence and the fact that the function

$$m \in \mathcal{P}(D(\mathbb{R}_+, \mathbb{R})) \mapsto \psi_1(m_{s_1})...\psi_k(m_{s_k}) \in \mathbb{R}$$

is bounded and continuous at μ, since μ is supported by continuous trajectories.

To prove the convergence (33), let us recall that we have already shown that

1. $\sup_N \sup_{0\leqslant s\leqslant t} \mathbb{E}\left[\mu^N_s(|x|^{3/2})\right] < \infty$,
2. $\sup_{0\leqslant s\leqslant t} \mathbb{E}\left[\mu_t(|x|^{3/2})\right] < \infty$,
3. $\mu(C(\mathbb{R}_+, \mathbb{R})) = 1$ *a.s.*
4. a.s. $\forall r$, μ^N_r converges to μ_r for the metric W_1,
5. for all $x, x' \in \mathbb{R}, y, y' \in \mathbb{R}^2, m, m' \in \mathcal{P}_1(\mathbb{R}), v \in \mathbb{R}$,
$$|L\varphi(y, m, x, v) - L\varphi(y', m', x', v)| \leqslant C(v)(||y - y'||_1 + |x - x'| + W_1(m, m')),$$
such that $\int_{\mathbb{R}} C(v)\nu_1(dv) < \infty$,
6. $$\int_{\mathbb{R}} \sup_{x,y,m} L\varphi(y, m, x, v)\nu_1(dv) < \infty.$$

In order to simplify the presentation, let us assume that the function G is of the form

$$G(\mu) = \int_{D^2} \mu \otimes \mu(d\gamma) \int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v)\nu_1(dv)\mu_r(dx)dr.$$

Now, let us show that $\mathbb{E}\left[|G(\mu^N) - G(\mu)|\right]$ vanishes as N goes to infinity.

$$|G(\mu) - G(\mu^N)| \leqslant \left|G(\mu) - \int_{D^2} \mu^N \otimes \mu^N(d\gamma) \left(\int_s^t \int_{\mathbb{R}} \int_{\mathbb{R}} L\varphi(\gamma_r, \mu_r, x, v)\nu_1(dv)\mu_r(dx)dr\right)\right|$$

$$
\begin{aligned}
&+\left|\int_{D^2}\mu^N\otimes\mu^N(d\gamma)\left(\int_s^t\int_{\mathbb{R}}\int_{\mathbb{R}}L\varphi(\gamma_r,\mu_r,x,v)\nu_1(dv)\mu_r(dx)dr\right)\right.\\
&\quad\left.-\int_{D^2}\mu^N\otimes\mu^N(d\gamma)\left(\int_s^t\int_{\mathbb{R}}\int_{\mathbb{R}}L\varphi(\gamma_r,\mu_r,x,v)\nu_1(dv)\mu_r^N(dx)dr\right)\right|\\
&+\left|G(\mu^N)-\int_{D^2}\mu^N\otimes\mu^N(d\gamma)\left(\int_s^t\int_{\mathbb{R}}\int_{\mathbb{R}}L\varphi(\gamma_r,\mu_r,x,v)\nu_1(dv)\mu_r^N(dx)dr\right)\right|\\
&=: A_1+A_2+A_3.
\end{aligned}
$$

We first show that A_1 vanishes a.s. (this implies that $\mathbb{E}[A_1]$ vanishes by dominated convergence). A_1 is of the form

$$A_1=\left|\int_{D^2}\mu\otimes\mu(d\gamma)H(\gamma)-\int_{D^2}\mu^N\otimes\mu^N(d\gamma)H(\gamma)\right|,$$

with

$$H:\gamma\in D^2\mapsto\int_s^t\int_{\mathbb{R}}\int_{\mathbb{R}}L\varphi(\gamma_r,\mu_r,x,v)\nu_1(dv)\mu_r(dx)dr\in\mathbb{R}.$$

We just have to prove that H is continuous and bounded. The boundedness is obvious, so let us verify the continuity.

Let $(\gamma^n)_n$ converge to γ in $D(\mathbb{R}_+,\mathbb{R})^2$. We have

$$
\begin{aligned}
|H(\gamma)-H(\gamma^n)|&\leqslant\int_s^t\int_{\mathbb{R}}\int_{\mathbb{R}}|H(\gamma_r,\mu_r,x,v)-H(\gamma_r^n,\mu_r,x,v)|\nu_1(dv)\mu_r(dx)dr\\
&\leqslant\int_s^t\int_{\mathbb{R}}\int_{\mathbb{R}}C(v)||\gamma_r-\gamma_r^n||_1\nu_1(dv)\mu_r(dx)dr\\
&\leqslant C\int_s^t||\gamma_r-\gamma_r^n||_1dr,
\end{aligned}
$$

which vanishes by dominated convergence: the integrand vanishes at every continuity point r of γ (whence for a.e. r), and, for n big enough, $\sup_{r\leqslant t}||\gamma_r^n||_1\leqslant 2\sup_{r\leqslant t}||\gamma_r||_1$.

Now we show that $\mathbb{E}[A_2]$ vanishes. We have

$$
\begin{aligned}
A_2\leqslant\int_{D^2}&\mu^N\otimes\mu^N(d\gamma)\\
&\left(\int_s^t\left|\int_{\mathbb{R}}\int_{\mathbb{R}}L\varphi(\gamma_r,\mu_r,x,v)\nu_1(dv)\mu_r(dx)-\int_{\mathbb{R}}\int_{\mathbb{R}}L\varphi(\gamma_r,\mu_r,x,v)\nu_1(dv)\mu_r^N(dx)\right|dr\right).
\end{aligned}
$$

Since the function $x\in\mathbb{R}\mapsto\int_{\mathbb{R}}L\varphi(\gamma_r,\mu_r,x,v)\nu_1(dv)$ is Lipschitz continuous (with Lipschitz constant independent of γ_r and μ_r), we have, by Kantorovich-Rubinstein duality (see e.g. Remark 6.5 of Villani (2008)),

$$A_2\leqslant C\int_{D^2}\mu^N\otimes\mu^N(d\gamma)\int_s^t W_1(\mu_r^N,\mu_r)dr=C\int_s^t W_1(\mu_r^N,\mu_r)dr.$$

Hence

$$\mathbb{E}[A_2]\leqslant C\int_s^t\mathbb{E}\left[W_1(\mu_r^N,\mu_r)\right]dr,$$

which vanishes by dominated convergence: the integrand vanishes thanks to *Step 2*, and the uniform integrability follows from the fact that

$$\sup_N \int_s^t \mathbb{E}\left[W_1(\mu_r^N,\mu_r)^{3/2}\right] dr \leqslant C(t-s)\sup_N \sup_{0\leqslant s\leqslant t} \mathbb{E}\left[\mu_s^N(|x|)^{3/2}\right] + C(t-s)\sup_{0\leqslant s\leqslant t} \mathbb{E}\left[\mu_s(|x|)^{3/2}\right].$$

We finally show that $\mathbb{E}[A_3]$ vanishes.

$$\begin{aligned} A_3 &\leqslant \int_{D^2} \mu^N\otimes\mu^N(d\gamma)\left(\int_s^t\int_{\mathbb{R}}\int_{\mathbb{R}} \left|L\varphi(\gamma_r,\mu_r^N,x,v)-L\varphi(\gamma_r,\mu_r,x,v)\right|\nu_1(dv)\mu_r^N(dx)dr\right)\\ &\leqslant \int_{D^2} \mu^N\otimes\mu^N(d\gamma)\left(\int_s^t\int_{\mathbb{R}}\int_{\mathbb{R}} C(v)W_1(\mu_r^N,\mu_r)\nu_1(dv)\mu_r^N(dx)dr\right). \end{aligned}$$

Then,

$$\mathbb{E}[A_3] \leqslant C\int_s^t \mathbb{E}\left[W_1(\mu_r^n,\mu_r)\right]dr,$$

which vanishes for the same reasons as in the previous step where we have shown that $\mathbb{E}[A_2]$ vanishes. □

Bibliography

Aldous, D. (1983). Exchangeability and related topics. In *Ecole d'Eté de Probabilités de Saint-Flour: XIII - 1983*. springer-verlag ed. *Lecture notes in mathematics* **1117**.

Andreis, L., Dai Pra, P. and Fischer, M. (2018). McKean–Vlasov limit for interacting systems with simultaneous jumps. *Stochastic Analysis and Applications* **36** 960–995. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07362994.2018.1486202.

Billingsley, P. (1999). *Convergence of Probability Measures*, Second ed. Wiley Series In Probability And Statistics.

Chevallier, J. and Ost, G. (2020). Fluctuations for spatially extended Hawkes processes. *Stochastic Processes and their Applications* **130** 5510-5542.

De Masi, A., Galves, A., Löcherbach, E. and Presutti, E. (2015). Hydrodynamic Limit for Interacting Neurons. *Journal of Statistical Physics* **158** 866–902.

Duarte, A., Ost, G. and Rodríguez, A. A. (2015). Hydrodynamic Limit for Spatially Structured Interacting Neurons. *Journal of Statistical Physics* **161** 1163–1202.

El Karoui, N. and Méléard, S. (1990). Martingale measures and stochastic calculus. *Probability Theory and Related Fields* **84** 83–101.

Erny, X., Löcherbach, E. and Loukianova, D. (2020). Conditional propagation of chaos for mean field systems of interacting neurons. *arXiv:1909.02925 [math]*.

Fournier, N. and Löcherbach, E. (2016). On a toy model of interacting neurons. *Annales de l'Institut Henri Poincaré - Probabilités et Statistiques* **52** 1844–1876.

Fournier, N. and Meleard, S. (2002). A Stochastic Particle Numerical Method for 3D Boltzmann Equations without Cutoff. *Mathematics of Computation* **71** 583–604.

Fournier, N. and Mischler, S. (2016). Rate of convergence of the Nanbu particle system for hard potentials and Maxwell molecules. *Ann. Probab.* **44** 589–627.

Graham, C. (1992). McKean-Vlasov Ito-Skorohod equations, and nonlinear diffusions with discrete jump sets. *Stochastic Processes and their Applications* **40** 69–82.

Graham, C. (2008). Chaoticity for Multiclass Systems and Exchangeability Within Classes. *Journal of Applied Probability* **45** 1196–1203. Publisher: Cambridge University Press.

Graham, C. and Méléard, S. (1997). Stochastic particle approximations for generalized Boltzmann models and convergence estimates. *Ann. Probab.* **25** 115–132.

Gärtner, J. (1988). On the McKean-Vlasov Limit for Interacting Diffusions. *Mathematische Nachrichten* **137** 197-248.

Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*, Second ed. Springer-Verlag BerlinHeidelberg NewYork.

Major, P. (1978). On the invariance principle for sums of independent identically distributed random variables. *Journal of Multivariate Analysis* **8** 487–517.

Meleard, Sylvie (1998). Stochastic approximations of the solution of a full Boltzmann equation with small initial data. *ESAIM: PS* **2** 23-40.

Sznitman, A.-S. (1989). Topics in propagation of chaos. In *Ecole d'Eté de Probabilités de Saint-Flour: XIX - 1989. Lecture notes in mathematics* **1464** 167–251. Springer, Berlin. OCLC: 23253880.

Tanaka, H. (1978). Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **46** 67–105.

Villani, C. (2008). *Optimal transport, old and new*. Springer.

Walsh, J. B. (1986). An introduction to stochastic partial differential equations. In *École d'Été de Probabilités de Saint Flour XIV - 1984. Lecture Notes in Mathematics* 265–439. Springer, Berlin, Heidelberg.