HAL open science

Conditional propagation of chaos for mean field systems of interacting neurons

Xavier Erny, Eva Löcherbach, Dasha Loukianova

To cite this version:

Xavier Erny, Eva Löcherbach, Dasha Loukianova. Conditional propagation of chaos for mean field systems of interacting neurons. 2019. hal-02280882v2

HAL Id: hal-02280882

https://hal.science/hal-02280882v2

Preprint submitted on 21 Sep 2019 (v2), last revised 2 Mar 2020 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Conditional propagation of chaos for mean field systems of interacting neurons

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

[*] *Laboratoire de Mathématiques et Modélisation d'Évry, Université d'Évry Val d'Essonne, UMR CNRS 8071 and*
[†] *Statistique, Analyse et Modélisation Multidisciplinaire, Université Paris 1 Panthéon-Sorbonne, EA 4543*

Abstract: We study the stochastic system of interacting neurons introduced in De Masi et al. (2015) and in Fournier and Löcherbach (2016) in a diffusive scaling. The system consists of N neurons, each spiking randomly with rate depending on its membrane potential. At its spiking time, the potential of the spiking neuron is reset to 0 and all other neurons receive an additional amount of potential which is a centred random variable of order $1/\sqrt{N}$. In between successive spikes, each neuron's potential follows a deterministic flow. We prove the convergence of the system, as $N \to \infty$, to a limit nonlinear jumping stochastic differential equation driven by Poisson random measure and an additional Brownian motion W which is created by the central limit theorem. This Brownian motion is underlying each particle's motion and induces a common noise factor for all neurons in the limit system. Conditionally on W, the different neurons are independent in the limit system. We call this property *conditional propagation of chaos*. We show the convergence in distribution, prove strong convergence with respect to an appropriate distance, and we get an explicit rate of convergence. The main technical ingredient of our proof is the famous coupling introduced in Komlós, Major and Tusnády (1976) of the point process representing the small jumps of the particle system with the limit Brownian motion.

MSC 2010 subject classifications: 60J75, 60K35, 60G55, 60G09.
Keywords and phrases: Multivariate nonlinear Hawkes processes with variable length memory, Mean field interaction, Piecewise deterministic Markov processes, Interacting particle systems, Propagation of chaos, Exchangeability, Hewitt Savage theorem, KMT approximation.

Introduction

This paper is devoted to the study of the Markov process $X_t^N = (X_t^{N,1}, \dots, X_t^{N,N})$ taking values in $\mathbb{R}^N$ and solving, for $i = 1, \dots, N$, for $t \geqslant 0$,

$$X_t^{N,i} = X_0^{N,i} + \int_0^t b(X_s^{N,i})ds - \int_0^t X_{s-}^{N,i} dZ_s^{N,i} + \frac{1}{\sqrt{N}} \sum_{j \neq i} \int_0^t U_j(s) dZ_s^{N,j}, \tag{1}$$

where $U_j(s)$ are i.i.d. centred random variables and where for each $1 \leqslant j \leqslant N$, $Z^{N,j}$ is a simple point process on $\mathbb{R}_+$ having stochastic intensity $s \mapsto f\left(X_{s-}^{N,j}\right)$.

The particle system (1) is a version of the model of interacting neurons considered in De Masi et al. (2015), inspired by Galves and Löcherbach (2013), and then further studied in Fournier and Löcherbach (2016) and Cormier, Tanré and Veltz (2018). The system consists of N interacting neurons. In (1), $Z_t^{N,j}$ represents the number of spikes emitted by the neuron j in the interval $[0, t]$ and $X_t^{N,j}$ the membrane potential of the neuron j at time t. Spiking occurs randomly following a point process of rate $f(x)$ for any neuron of which the membrane potential equals x. Each time a neuron emits a spike, the potentials of all other neurons receive an additional amount of potential. In De Masi et al. (2015), Fournier and Löcherbach (2016) and Cormier, Tanré and Veltz (2018) this

amount is of order N^{-1}, leading to classical mean field limits as $N \to \infty$. On the contrary to this, in the present article we study a *diffusive scaling* where each neuron j receives the amount $U_i(t)/\sqrt{N}$ at spike times t of neuron $i, i \neq j$. The variable $U_i(t)$ is centred modeling the fact that the synaptic weights are balanced. Moreover, right after its spike, the potential of the spiking neuron i is reset to 0, interpreted as resting potential. Finally, in between successive spikes, each neuron's potential follows a deterministic flow with drift b.

Equations similar to (1) appear also in the frame of multivariate Hawkes processes with mean field interactions. Indeed, if $\left(Z^{N,i}\right)_{1\leqslant i\leqslant N}$ is a multivariate Hawkes process where the stochastic intensity of each $Z^{N,i}$ is given by $f\left(X^N_{t-}\right)_t$ with

$$X^N_t = e^{-\alpha t}X^N_0 + \frac{1}{\sqrt{N}}\sum_{j=1}^N \int_0^t e^{-\alpha(t-s)}U_j(s)dZ^{N,j}_s, \tag{2}$$

then X^N satisfies

$$X^N_t = X^N_0 - \alpha\int_0^t X^N_s ds + \frac{1}{\sqrt{N}}\sum_{j=1}^N \int_0^t U_j(s)dZ^{N,j}_s,$$

which corresponds to equation (1) with $b(x) = -\alpha x$, but without the big jumps, i.e. without the reset to 0 after each spike.

The above model of Hawkes processes has been studied in our previous paper Erny, Löcherbach and Loukianova (2019). There we have shown firstly that X^N converges in distribution in $D(\mathbb{R}_+, \mathbb{R})$ to a limit process $\bar{X}$ solving

$$d\bar{X}_t = -\alpha\bar{X}_t dt + \sigma\sqrt{f\left(\bar{X}_t\right)}dW_t, \tag{3}$$

and secondly that the sequence of multivariate counting processes $\left(Z^{N,i}\right)_i$ converges in distribution in $D(\mathbb{R}_+, \mathbb{R})^{\mathbb{N}^*}$ to a limit sequence of counting processes $\left(\bar{Z}^i\right)_i$. Here, every $\bar{Z}^i$ is driven by its own Poisson random measure and has the same intensity $\left(f(\bar{X}_{t-})\right)_t$, $\bar{X}$ the strong solution of (3) with respect to some Brownian motion W. Consequently, the processes $\bar{Z}^i$ $(i \geqslant 1)$ are conditionally independent given the Brownian motion W.

In the present paper we add the reset term in (1) that forces the potential $X^{N,i}$ of neuron i to go back to 0 at each jump time of $Z^{N,i}$. This models the well-known biological fact that right after its spike, the membrane potential of the spiking neuron is reset to a resting potential which we choose to be equal to 0. From a mathematical point of view, this reset to 0 induces a de-synchronization of the processes $X^{N,i}$ $(1 \leqslant i \leqslant N)$. In terms of Hawkes processes, it means that in (2), the process X^N_t has been replaced by

$$X^{N,i}_t = \frac{1}{\sqrt{N}}\sum_{j=1}^N \int_{L^i_t}^t e^{-\alpha(t-s)}U_j(s)dZ^{N,j}_s, \quad \text{where } L^i_t = \sup\{s \leqslant t : \Delta Z^{N,i}_s = 1\}$$

is the last spiking time of neuron i before time t.[1] Thus the integral over the past, starting from 0 in (2), is replaced by an integral starting at the last jump time before the present time. In Galves and Löcherbach (2013), such processes are termed being of *variable length memory*, in reminiscence of Rissanen (1983), and we are thus considering *multivariate Hawkes processes with mean field interactions and variable length memory*. As a consequence, on the contrary to the situation in

[1] In the present paper, the drift $b(x) = -\alpha x$ of (3) has been replaced by a general drift coefficient.

Erny, Löcherbach and Loukianova (2019), the point processes $Z^{N,i}$ $(1 \leqslant i \leqslant N)$ do not share the same stochastic intensity. It turns out that the reset term in (1) is a jump term that survives in the limit $N \to \infty$.

Before introducing the exact limit equation for the system (1), let us explain informally how the limit particle system associated to $\left(X^{N,i}\right)_{1\leqslant i\leqslant N}$ should a priori look like. So suppose that there exists a process $(\bar{X}^1, \bar{X}^2, \bar{X}^3, \ldots) \in D(\mathbb{R}_+, \mathbb{R})^{\mathbb{N}^*}$ such that for all $K > 0$, we have weak convergence $\mathcal{L}(X^{N,1,}, \ldots, X^{N,K}) \to \mathcal{L}(\bar{X}^1, \ldots, \bar{X}^K)$ in $D(\mathbb{R}_+, \mathbb{R})^K$, as $N \to \infty$. In equation (1) the only term that depends on N is the martingale term which is approximately given by

$$M_t^N = \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_0^t U_j(s) dZ_s^{N,j}.$$

Each $\bar{X}^i$ should then solve the equation (1), where the term M_t^N is replaced by $M_t := \lim_{N\to\infty} M_t^N$. Because of the scaling in $N^{-1/2}$, the limit martingale M_t should be a stochastic integral with respect to some Brownian motion, and its variance should be the limit of

$$\mathbb{E}\left[(M_t^N)^2\right] = \sigma^2 \int_0^t \mathbb{E}\left[\frac{1}{N}\sum_{j=1}^N f(X_s^{N,j})\right] ds,$$

where σ^2 is the variance of $U_j(s)$. Therefore, the limit martingale should be of the form

$$M_t = \sigma \int_0^t \sqrt{\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^N f\left(X_s^{N,j}\right)} dW_s = \sigma \int_0^t \sqrt{\lim_{N\to\infty} \mu_s^N(f)} dW_s,$$

where μ_s^N is the empirical measure of the system $\left(X_s^{N,j}\right)_{1\leqslant j\leqslant N}$.

Since the law of the $N-$particle system $(X^{N,1}, \ldots, X^{N,N})$ is symmetric, the law of the limit system $\bar{X} = (\bar{X}^1, \bar{X}^2, \bar{X}^3, \ldots)$ must be exchangeable, that is, for all finite permutations σ, we have that $\mathcal{L}(\bar{X}^{\sigma(1)}, \bar{X}^{\sigma(2)}, \ldots) = \mathcal{L}(\bar{X})$. In particular, the theorem of Hewitt-Savage, see Hewitt and Savage (1955), implies that the random limit

$$\mu_s := \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^N \delta_{\bar{X}_s^i} \tag{4}$$

exists. Supposing that μ_s^N converges, it necessarily converges towards μ_s. Therefore, $\bar{X}$ should solve the limit system

$$\bar{X}_t^i = \bar{X}_0^i + \int_0^t b(\bar{X}_s^i) ds - \int_0^t \bar{X}_{s-}^i d\bar{Z}_s^i + \sigma \int_0^t \sqrt{\mu_s(f)} dW_s, i \in \mathbb{N}, \tag{5}$$

where $(W_t)_{t\geqslant 0}$ is a standard one-dimensional Brownian motion, where each $\bar{Z}^i$ has intensity $t \mapsto f(\bar{X}_{t-}^i)$, and where μ_s is given by (4).

Analogously to Erny, Löcherbach and Loukianova (2019), the scaling in $N^{-1/2}$ in (1) creates a Brownian motion W in the limit system (5). We will show that the presence of this Brownian motion entails a *conditional propagation of chaos*, that is the conditional independence of the particles

given W. In particular, the limit measure μ_s will be random. This differs from the classical framework, where the scaling is in N^{-1} (see e.g. Delattre, Fournier and Hoffmann (2016), Ditlevsen and Löcherbach (2017) in the framework of Hawkes processes, and De Masi et al. (2015), Fournier and Löcherbach (2016) and Cormier, Tanré and Veltz (2018) in the framework of systems of interacting neurons), leading to a deterministic limit measure μ_s and the true propagation of chaos property implying that the particles of the limit system are independent.

This is not the first time that conditional propagation of chaos is studied in the literature; it has already been considered e.g. in Carmona, Delarue and Lacker (2016), Coghi and Flandoli (2016) and Dermoune (2003). But in these papers the common noise, represented by a common (maybe infinite dimensional) Brownian motion, is already present at the level of the finite particle system, the mean field interactions act on the drift of each particle, and the scaling is the classical one in N^{-1}. On the contrary to this, in our model, this common Brownian motion, leading to conditional propagation of chaos, is only present in the limit, and it is created by the central limit theorem as a consequence of the joint action of the small jumps of the finite size particle system. Moreover, in our model, the interactions survive as a variance term in the limit system as a consequence of the diffusive scaling in $N^{-1/2}$.

Now let us discuss the form of μ_s, which is the limit of the empirical measures of the limit system $\left(\bar{X}^i_s\right)_{i\geqslant 1}$. The theorem of Hewitt-Savage, Hewitt and Savage (1955), implies that the law of $\left(\bar{X}^i_s\right)_{i\geqslant 1}$ is a mixture directed by the law of μ_s. As it has been remarked by Carmona, Delarue and Lacker (2016) and Coghi and Flandoli (2016), this conditioning reflects the dependencies between the particles.

Since the variables $\bar{X}^i$ are conditionally independent given the Brownian motion W, μ_s will be shown to be the conditional law of the solution given the Brownian motion, that is, $P-$almost surely,

$$\mu_s(\cdot) = P(\bar{X}^i_s \in \cdot | (W_t)_{0\leqslant t\leqslant s}) = P(\bar{X}^i_s \in \cdot | W), \tag{6}$$

for any $i \in \mathbb{N}$. Equation (5) together with (6) gives a precise definition of the limit system.

The nonlinear SDE (5) is not clearly well-posed, and our first main result, Theorem 1.1, gives appropriate conditions on the coefficients b and f of the system that guarantee pathwise uniqueness and the existence of a strong solution to (5). We then establish the convergence of the system $\left(X^{N,i}\right)_{1\leqslant i\leqslant N}$ to $\left(\bar{X}^i\right)_{i\geqslant 1}$. We prove strong convergence with respect to an appropriate distance in an L^1-sense together with a rate of convergence in Theorem 1.6, and convergence in distribution in Theorem 1.3.

To prove the strong convergence, we couple the point processes of (1) with the Brownian motion appearing in the limit equation (5) using ideas that go back to Kurtz (1978). This coupling is based on a corollary of the KMT inequality (see Theorem 1 of Komlós, Major and Tusnády (1976)). To the best of our knowledge, this strategy of proof is completely new and has neither been used in Erny, Löcherbach and Loukianova (2019) nor in the frame of classical mean field limits where the scaling is in N^{-1}.

Finally, Proposition 1.8 states the convergence in law of the sequence of empirical measures $\mu^N = N^{-1}\sum_{i=1}^N \delta_{(X^{N,i}_t)_{t\geqslant 0}}$, in $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$, to the random limit $\mu = P((\bar{X}_t)_{t\geqslant 0} \in \cdot | W)$. This random limit measure μ satisfies the following nonlinear stochastic PDE in weak form: for any test function $\varphi \in C^2_b(\mathbb{R})$, the set of C^2-functions on $\mathbb{R}$ such that φ, φ' and φ'' are bounded, for any $t \geqslant 0$,

$$\int_{\mathbb{R}} \varphi(x)\mu_t(dx) = \int_{\mathbb{R}} \varphi(x)\mu_0(dx) + \int_0^t \left(\int_{\mathbb{R}} \varphi'(x)\mu_s(dx)\right) \sqrt{\mu_s(f)}dW_s$$

$$+\int_0^t\int_{\mathbb{R}}\Big([\varphi(0)-\varphi(x)]f(x)+\varphi'(x)b(x)+\frac{1}{2}\varphi''(x)\mu_s(f)\Big)\mu_s(dx)ds.$$

Organisation of the paper. In Section 1, we introduce formally the systems that we will study, we fix some notations, and we state the main results, Theorems 1.1, 1.3 and 1.6 and Proposition 1.8. Section 2 is devoted to the proofs of the main results. Finally, in Appendix Section 3, we prove some important technical results that we use in the paper, in particular we complete the proof of the well-posedness of the limit system together with some useful a priori estimates.

1. Notation, Model and main results

1.1. Notation

Let us introduce some notation we use throughout the paper.

If E is a metric space, we note:

- $\mathcal{P}(E)$ the space of probability measures on E endowed with the topology of the weak convergence,
- $C_b^n(E)$ the set of the functions g which are n times continuously differentiable such that $g^{(k)}$ is bounded for each $0\leqslant k\leqslant n$,
- $C_c^n(E)$ the set of functions $g\in C_b^n(E)$ that have a compact support.

In addition, in what follows $D(\mathbb{R}_+,\mathbb{R})$ denotes the space of càdlàg functions from $\mathbb{R}_+$ to $\mathbb{R}$, endowed with the topology of the uniform convergence on every compact set, 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.

In the sequel, ν will denote a probability measure on $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ with $\int_{\mathbb{R}}u\nu(du)=0$ and with $\int_{\mathbb{R}}u^2\nu(du)=\sigma^2$.

1.2. The finite system

We consider, for each $N\geqslant 1$, a family of i.i.d. Poisson measures $(\pi^i(ds,dz,du))_{i=1,\dots,N}$ on $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}$ having intensity measure $dsdz\nu(du)$, as well as an i.i.d. family $(X_0^{N,i})_{i=1,\dots,N}$ of $\mathbb{R}$-valued random variables independent of the Poisson measures. The object of this paper is to study the convergence of the Markov process $X_t^N=(X_t^{N,1},\dots,X_t^{N,N})$ taking values in $\mathbb{R}^N$ and solving, for $i=1,\dots,N$, for $t\geqslant 0$,

$$\begin{cases} X_t^{N,i} &= X_0^{N,i}+\int_0^t b(X_s^{N,i})ds-\int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}}X_{s-}^{N,i}\mathbb{1}_{\{z\leqslant f(X_{s-}^{N,i})\}}\pi^i(ds,dz,du)\\ &\quad+\frac{1}{\sqrt{N}}\sum_{j\neq i}\int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}}u\mathbb{1}_{\{z\leqslant f(X_{s-}^{N,j})\}}\pi^j(ds,dz,du),\\ X_0^{N,i} &\sim \nu_0. \end{cases}\tag{7}$$

The coefficients of this system are the drift function $b:\mathbb{R}\to\mathbb{R}$, the jump rate function $f:\mathbb{R}\mapsto\mathbb{R}_+$ and the probability measures ν and ν_0. The generator of the process X^N is given for any smooth test function $\varphi:\mathbb{R}^N\to\mathbb{R}$ by

$$L\varphi(x)=\sum_{i=1}^N\partial_{x_i}\varphi(x)b(x_i)+\sum_{i=1}^N f(x_i)\int_{\mathbb{R}}\nu(du)\left(\varphi(x-x_ie_i+\sum_{j\neq i}\frac{u}{\sqrt{N}}e_j)-\varphi(x)\right),$$

where $x = (x_1, \dots, x_N)$ and where e_j denotes the $j-$th unit vector in $\mathbb{R}^N$.

In order to guarantee existence and uniqueness of a strong solution of (7), we introduce the following hypothesis.

Assumption 1. *The functions f and b are Lipschitz continuous.*

In addition, we also need the following condition to obtain a priori bounds on some moments of the process $\left(X^{N,i}\right)_{1\leqslant i\leqslant N}$.

Assumption 2. *We assume that $\int_{\mathbb{R}} x d\nu(x) = 0$, $\int_{\mathbb{R}} x^2 d\nu(x) < +\infty$, and $\int_{\mathbb{R}} x^2 d\nu_0(x) < +\infty$.*

Under Assumptions 1 and 2, existence and uniqueness of strong solutions of (7) follow from Theorem IV.9.1 of Ikeda and Watanabe (1989), exactly in the same way as in Proposition 6.6 of Erny, Löcherbach and Loukianova (2019).

1.3. The limit system

The limit system $\left(\bar{X}^i\right)_{i\geqslant 1}$ satisfies the following dynamic

$$\begin{cases} \bar{X}^i_t &= \bar{X}^i_0 + \int_0^t b(\bar{X}^i_s)ds - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \bar{X}^i_{s-}\mathbb{1}_{\{z\leqslant f(\bar{X}^i_{s-})\}}\pi^i(ds,dz,du) \\ & \quad +\sigma\int_0^t \sqrt{\mathbb{E}\left[f\left(\bar{X}^i_s\right)|\,\mathcal{W}_s\right]}dW_s, \\ \bar{X}^i_0 &\sim \nu_0. \end{cases} \tag{8}$$

In the above equation, $(W_t)_{t\geqslant 0}$ is a standard one-dimensional Brownian motion which is independent of the Poisson random measures, and $\mathcal{W}_s = \sigma\{W_t, t\leqslant s\}$. Moreover, the initial positions $\bar{X}^i_0, i\geqslant 1$, are i.i.d., independent of W and of the Poisson random measures, distributed according to ν_0 which is the same probability measure as in (7). The common jumps of the particles in the finite system, due to their scaling in $1/\sqrt{N}$ and the fact that they are centred, by the Central Limit Theorem, create this single Brownian motion W_t which is underlying each particle's motion and which induces the common noise factor for all particles in the limit.

The limit equation (8) is not clearly well-posed and requires more conditions on the rate function f. Let us briefly comment on the type of difficulties that one encounters when proving trajectorial uniqueness of (8). Roughly speaking, the jump terms demand to work in an L^1-framework, whereas the diffusive terms demand to work in an L^2-framework. Graham (1992) proposes a unified approach to deal both with jump and with diffusion terms in a non-linear framework, and we shall rely on his ideas in the sequel. The presence of the random volatility term which involves conditional expectation causes however additional technical difficulties. Finally, another difficulty comes from the fact that the jumps induce non-Lipschitz terms of the form $\bar{X}^i_s f(\bar{X}^i_s)$. For this reason a classical Wasserstein-1$-$coupling is not appropriate for the jump terms. Therefore we propose a different distance which is inspired by the one already used in Fournier and Löcherbach (2016). To do so, we need to work under the following additional assumption.

Assumption 3. *1. We suppose that* $\inf f > 0$.
2. There exists a function $a \in C^2(\mathbb{R},\mathbb{R}_+)$, strictly increasing and bounded, such that, for a suitable constant C, for all $x, y \in \mathbb{R}$,

$$|a''(x) - a''(y)| + |a'(x) - a'(y)| + |b(x) - b(y)| + |f(x) - f(y)| \leqslant C|a(x) - a(y)|.$$

Note that Assumption 3 implies Assumption 1 as well as the boundedness of the rate function f. An example where Assumption 3 is satisfied is $f(x) = c + d\arctan(x)$, where $c > d\frac{\pi}{2}$, $d > 0$, with a similar choice for b. In this case, we choose $a = Cf$.

Under these additional assumptions we obtain the well-posedness of each coordinate of the limit system (8), that is, of the $(\mathcal{F}_t)_t-$ adapted process $(\bar{X}_t)_t$ which is solution of the SDE

$$\begin{cases} d\bar{X}_t &= b\left(\bar{X}_t\right)dt - \bar{X}_{t-}\int_{\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\{z\leqslant f(\bar{X}_{t-})\}}\pi(dt,dz,du) + \sigma\sqrt{\mu_t(f)}dW_t, \\ \bar{X}_0 &\sim \nu_0, \end{cases} \tag{9}$$

where $\mu_t(f) = \mathbb{E}\left[f\left(\bar{X}_t\right)\middle|\mathcal{W}_t\right]$ and where $\mathcal{F}_t = \sigma\{\pi([0,s]\times A), s\leqslant t, A\in\mathcal{B}(\mathbb{R}_+\times\mathbb{R})\}\vee\mathcal{W}_t$.

Theorem 1.1. *Grant Assumption 3.*
1. Pathwise uniqueness holds for the nonlinear SDE (9).
2. If additionally, $\int_{\mathbb{R}} x^2 d\nu_0(x) < +\infty$, then there exists a strong solution $(\bar{X}_t)_{t\geqslant 0}$ of the nonlinear SDE (9) that satisfies, for every $t > 0$,

$$\sup_{0\leqslant s\leqslant t}\mathbb{E}\left[\bar{X}_s^2\right] < +\infty.$$

In what follows we just give the proof of Item 1. of the above theorem since its arguments are important for the sequel. We postpone the rather classical proof of Item 2. to Appendix.

Proof of Item 1. of Theorem 1.1. Consider two solutions $(\widehat{X}_t)_{t\geqslant 0}$ and $(\check{X}_t)_{t\geqslant 0}$, defined on the same probability space and driven by the same Poisson random measure π and the same Brownian motion W, and with $\widehat{X}_0 = \check{X}_0$. We consider $Z_t := a(\widehat{X}_t) - a(\check{X}_t)$, for all $t\leqslant T$. Recall $\widehat{\mu}_s(f) = \mathbb{E}[f(\widehat{X}_s)|\mathcal{W}_s]$ and denote $\check{\mu}_s(f) = \mathbb{E}[f(\check{X}_s)|\mathcal{W}_s]$.
Using Ito's formula, we can write

$$\begin{aligned} Z_t = \int_0^t \left(b(\widehat{X}_s)a'(\widehat{X}_s) - b(\check{X}_s)a'(\check{X}_s)\right)ds + \frac{1}{2}\int_0^t (a''(\widehat{X}_s)\widehat{\mu}_s(f) - a''(\check{X}_s)\check{\mu}_s(f))\sigma^2 ds \\ + \int_0^t (a'(\widehat{X}_s)\sqrt{\widehat{\mu}_s(f)} - a'(\check{X}_s)\sqrt{\check{\mu}_s(f)})\sigma dW_s \\ - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} [a(\widehat{X}_{s-}) - a(\check{X}_{s-})]\mathbb{1}_{\{z\leqslant f(\widehat{X}_{s-})\wedge f(\check{X}_{s-})\}}\pi(ds,dz,du) \\ + \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} [a(0) - a(\widehat{X}_{s-})]\mathbb{1}_{\{f(\check{X}_{s-})<z\leqslant f(\widehat{X}_{s-})\}}\pi(ds,dz,du) \\ + \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} [a(\check{X}_{s-}) - a(0)]\mathbb{1}_{\{f(\widehat{X}_{s-})<z\leqslant f(\check{X}_{s-})\}}\pi(ds,dz,du) =: A_t + M_t + \Delta_t, \end{aligned}$$

where A_t denotes the bounded variation part of the evolution, M_t the martingale part and Δ_t the sum of the three jump terms. Notice that

$$M_t = \int_0^t (a'(\widehat{X}_s)\sqrt{\widehat{\mu}_s(f)} - a'(\check{X}_s)\sqrt{\check{\mu}_s(f)})\sigma dW_s$$

is a square integrable martingale since f and a' are bounded.

We wish to obtain a control on $|Z_t^*| := \sup_{s\leqslant t}|Z_s|$. We first take care of the jumps of $|Z_t|$. Notice first that, since f and a are bounded,

$$\Delta(x,y) := (f(x)\wedge f(y))|a(x)-a(y)| + |f(x)-f(y)|\left||a(0)-a(y)|+|a(0)-a(x)|\right| \leqslant C|a(x)-a(y)|,$$

implying that

$$\mathbb{E}\sup_{s\leqslant t}|\Delta_s| \leqslant C\mathbb{E}\int_0^t |a(\widehat{X}_s)-a(\check{X}_s)|ds \leqslant Ct\,\mathbb{E}|Z_t^*|.$$

Moreover, for a constant C depending on σ^2, $\|f\|_\infty, \|a\|_\infty, \|a'\|_\infty, \|a''\|_\infty$ and $\|b\|_\infty$,

$$\mathbb{E}\sup_{s\leqslant t}|A_s| \leqslant C\int_0^t \mathbb{E}|b(\widehat{X}_s)-b(\check{X}_s)|ds + C\int_0^t \mathbb{E}|a'(\widehat{X}_s)-a'(\check{X}_s)|ds + C\left[\int_0^t |a''(\widehat{X}_s)-a''(\check{X}_s)|ds + \int_0^t |\widehat{\mu}_s(f)-\check{\mu}_s(f)|ds\right].$$

We know that $|b(\widehat{X}_s)-b(\check{X}_s)| + |a'(\widehat{X}_s)-a'(\check{X}_s)| + |a''(\widehat{X}_s)-a''(\check{X}_s)| \leqslant C|a(\widehat{X}_s)-a(\check{X}_s)| = C|Z_s|$. Therefore,

$$\mathbb{E}\sup_{s\leqslant t}|A_s| \leqslant C\mathbb{E}\left[\int_0^t |Z_s|ds + \int_0^t |\widehat{\mu}_s(f)-\check{\mu}_s(f)|ds\right].$$

Moreover,

$$|\widehat{\mu}_s(f)-\check{\mu}_s(f)| = \left|\mathbb{E}\left(f(\widehat{X}_s)-f(\check{X}_s)|\mathcal{W}_s\right)\right| \leqslant \mathbb{E}\left(|f(\widehat{X}_s)-f(\check{X}_s)||\mathcal{W}_s\right) \leqslant \mathbb{E}(|Z_s||\mathcal{W}_s),$$

and thus,

$$\mathbb{E}\int_0^t |\widehat{\mu}_s(f)-\check{\mu}_s(f)|ds \leqslant \mathbb{E}\int_0^t |Z_s|ds \leqslant t\mathbb{E}|Z_t^*|.$$

Putting all these upper bounds together we conclude that for a constant C not depending on t,

$$\mathbb{E}\sup_{s\leqslant t}|A_s| \leqslant Ct\mathbb{E}|Z_t^*|.$$

Finally, we treat the martingale part using the Burkholder-Davis-Gundy inequality, and we obtain

$$\mathbb{E}\sup_{s\leqslant t}|M_s| \leqslant C\mathbb{E}\left[\left(\int_0^t (a'(\widehat{X}_s)\sqrt{\widehat{\mu}_s(f)}-a'(\check{X}_s)\sqrt{\check{\mu}_s(f)})^2 ds\right)^{1/2}\right].$$

But

$$(a'(\widehat{X}_s)\sqrt{\widehat{\mu}_s(f)}-a'(\check{X}_s)\sqrt{\check{\mu}_s(f)})^2 \leqslant C\left[((a'(\widehat{X}_s)-a'(\check{X}_s))^2 + (\sqrt{\widehat{\mu}_s(f)}-\sqrt{\check{\mu}_s(f)})^2\right] \leqslant C|Z_t^*|^2 + C(\sqrt{\widehat{\mu}_s(f)}-\sqrt{\check{\mu}_s(f)})^2, \quad (10)$$

where we have used once more that $|a'(x)-a'(y)| \leqslant C|a(x)-a(y)|$ and that f and a' are bounded.

Finally, since $\inf f > 0$,

$$|\sqrt{\widehat{\mu}_s(f)} - \sqrt{\check{\mu}_s(f)}|^2 \leqslant C|\widehat{\mu}_s(f) - \check{\mu}_s(f)|^2 \leqslant C\left(\mathbb{E}(|Z_s^*||\mathcal{W}_s)\right)^2.$$

We use that $(Z_t)_t$ is $(\mathcal{F}_t)_t-$adapted to obtain that $\mathbb{E}(|Z_s^*||\mathcal{W}_s) = \mathbb{E}(|Z_s^*||\mathcal{W}_t)$ for all $t \geqslant s$. Moreover, $|Z_s^*| \leqslant |Z_t^*|$, implying that $\mathbb{E}(|Z_s^*||\mathcal{W}_s) = \mathbb{E}(|Z_s^*||\mathcal{W}_t) \leqslant \mathbb{E}(|Z_t^*||\mathcal{W}_t)$. Therefore we obtain the upper bound

$$|\sqrt{\widehat{\mu}_s(f)} - \sqrt{\check{\mu}_s(f)}|^2 \leqslant C\left(\mathbb{E}(|Z_t^*||W)\right)^2$$

for all $s \leqslant t$, which implies the control of

$$\mathbb{E}\sup_{s\leqslant t}|M_s| \leqslant C\sqrt{t}\mathbb{E}|Z_t^*|.$$

The above upper bounds imply that, for a constant C not depending on t nor on the initial condition,

$$\mathbb{E}|Z_t^*| \leqslant C(t+\sqrt{t})\mathbb{E}|Z_t^*|,$$

and therefore, for t_1 sufficiently small, $\mathbb{E}|Z_{t_1}^*| = 0$. We can repeat this argument on intervals $[t_1, 2t_1]$, with initial condition $\hat{X}_{t_1}$, and iterate it up to any finite T because t_1 does only depend on the coefficients of the system but not on the initial condition. This implies the assertion. □

Corollary 1.2. *Grant Assumption 3 and suppose that $\int_{\mathbb{R}} x^2 d\nu_0(x) < +\infty$. Then the measure $\mu = P((\bar{X}_t)_{t\geqslant 0} \in \cdot|W)$ satisfies the following nonlinear stochastic PDE in weak form: for any $\varphi \in C_b^2(\mathbb{R})$, for any $t \geqslant 0$,*

$$\int_{\mathbb{R}} \varphi(x)\mu_t(dx) = \int_{\mathbb{R}} \varphi(x)\nu_0(dx) + \int_0^t \left(\int_{\mathbb{R}} \varphi'(x)\mu_s(dx)\right) \sqrt{\mu_s(f)}dW_s$$
$$+ \int_0^t \int_{\mathbb{R}} \Big([\varphi(0) - \varphi(x)]f(x) + \varphi'(x)b(x) + \frac{1}{2}\varphi''(x)\mu_s(f)\Big)\mu_s(dx)ds.$$

The proof of the above corollary is given in Appendix.

1.4. Convergence in distribution

The main results of this paper concern the convergence of the system $\left(X^{N,i}\right)_{1\leqslant i\leqslant N}$ to $\left(\bar{X}^i\right)_{i\geqslant 1}$. The first one proves that convergence in distribution holds. In order to state it, we need some additional integrability assumption on the measure ν.

Assumption 4. *We assume that $\int_{\mathbb{R}} e^{ax}\nu(dx) < \infty$ for all $|a| \leqslant a_0$ for some $a_0 > 0$.*

Theorem 1.3. *Grant Assumptions 2, 3 and 4. Then the sequence of processes $\left(X^{N,i}\right)_{i\in\mathbb{N}^*}$ converges to $\left(\bar{X}^i\right)_{i\in\mathbb{N}^*}$ in distribution in the space $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}$ endowed with the product topology, where $D(\mathbb{R}_+,\mathbb{R})$ is endowed with the topology of uniform convergence on every compact set.*

Theorem 1.3 is proved in Section 2.3. The second main result is a strong convergence result stated with respect to an appropriate L^1-norm, relying on an explicit coupling. To construct this coupling, we first introduce an auxiliary particle system.

Remark 1.4. *In Theorem 1.3, we implictly define $X^{N,i} = 0$ for every $i \geqslant N+1$.*

1.5. An auxiliary particle system and a strong approximation result

In what follows we exploit the old idea of time change that goes back at least to Kurtz (1978) and rewrite the evolution of the finite particle system in a different way. For that sake, we consider a standard Poisson process $\mathbf{N}_t$ of rate 1 and a family of i.i.d. variables $(U_n)_{n\geqslant 1}$ distributed according to ν, independent of everything else, as well as a family of i.i.d. variables $(V_n)_{n\geqslant 1}$ uniformly distributed on $[0,1]$, independent of the previous variables. We also define

$$\mathbf{Z}_t := \sum_{n=1}^{\mathbf{N}_t} (U_n, V_n) = (\mathbf{Z}_t^1, \mathbf{Z}_t^2) \tag{11}$$

which is a compound Poisson process. Notice that its first coordinate process $\mathbf{Z}_t^1$ is centred since $\mathbb{E}[U_n] = 0$.

Then, according to Theorem 7.4.I of Daley and Vere-Jones (2003), instead of writing the dynamics of $X^{N,i}, i = 1, \dots, N$, as solution of a SDE driven by N independent Poisson random measures as in (7) above, we rather describe their dynamic by solving a SDE driven by a time change of the compound Poisson process $\mathbf{Z}$. This leads to the following representation

$$X_t^{N,i} = X_0^{N,i} + \int_0^t b(X_s^{N,i})ds - \int_0^t X_{s-}^{N,i} dZ_s^{N,i} + \frac{1}{\sqrt{N}} \mathbf{Z}^1_{A_t^{N,X}} - \frac{1}{\sqrt{N}} R_t^{N,i}. \tag{12}$$

In the above equation, the random time change $A_t^{N,X}$ is given by

$$A_t^{N,X} = \sum_{j=1}^N \int_0^t f(X_s^{N,j})ds.$$

The counting processes $Z_t^{N,i}, 1 \leqslant i \leqslant N$, are defined by the classical thinning of $\mathbf{N}_{A_t^{N,X}}$ which represents the total number of jumps (spikes) during $[0,t]$. To define $Z^{N,i}$, each jump time t of $\mathbf{N}_{A_t^{N,X}}$ is accepted as jump of neuron i, that is, of $Z^{N,i}$, with probability

$$\frac{f(X_{t-}^{N,i})}{\sum_{j=1}^N f(X_{t-}^{N,j})}.$$

To realise these probabilities we use the uniform random variables V_n which are given by the second coordinate process $\mathbf{Z}_t^2$. More precisely, introducing for any $1 \leqslant i \leqslant N$ and $x \in \mathbb{R}^N$,

$$F_i(x) := \frac{\sum_{j=1}^i f(x_j)}{\sum_{j=1}^N f(x_j)}, F_0(x) := 0,$$

the process $Z_t^{N,i}$ is given by

$$Z_t^{N,i} = \int_0^t \mathbb{1}_{\left\{\mathbf{Z}^2_{A_s^{N,X}} \in [F_{i-1}(X_{s-}^N), F_i(X_{s-}^N)[\right\}} d\mathbf{N}_{A_s^{N,X}}$$

and the remainder terms by

$$R_t^{N,i} = \int_0^t \mathbf{Z}^1_{A_s^{N,X}} dZ_s^{N,i}.$$

It is straightforward to show that (12) defines the same dynamic as (7).

The important point is that we can couple the centred coordinate $\mathbf{Z}^1$ of the compound Poisson process $\mathbf{Z}$ with a Brownian motion. Indeed Corollary 7.5.5 of Ethier and Kurtz (2005), based on Komlós, Major and Tusnády (1976), gives the following

Lemma 1.5. *Grant Assumption 4. Then $\mathbf{Z}_t$ can be constructed on the same probability space as a standard one-dimensional Brownian motion B_t, such that*

$$\sup_{t\geqslant 0} \frac{|\mathbf{Z}_t^1 - \sigma B_t|}{\log t \vee 2} \leqslant K < \infty$$

almost surely, where K is a random variable having exponential moments, and $\sigma^2 = \mathbb{V}[U_1]$.

Applying the above result, we know that $\mathbf{Z}^1_{A_t^{N,X}}$ behaves, for large N, as $\sigma B_{A_t^{N,X}}$. The process $B_{A_t^{N,X}}$ can be written as

$$B_{A_t^{N,X}} = \int_0^t \sqrt{\sum_{j=1}^N f\left(X_s^{N,j}\right)} dW_s^N, \tag{13}$$

where W^N is another one-dimensional standard Brownian motion.

Therefore, we will be able to show that, for N large enough, $(X^{N,1}, \dots X^{N,N})$ behaves as the auxiliary process $(\tilde{X}^{N,1}, \dots \tilde{X}^{N,N})$ where

$$\begin{cases} \tilde{X}_t^{N,i} &= X_0^{N,i} + \int_0^t b(\tilde{X}_s^{N,i})ds - \int_0^t \tilde{X}_{s-}^{N,i}\, d\tilde{Z}_s^{N,i} + \sigma \int_0^t \sqrt{\frac{1}{N}\sum_{j=1}^N f(\tilde{X}_s^{N,j})} dW_s^N, \\ \tilde{X}_0^{N,i} &\sim \nu_0, \end{cases} \tag{14}$$

and where $d\tilde{Z}_s^{N,i}$ has compensator $f(\tilde{X}_s^{N,i})ds$.

The well-posedness of (14) holds true under Assumptions 1 and 2 if we suppose moreover that $\inf f > 0$. This can be proved with the same reasoning as for (7), using Theorem IV.9.1 of Ikeda and Watanabe (1989).

Obviously, (14) is a mean field particle version of the limit system (8), constructed with a particular choice of underlying Brownian motion. In the following, we denote by $\bar{X}^N$ the strong solution of the system (8) defined with respect to the Brownian W^N. Moreover we will denote by $\bar{X}$ any solution of the system (8) defined for some Brownian W that does not depend on N. We can now state the second main result of this paper.

Theorem 1.6. *If Assumptions 2, 3 and 4 hold, then, for each $N \in \mathbb{N}^*$, there exists a one-dimensional standard Brownian motion W^N such that, for every $t > 0, i \leqslant N$,*

$$\mathbb{E}\left[\sup_{0\leqslant s\leqslant t} \left|a\left(X_s^{N,i}\right) - a\left(\bar{X}_s^{N,i}\right)\right|\right] \leqslant C_t \left(w_a\left(\frac{(\ln N)^{1/2}}{N^{1/4}}\right) + \frac{1}{\sqrt{N}}\right), \tag{15}$$

where a is the function given in Assumption 3, w_a its modulus of continuity, and $\left(\bar{X}^{N,i}\right)_{1\leqslant i\leqslant N}$ is the solution of (8) with respect to the Brownian motion W^N and the initial condition $\bar{X}_0^{N,i} = X_0^{N,i}$.

Remark 1.7. *Let us emphasise the fact that the expression in (15) vanishes as N goes to infinity. Indeed, under Assumption 3, the function a is Lipschitz continuous, so its modulus of continuity vanishes.*

Theorem 1.6 is proved in Section 2.2. A consequence of the above result is the following

Proposition 1.8. *Grant the assumptions of Theorem 1.6. If $a'(x) \neq 0$ for all $x \in \mathbb{R}$, then for each $t \geqslant 0$, the sequence of empirical measures $\mu^N = N^{-1}\sum_{i=1}^N \delta_{(X_t^{N,i})_{t\geqslant 0}}$ converges in distribution in $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$ to $\mu = P((\bar{X}_t)_{t\geqslant 0} \in \cdot|W)$.*

The proof of Proposition 1.8 is postponed to Appendix.

2. Proofs of the main results

2.1. Useful properties of the limit system

In the proof of Theorem 1.6, we use an important property of the limit system (8), which is the conditional independence of the processes $\bar{X}^i$ $(i \geqslant 1)$ given the Brownian motion W.

Proposition 2.1. *If Assumption 3 holds and $\int_{\mathbb{R}} x^2 d\nu_0(x) < +\infty$, then*

(i) for all $N \in \mathbb{N}^*$ there exists a strong solution $\left(\bar{X}^i\right)_{1\leqslant i\leqslant N}$ of (8), and pathwise uniqueness holds,
(ii) $\bar{X}^1, \dots, \bar{X}^N$ are independent conditionally to W,
(iii) for all $t \geqslant 0$, almost surely, the weak limit of $\frac{1}{N}\sum_{i=1}^N \delta_{\bar{X}^i_{|[0,t]}}$ is given by $\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N \delta_{\bar{X}^i_{|[0,t]}} = P(\bar{X}^i_{|[0,t]} \in \cdot|\mathcal{W}_t) = P(\bar{X}^i_{|[0,t]} \in \cdot|W)$.

The proof of Proposition 2.1 is postponed to Appendix, in Section 3.1.

2.2. Strong convergence

We prove the convergence of the finite system (7) to the limit system (8), by controlling the distance between these systems and the auxiliary system (14). This is done by introducing a suitable coupling between (7) and (14). .

Proposition 2.2. *Suppose that $X_0^{N,i} = \tilde{X}_0^{N,i}$, for all $1 \leqslant i \leqslant N$. Grant Assumptions 2, 3 and 4. Then there exists a coupling of $X^{N,i}$ and $\tilde{X}^{N,i}$ such that for all $t > 0$, for all $i = 1, \dots, N$,*

$$\mathbb{E}\left[\sup_{0\leqslant s\leqslant t}\left|\tilde{X}_s^{N,i} - X_s^{N,i}\right|\right] \leqslant C_t \frac{(\ln N)^{1/2}}{N^{1/4}}.$$

Proof. By exchangeability, it suffices to prove the result for $i = 1$. We couple the two processes by using the KMT approximation of Lemma 1.5 and then using a total variation coupling of the two jump processes $Z^{N,1}$ and $\tilde{Z}^{N,1}$.
Step 1. Construction of the coupling. We construct the initial process $X^N = (X^{N,1}, \dots, X^{N,N})$ driven by the underlying compound Poisson process $\mathbf{Z}_t$ as in (12). Then we couple $\mathbf{Z}$ with the Brownian motion B according to Lemma 1.5 and thus, by time change, with the Brownian motion W^N of (13). Therefore, in what follows, we shall work with the filtration

$$\mathcal{F}_t^N = \mathcal{F}^{\mathbf{Z}}_{A_t^{N,X}} \vee \sigma\{W_s^N, s \leqslant t\}, \tag{16}$$

where $\mathcal{F}_t^{\mathbf{Z}}$ is the natural filtration of the compound Poisson process $\mathbf{Z}$.

To construct the total variation coupling of $Z^{N,1}$ and $\widetilde{Z}^{N,1}$, we complete the jumps of $Z_t^{N,1}$, using the construction of Lemma 4 of Brémaud and Massoulié (1996), to a Poisson random measure $\pi^1(dt,dz)$ on $\mathbb{R}_+ \times \mathbb{R}_+$ having intensity $dtdz$. This PRM π^1 depends on $Z_t^{N,1}$, by copying all of its points T_n, adding to them a random mark z which is placed uniformly on the strip $z \in [0, f(X_{T_n-}^{N,1})[$, independently of anything else. Finally, we add independent PRM marks on the missing domain $\{(t,z) \in \mathbb{R}_+^2 : z \geqslant f(X_{t-}^{N,1})\}$. Notice that the PRM π^1 depends on $Z^{N,1}$, and thus on the compound Poisson process $\mathbf{Z}$ of (11). We use the same construction for all other coordinates $i > 1$, using the same underlying $\mathbf{Z}$ and independent PRM's on the missing domains.

We are now able to define the dynamics of $\widetilde{X}^{N,1}$, coupled to $X^{N,1}$, by

$$\left\{ \begin{array}{rcl} \widetilde{X}_t^{N,1} & = & X_0^{N,1} + \displaystyle\int_0^t b(\widetilde{X}_s^{N,1})ds - \int_{[0,t]\times\mathbb{R}_+} \widetilde{X}_{s-}^{N,1} \mathbb{1}_{\{z \leqslant f(\widetilde{X}_{s-}^{N,1})\}} d\pi^1(s,z) \\ & & +\sigma \displaystyle\int_0^t \sqrt{\frac{1}{N}\sum_{j=1}^N f(\widetilde{X}_s^{N,j})} dW_s^N. \end{array} \right. \tag{17}$$

Using this construction of π^1 guarantees that $X^{N,1}$ and $\widetilde{X}^{N,1}$ have a maximal number of common jumps.

Step 2.

Let $(M_N)_N$ be an increasing sequence of positive numbers that goes to infinity. Then

$$\begin{aligned} \left|\widetilde{X}_t^{N,1} - X_t^{N,1}\right| \leqslant & \int_0^t \left| b\left(\widetilde{X}_s^{N,1}\right) - b\left(X_s^{N,1}\right)\right| ds \\ & + \int_{]0,t]\times\mathbb{R}_+} \left| \widetilde{X}_s^{N,1} \mathbb{1}_{\{z \leqslant f(\widetilde{X}_s^{N,1})\}} - X_s^{N,1} \mathbb{1}_{\{z \leqslant f(X_s^{N,1})\}} \right| d\pi^1(s,z) \\ & + \frac{\sigma}{\sqrt{N}} \left| \int_0^t \sqrt{\sum_{j=1}^N f(\widetilde{X}_s^{N,j})} dW_s^N - B_{A_t^{N,X}} \right| \\ & + \frac{1}{\sqrt{N}} \left| \sigma B_{A_t^{N,X}} - \mathbf{Z}^1_{A_t^{N,X}} \right| + \frac{1}{\sqrt{N}} |R_t^{N,1}|. \end{aligned}$$

Let

$$u_t^N := \mathbb{E}\left[\sup_{0 \leqslant s \leqslant t} \left| \widetilde{X}_s^{N,1} - X_s^{N,1} \right| \right];$$

note that, for all $t \geqslant 0, u_t^N < +\infty$ thanks to the points (ii) and (iv) of Lemma 3.1.

Then, using that the expectation of the jump term

$$\int_{]0,t]\times\mathbb{R}_+} \left| \widetilde{X}_s^{N,1} \mathbb{1}_{\{z \leqslant f(\widetilde{X}_s^{N,1})\}} - X_s^{N,1} \mathbb{1}_{\{z \leqslant f(X_s^{N,1})\}} \right| d\pi^1(s,z)$$

above is bounded by

$$\mathbb{E}\left[||f||_\infty \int_0^t |\widetilde{X}_s^{N,1} - X_s^{N,1}| ds \right] + \int_0^t \mathbb{E}\left[|f(\widetilde{X}_s^{N,1}) - f(X_s^{N,1})| (|\widetilde{X}_s^{N,1}| + |X_s^{N,1}|) \right] ds, \tag{18}$$

and inserting

$$1 = \mathbb{1}_{\{|\widetilde{X}_s^{N,1}| + |X_s^{N,1}| \leqslant M_N\}} + \mathbb{1}_{\{|\widetilde{X}_s^{N,1}| + |X_s^{N,1}| > M_N\}}$$

in the second term of (18), we obtain (with Markov inequality and points (i) and (iii) of Lemma 3.1)

$$u_t^N \leqslant Ctu_t^N + C(1+M_N)tu_t^N + \frac{C_T}{M_N^2} + \sigma\mathbb{E}\left[\left(\int_0^t \left(\sqrt{\frac{1}{N}\sum_{j=1}^N f(\tilde{X}_s^{N,j})} - \sqrt{\frac{1}{N}\sum_{j=1}^N f(X_s^{N,j})}\right)^2 ds\right)^{1/2}\right] + K\frac{\ln N}{\sqrt{N}},$$

for some constants $C, K > 0$. Here, we have used Lemma 1.5. Hence, using that $\inf f > 0$, for all $t \in [0,T]$,

$$u_t^N \leqslant C(t+M_N t)u_t^N + \frac{C_T}{M_N^2} + K\frac{\ln N}{\sqrt{N}} + C\sigma\mathbb{E}\left[\left(\int_0^t \left(\frac{1}{N}\sum_{j=1}^N \left|\tilde{X}_s^{N,j} - X_s^{N,j}\right|\right)^2 ds\right)^{1/2}\right].$$

Now, introducing $v_t^{N,j} = \sup_{0\leqslant s\leqslant t} |\tilde{X}_s^{N,j} - X_s^{N,j}|$ and using that $\mathbb{E}\left[v_t^{N,j}\right] = u_t^N$ for all j, we have

$$\begin{aligned} u_t^N \leqslant & C(t+M_N t)u_t^N + \frac{C_T}{M_N^2} + K\frac{\ln N}{\sqrt{N}} + C\sqrt{t}\frac{1}{N}\sum_{j=1}^N \mathbb{E}\left[v_t^{N,j}\right] \\ \leqslant & C(t+M_N t+\sqrt{t})u_t^N + \frac{C_T}{M_N^2} + K\frac{\ln N}{\sqrt{N}}. \end{aligned}$$

Choose now $t_N = \frac{1}{16M_N C^2}$, such that (assuming $C \geqslant 1$ and $M_N \geqslant 1$)

$$C(t_N + M_N t_N + \sqrt{t_N}) \leqslant C\left(\frac{1}{16C} + \frac{1}{16C} + \frac{1}{4C}\right) \leqslant 1/2.$$

Then for all $t \leqslant t_N \leqslant T$,

$$u_t^N \leqslant 2\frac{C_T}{M_N^2} + 2K\frac{\ln N}{\sqrt{N}}.$$

The above argument can be iterated such that for each $n \in \mathbb{N}^*$, for all $t \leqslant nt_N \leqslant T$,

$$u_t^N \leqslant 2n\left(\frac{C_T}{M_N^2} + K\frac{\ln N}{\sqrt{N}}\right),$$

which implies in turn that for all $T > 0$, for all $t \in [0,T]$,

$$u_t^N \leqslant 2\lceil T/t_N\rceil\left(\frac{C_T}{M_N^2} + K\frac{\ln N}{\sqrt{N}}\right) \leqslant K_T\left(\frac{1}{M_N} + \frac{(\ln N)M_N}{\sqrt{N}}\right).$$

Finally, choosing $M_N = N^{1/4}/(\ln N)^{1/2}$ proves the proposition. □

Now we control the distance between the auxiliary system and the limit system. For that sake we construct the auxiliary system and the limit system using the same Poisson random measures $\pi^i(ds,dz)$ as those used in (17). Our argument relies on the conditional independence of the coordinates of the limit system.

Proposition 2.3. *Grant Assumptions 2 and 3, and suppose that $\tilde{X}_0^{N,i} = \bar{X}_0^{N,i}$ for all i. Then for all $s \leqslant t$,*

$$\mathbb{E}\sup_{s\leqslant t}|a(\tilde{X}_s^{N,1}) - a(\bar{X}_s^{N,1})| \leqslant C_t N^{-1/2}.$$

Proof. The proof is done by decomposing the evolution of $a(\bar{X}_t^{N,1})$ in the following way.

$$a(\bar{X}_t^{N,1}) = a(\bar{X}_0^{N,1}) + \int_0^t a'(\bar{X}_s^{N,1})b(\bar{X}_s^{N,1})ds + \int_{[0,t]\times\mathbb{R}_+}\left(a(0) - a(\bar{X}_{s-}^{N,1})\right)\mathbb{1}_{\{z\leqslant f(\bar{X}_{s-}^{N,1})\}}\pi^1(ds,dz)$$
$$+ \frac{\sigma^2}{2}\int_0^t a''(\bar{X}_s^{N,1})\frac{1}{N}\sum_{j=1}^N f(\bar{X}_s^{N,j})ds - B_t^N + \sigma\int_0^t a'(\bar{X}_s^{N,1})\sqrt{\frac{1}{N}\sum_{j=1}^N f(\bar{X}_s^{N,j})}dW_s^N - M_t^N,$$

where

$$B_t^N = \frac{\sigma^2}{2}\int_0^t a''(\bar{X}_s^{N,1})\left(\frac{1}{N}\sum_{j=1}^N f(\bar{X}_s^{N,j}) - \mathbb{E}\left[f(\bar{X}_s^{N,1})\middle| W^N\right]\right)ds,$$

and

$$M_t^N = \sigma\int_0^t a'(\bar{X}_s^{N,1})\left(\sqrt{\frac{1}{N}\sum_{j=1}^N f(\bar{X}_s^{N,j})} - \sqrt{\mathbb{E}\left[f(\bar{X}_s^{N,1})|W^N\right]}\right)dW_s^N.$$

Since

$$< M^N >_t \leqslant \sigma^2\left(\sup_{x\in\mathbb{R}}|a'(x)^2|\right)\int_0^t\left(\sqrt{\frac{1}{N}\sum_{j=1}^N f(\bar{X}_s^{N,j})} - \sqrt{\mathbb{E}\left[f(\bar{X}_s^{N,1})|W^N\right]}\right)^2 ds,$$

recalling that the variables $\bar{X}_s^{N,j}$ $(1 \leqslant j \leqslant N)$ are i.i.d. conditionally to W^N (see Proposition 2.1), taking conditional expectation $\mathbb{E}(\cdot|W^N)$ implies that

$$\mathbb{E}\left[< M^N >_t\right] \leqslant C_t N^{-1} \text{ and } \mathbb{E}\left[B_t^N\right] \leqslant C_t N^{-1},$$

and this implies the result, with the same reasoning as in the proof of Theorem 1.1. □

We conclude with the

Proof of Theorem 1.6. The result is now a straightforward consequence of Propositions 2.2 and 2.3. □

2.3. Weak convergence

Now we prove the convergence in distribution of the finite system (7) to the limit system (8) in the topology of the uniform convergence on every compact set.

Lemma 2.4. *Let $(a_i)_{i\in\mathbb{N}}$ be a sequence of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}$. Then the function Φ defined as*

$$\Phi : (x_i)_i \in D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}} \mapsto (a_i \circ x_i)_i \in D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}}$$

is continuous, where $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}}$ is endowed with the product topology with respect to the uniform convergence on every compact set.

Proof. Let $\left((x_i^n)_{i\in\mathbb{N}}\right)_n$ be a sequence that converges to some $(x_i)_i$. This means that, for every i, x_i^n converges to x_i in $D(\mathbb{R}_+,\mathbb{R})$ uniformly on every interval $[0,T]$ $(T>0)$.

Consequently, for every i, $a_i\circ x_i^n$ converges to $a_i\circ x_i$ uniformly on every compact set. This implies the convergence of $\Phi\left((x_i^n)_i\right)$ to $\Phi\left((x_i)_i\right)$. □

Proof of Theorem 1.3. Let X^N and $\bar{X}^N$ be defined as in Theorem 1.6. We define a metric d_U on $D(\mathbb{R}_+,\mathbb{R})$ that defines the topology of the uniform convergence on every compact set by

$$d_U(x,y) := \sum_{n=1}^{+\infty}\frac{1}{2^n}\left(1\wedge\sup_{0\leqslant s\leqslant n}|x(s)-y(s)|\right).$$

Then we define the following metric d on $D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}$, that defines the product topology with respect to d_U

$$d((x_i)_i,(y_i)_i) := \sum_{i=1}^{+\infty}\frac{1}{2^i}\,d_U(x_i,y_i).$$

Now we prove that $\left(a\circ X^{N,i}\right)_i$ converges to $\left(a\circ\bar{X}^i\right)_i$ in distribution in the topology of d (that is the product topology with respect to the topology of the uniform convergence on the compact sets), where $\left(\bar{X}^i\right)_{i\in\mathbb{N}^*}$ is a solution of (8) for any Brownian motion W that does not depend on N.

Let $g : D(\mathbb{R}_+,\mathbb{R})^{\mathbb{N}^*}\to\mathbb{R}$ be any bounded and uniformly continuous function. We want to prove that $\mathbb{E}\left[g\left(\Phi(X^N)\right)\right]$ converges to $\mathbb{E}\left[g\left(\Phi(\bar{X})\right)\right]\left(=\mathbb{E}\left[g\left(\Phi(\bar{X}^N)\right)\right]\right)$ as N goes to infinity, where $\Phi((x_i)_i) := (a\circ x_i)_i$.

Then we have,

$$\begin{aligned}\left|\mathbb{E}\left[g\left(\Phi(X^N)\right)\right]-\mathbb{E}\left[g\left(\Phi(\bar{X})\right)\right]\right| &\leqslant \mathbb{E}\left[\left|g(\Phi(X^N))-g(\Phi(\bar{X}^N))\right|\right]\\ &\leqslant \mathbb{E}\left[w_g\left(d\left(\Phi(X^N),\Phi(\bar{X}^N)\right)\right)\right],\end{aligned}\tag{19}$$

where w_g is the modulus of continuity of g (with respect to the metric d).

Thanks to Theorem 1.6, for any increasing function $\varphi:\mathbb{N}^*\to\mathbb{N}^*$, there exists another one ψ such that $d\left(\Phi(X^{\varphi(\psi(N))}),\Phi(\bar{X}^{\varphi(\psi(N))})\right)$ vanishes almost surely as N goes to infinity. Then, using (19), $\mathbb{E}\left[g\left(\Phi(X^{\varphi(\psi(N))})\right)\right]$ converges to $\mathbb{E}\left[g\left(\Phi(\bar{X})\right)\right]$. This proves that,

$$\mathbb{E}\left[g\left(\Phi(X^N)\right)\right]\underset{N\to\infty}{\longrightarrow}\mathbb{E}\left[g\left(\Phi(\bar{X})\right)\right].$$

As the previous convergence holds for any bounded and uniformly continuous function g, we know, by Portmanteau theorem (see Theorem 2.1 of Billingsley (1999)), that $\left(a\circ X^{N,i}\right)_i$ converges to $\left(a\circ\bar{X}^i\right)_i$ in distribution in the product topology with respect to the uniform convergence on every compact set, as N goes to infinity.

Then, applying Lemma 2.4 with $a_i := a^{-1}$ that is continuous, we obtain the result. □

3. Appendix

3.1. Properties of the limit system

We start with the

Proof of Item 2. of Theorem 1.1. The proof is done using a classical Picard-iteration. For that sake we introduce the sequence of processes $\bar{X}_t^{[0]} \equiv \bar{X}_0$, and

$$\bar{X}_t^{[n+1]} := \bar{X}_0 + \int_0^t b(\bar{X}_s^{[n]})ds - \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \bar{X}_{s-}^{[n+1]} \mathbb{1}_{\{z\leqslant f(\bar{X}_{s-}^{[n]})\}} \pi(ds,dz,du) + \sigma \int_0^t \sqrt{\mu_s^n(f)} dW_s,$$

where

$$\mu_s^n = P(\bar{X}_s^{[n]} \in \cdot | \mathcal{W}_s).$$

Let us first prove a control on the moments of $\bar{X}^{[n]}$ uniformly in n. We define, for each $k \in \mathbb{N}^*$,

$$\tau_k := \inf_{n\in\mathbb{N}} \inf\{t > 0 \ : \ |\bar{X}_t^{[n]}| > k(n+1)\}.$$

Applying Ito's formula we have

$$\mathbb{E}\left[\left(\bar{X}_{t\wedge\tau_k}^{[n+1]}\right)^2\right] \leqslant \mathbb{E}\left[\bar{X}_0^2\right] + 2\int_0^t \mathbb{E}\left[\bar{X}_{s\wedge\tau_k}^{[n+1]} b\left(\bar{X}_{s\wedge\tau_k}^{[n]}\right)\right] ds + \sigma^2 \int_0^t \mathbb{E}\left[\mu_{s\wedge\tau_k}^n(f)\right] ds.$$

Using that f and b are bounded, we have

$$u_t^{[n+1]} := \mathbb{E}\left[\left(\bar{X}_{t\wedge\tau_k}^{[n+1]}\right)^2\right] \leqslant C(1+t) + C\int_0^t u_s^{[n+1]} ds.$$

Then, by Grönwall's lemma, we know that, for all $t > 0$

$$\sup_{k\in\mathbb{N}} \sup_{n\in\mathbb{N}} \sup_{0\leqslant s\leqslant t} \mathbb{E}\left[\left(\bar{X}_{s\wedge\tau_k}^{[n]}\right)^2\right] < +\infty.$$

Besides, $(\tau_k)_k$ is nondecreasing, so it converges almost surely to some τ, which is almost surely infinite since

$$\begin{aligned} P(\tau \leqslant t) = \lim_{k\to\infty} P(\tau_k \leqslant t) &\leqslant \lim_{k\to\infty} P\left(\exists n \in \mathbb{N}, \left|\bar{X}_{t\wedge\tau_k}^{[n]}\right| \geqslant k(n+1)\right) \\ &\leqslant \lim_{k\to\infty} \sum_{n\in\mathbb{N}} P\left(\left|\bar{X}_{t\wedge\tau_k}^{[n]}\right| \geqslant k(n+1)\right) \\ &\leqslant C_t \lim_{k\to\infty} \sum_{n\in\mathbb{N}} \frac{1}{k^2(n+1)^2} \leqslant \lim_{k\to\infty} \frac{C_t}{k^2} = 0. \end{aligned}$$

Then, by Fatou's lemma, we know that

$$\sup_{n\in\mathbb{N}} \sup_{0\leqslant s\leqslant t} \mathbb{E}\left[\left(\bar{X}_s^{[n]}\right)^2\right] < +\infty. \tag{20}$$

Now, we prove the convergence of $\bar{X}_t^{[n]}$. The same strategy as the one of the proof of Item 1. of Theorem 1.1 allows to show that

$$\delta_t^n := \mathbb{E}\sup_{s\leqslant t} |a(\bar{X}_s^{[n]}) - a(\bar{X}_s^{[n-1]})|$$

satisfies

$$\delta_t^n \leqslant C(t+\sqrt{t})\delta_t^{n-1},$$

for all $n \geqslant 1$, for a constant C only depending on the parameters of the model, but not on n, neither on t. Choose t_1 such that

$$C(t_1 + \sqrt{t_1}) \leqslant \frac{1}{2}.$$

Since $\sup_{s \leqslant t_1} |a(\bar{X}_s^{[0]})| = a(\bar{X}_0) \leqslant \|a\|_\infty$, we deduce from this that

$$\delta_{t_1}^n \leqslant 2^{-n} \|a\|_\infty.$$

This implies the almost sure convergence of $a\left(\bar{X}_t^{[n]}\right)_n$ to some random variable Z_t for all $t \in [0, t_1]$. As a is an increasing function, this implies the almost sure convergence of $\bar{X}_t^{[n]}$ to some (possibly infinite) random variable $\bar{X}_t$. The almost sure finiteness of $\bar{X}_t$ is then guaranteed by Fatou's lemma and (20).

It remains to prove that $\bar{X}$ is solution of the limit equation (9) which follows by standard arguments (note that the jump term does not cause troubles because it is of finite activity). The most important point is to notice that

$$\mu_t^n(f) = \mathbb{E}(f(\bar{X}_t^{[n]})|\mathcal{W}_t) \to \mathbb{E}(f(\bar{X}_t)|\mathcal{W}_t)$$

almost surely, which follows from the almost sure convergence of $f(\bar{X}_t^{[n]}) \to f(\bar{X}_t)$, using dominated convergence.

Finally, once the convergence is proven on the time interval $[0, t_1]$, we can proceed iteratively over successive intervals $[kt_1, (k+1)t_1]$ to conclude the proof. □

We just proved existence and uniqueness of strong solution of the SDE (9). In the paper, we also need to know some properties about the joint distribution of the limit system given by (8), not only each of its coordinate individually.

Proof of Proposition 2.1. (i) Given a Brownian motion W and i.i.d. Poisson measures π^i, the same proof as the one of Theorem 1.1 implies the existence and the uniqueness of the system given in (8) for $1 \leqslant i \leqslant N$.

(ii) The construction of the proof of Item 2. of Theorem 1.1, together with the proof of Theorem 1.1 of Chapter IV.1 and of Theorem 9.1 in Chapter IV.9 of Ikeda and Watanabe (1989), imply the existence of a measurable function Φ that does not depend on $k = 1, \dots, N$, and that satisfies, for each $1 \leqslant k \leqslant N$,

$$\bar{X}^k = \Phi(\bar{X}_0^k, \pi^k, W)$$

and for all $t \geqslant 0$,

$$\bar{X}^k_{|[0,t]} = \Phi_t(\bar{X}_0^k, \pi^k_{|[0,t]\times\mathbb{R}_+\times\mathbb{R}}, (W_s)_{s\leqslant t}); \tag{21}$$

in other words, our process is non-anticipative and does only depend on the underlying noise up to time t.

Then we can write, for all continuous bounded functions g, h,

$$\mathbb{E}\left[g(\bar{X}^i)h(\bar{X}^j)\middle| W\right] = \mathbb{E}\left[g(\Phi(\bar{X}_0^i, \pi^i, W))h(\Phi(\bar{X}_0^j, \pi^j, W))\middle| W\right] = \psi(W),$$

where $\psi(w) := \mathbb{E}\left[g(\Phi(\bar{X}_0^i, \pi^i, w))h(\Phi(\bar{X}_0^j, \pi^j, w))\right] = \mathbb{E}\left[g(\Phi(\bar{X}_0^i, \pi^i, w))\right]\mathbb{E}\left[h(\Phi(\bar{X}_0^j, \pi^j, w))\right] =: \psi_i(w)\psi_j(w)$. With the same reasoning, we show that $\mathbb{E}\left[g(\bar{X}^i)\middle| W\right] = \psi_i(W)$ and $\mathbb{E}\left[h(\bar{X}^j)\middle| W\right] = \psi_j(W)$. The same arguments prove the mutual independence of $\bar{X}^1, \dots \bar{X}^N$ conditionally to W.

(iii) Using the representation $\bar{X}^k_{|[0,t]} = \Phi_t(\bar{X}^k_0, \pi^k, W)$, we can write for any continuous and bounded function $g : D([0,t],\mathbb{R}) \to \mathbb{R}$,

$$\int_{\mathbb{R}} g d(N^{-1}\sum_{i=1}^N \delta_{\bar{X}^i_{|[0,t]}}) = \frac{1}{N}\sum_{i=1}^N g(\bar{X}^i_{|[0,t]}) = \frac{1}{N}\sum_{i=1}^N g\circ\Phi_t(\bar{X}^i_0, \pi^i, W).$$

Using the law of large numbers on the account of the sequence of i.i.d. PRM's and working conditionally on W, we obtain that

$$\lim_{N\to\infty}\int_{\mathbb{R}} g d(N^{-1}\sum_{i=1}^N \delta_{\bar{X}^i_{|[0,t]}}) = \mathbb{E}\left[g\circ\Phi_t(\bar{X}^1_0, \pi^1, W)|W\right] = \mathbb{E}\left[g(\bar{X}^1_{|[0,t]})|W\right] = \mathbb{E}\left[g(\bar{X}^1_{|[0,t]})|(W_s)_{s\leqslant t}\right],$$

where we have used (21). □

3.2. Proof of Proposition 1.8

For $m = \mathcal{L}(X) \in \mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$, for every $t_1,\dots,t_k \in \mathbb{R}_+$, let $m_{(t_1,\dots,t_k)} = \mathcal{L}(X_{t_1},\dots,X_{t_k})$, and $\pi_{(t_1,\dots,t_k)}(m) = m_{(t_1,\dots,t_k)}$. One can note that $\pi_{(t_1,\dots,t_k)}$ is continuous on $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$.

Step 1. In a first time, we prove the convergence in distribution of $\mu^N_{(t_1,\dots,t_k)}$ to $\mu_{(t_1,\dots,t_k)}$ for any $0 \leqslant t_1 \leqslant \dots \leqslant t_k$. For this purpose, let us consider the algebra $\mathcal{M}$ composed of the functions Φ of the form

$$\Phi : m \in \mathcal{P}(\mathbb{R}^k) \longmapsto h\left(\int \lambda_1 dm, \dots, \int \lambda_r dm\right), \tag{22}$$

where $h : \mathbb{R}^r \to \mathbb{R}$ is Lipschitz continuous and bounded, and $\lambda_i \in C_b(\mathbb{R}^k)$ satisfies, for every $x_1,\dots,x_k,y_1,\dots,y_k \in \mathbb{R}$,

$$|\lambda_i(x_1,\dots,x_k) - \lambda_i(y_1,\dots,y_k)| \leqslant C\sum_{j=1}^k |a(x_j) - a(y_j)|.$$

Let us prove that, for all $\Phi \in \mathcal{M}$,

$$\mathbb{E}\left[\Phi(\mu^N_{(t_1,\dots,t_k)})\right] \underset{N\to\infty}{\longrightarrow} \mathbb{E}\left[\Phi(\mu_{(t_1,\dots,t_k)})\right]. \tag{23}$$

For Φ in the form (22), we have

$$\begin{aligned}
\left|\mathbb{E}\left[\Phi(\mu^N_{(t_1,\dots,t_k)})\right] - \mathbb{E}\left[\Phi(\mu_{(t_1,\dots,t_k)})\right]\right| &= \left|\mathbb{E}\left[\Phi(\mu^N_{(t_1,\dots,t_k)})\right] - \mathbb{E}\left[\Phi(\mathcal{L}(\bar{X}^{N,1}_{t_1},\dots,\bar{X}^{N,1}_{t_k}|W^N))\right]\right| \\
&\leqslant \mathbb{E}\left[\left|h\left(\frac{1}{N}\sum_{j=1}^N \lambda_1(X^{N,j}_{t_1},\dots,X^{N,j}_{t_k}),\dots,\frac{1}{N}\sum_{j=1}^N \lambda_r(X^{N,j}_{t_1},\dots,X^{N,j}_{t_k})\right)\right.\right. \\
&\quad \left.\left. -h\left(\frac{1}{N}\sum_{j=1}^N \lambda_1(\bar{X}^{N,j}_{t_1},\dots,\bar{X}^{N,j}_{t_k}),\dots,\frac{1}{N}\sum_{j=1}^N \lambda_r(\bar{X}^{N,j}_{t_1},\dots,\bar{X}^{N,j}_{t_k})\right)\right|\right] \\
&\quad + \mathbb{E}\left[\left|h\left(\frac{1}{N}\sum_{j=1}^N \lambda_1(\bar{X}^{N,j}_{t_1},\dots,\bar{X}^{N,j}_{t_k}),\dots,\frac{1}{N}\sum_{j=1}^N \lambda_r(\bar{X}^{N,j}_{t_1},\dots,\bar{X}^{N,j}_{t_k})\right)\right.\right.
\end{aligned}$$

$$-h\left(\mathbb{E}\left[\lambda_1(\bar{X}^{N,1}_{t_1},\dots,\bar{X}^{N,1}_{t_k})\middle| W^N\right],\dots,\mathbb{E}\left[\lambda_r(\bar{X}^{N,1}_{t_1},\dots,\bar{X}^{N,1}_{t_k})\middle| W^N\right]\right)\Big|\Big]$$
$$\leqslant C\sum_{i=1}^{k}\mathbb{E}\left[\left|a\left(X^{N,1}_{t_i}\right)-a\left(\bar{X}^{N,1}_{t_i}\right)\right|\right]$$
$$+C\sum_{i=1}^{r}\mathbb{E}\left[\left|\frac{1}{N}\sum_{j=1}^{N}\lambda_i(\bar{X}^{N,j}_{t_1},\dots,\bar{X}^{N,j}_{t_k})-\mathbb{E}\left[\lambda_i(\bar{X}^{N,j}_{t_1},\dots,\bar{X}^{N,j}_{t_k})\middle| W^N\right]\right|\right].$$

Then, Theorem 1.6 and Proposition 2.1.(iii) imply (23).

Besides, the sequence $\left(\mu^N_{(t_1,\dots,t_k)}\right)_N$ is tight. Indeed, it is well-known that this is equivalent to the tightness of the sequence $\left(X^{N,1}_{t_1},\dots,X^{N,1}_{t_k}\right)_N$ (see Proposition 2.2 of Sznitman (1991)), and this is a mere consequence of Theorem 1.3.

Now, to obtain the convergence in distribution of $\mu^N_{(t_1,\dots,t_k)}$ to $\mu_{(t_1,\dots,t_k)}$, it is sufficient to show that the algebra $\mathcal{M}$ separates the points. Indeed, if this is the case, Theorem 3.4.5.(a) and Lemma 3.4.3 of Ethier and Kurtz (2005) imply the result.

Let m, m' be two distinct probabilities on $\mathbb{R}^k$. There exist $\alpha_i < \beta_i$ $(1 \leqslant i \leqslant k)$ such that $m(C) \neq m'(C)$ with $C = \prod_{i=1}^k[\alpha_i,\beta_i]$. Let us assume that $m(C) > m'(C)$. This implies the existence of some $\delta > 0$ that satisfies $m(\prod_{i=1}^k[\alpha_i,\beta_i]) > m'(\prod_{i=1}^k[\alpha_i-\delta,\beta_i+\delta])$.

Let us consider $\lambda_i \in C_c^\infty(\mathbb{R})$ such that $\mathbb{1}_{[\alpha_i,\beta_i]}(x) \leqslant \lambda_i(x) \leqslant \mathbb{1}_{[\alpha_i-\delta,\beta_i+\delta]}(x)$. Defining $\lambda(x) := \prod_{i=1}^k \lambda_i(x)$, we obtain,

$$\int \lambda dm \geqslant m\left(\prod_{i=1}^k[\alpha_i,\beta_i]\right) > m'\left(\prod_{i=1}^k[\alpha_i-\delta,\beta_i+\delta]\right) \geqslant \int \lambda dm'.$$

Considering $\Phi(m'') := h(\int \lambda dm'')$, with $h(x) := x$ if $|x| \leqslant 1$ and $h(x) = x/|x|$ if $|x| > 1$, we have $\Phi(m) \neq \Phi(m')$. It only remains to prove that, $\Phi \in \mathcal{M}$, that is, $|\lambda_i(x) - \lambda_i(y)| \leqslant C|a(x) - a(y)|$. This is a straightforward consequence of the fact that $\sup_{x\in\mathbb{R}} \frac{|\lambda_i'(x)|}{|a'(x)|} < \infty$, since λ_i belongs to $C_c^1(\mathbb{R})$, and $a'(x) \neq 0$, for any $x \in \mathbb{R}$.

Step 2. Now we can deduce, from Step 1, the convergence in distribution of μ^N to μ. As a consequence of Proposition 2.2 of Sznitman (1991) and Theorem 1.3, the sequence μ^N is tight. Let $\hat{\mu}$ be any limit of a converging subsequence of μ^N. The continuity of $\pi_{(t_1,\dots,t_n)}$ and Step 1 imply that, for all $t_1,\dots,t_n \in \mathbb{R}_+$,

$$\mathcal{L}(\mu_{(t_1,\dots,t_n)}) = \mathcal{L}(\hat{\mu}_{(t_1,\dots,t_n)}). \tag{24}$$

To conclude the proof, we just have to show that $\mathcal{L}(\mu) = \mathcal{L}(\hat{\mu})$. Let us consider $\mathcal{N}$ the algebra composed of the functions of the form

$$\Phi : m \in \mathcal{P}(D(\mathbb{R}_+,\mathbb{R})) \longmapsto h\left(\int \lambda_1 dm_{(t^1_1,\dots,t^1_{k_1})},\dots,\int \lambda_n dm_{(t^n_1,\dots,t^n_{k_n})}\right),$$

where $\lambda_i \in C_b(\mathbb{R}^{k_i}), h \in C_b(\mathbb{R}^n), t^j_i \in \mathbb{R}_+$.

By (24), for all $\Phi \in \mathcal{N}$, $\mathbb{E}[\Phi(\hat{\mu})] = \mathbb{E}[\Phi(\mu)]$. Now, using Theorem 3.4.5.(a) of Ethier and Kurtz (2005), we just have to show that $\mathcal{N}$ separates the points of $\mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$.

This last point is straightforward: let $m, m' \in \mathcal{P}(D(\mathbb{R}_+,\mathbb{R}))$ such that $m \neq m'$. This implies the existence of $t_1,\dots,t_k$ such that $\mathcal{L}(X_{(t_1,\dots,t_k)}) \neq \mathcal{L}(Y_{(t_1,\dots,t_k)})$, that is, $\int \lambda dm_{(t_1,\dots,t_k)} \neq \int \lambda dm'_{(t_1,\dots,t_k)}$

for some $\lambda \in C_b(\mathbb{R}^k)$. Now let $h : \mathbb{R} \mapsto \mathbb{R}$ be defined as $h(x) = x$ if $|x| \leqslant ||\lambda||_\infty + 1$, and $h(x) = \frac{x}{|x|}(||\lambda||_\infty + 1)$ otherwise. Then $\Phi : m'' \mapsto h(\int \lambda dm''_{(t_1,\dots t_k)})$ belongs to $\mathcal{N}$ and satisfies $\Phi(m) \neq \Phi(m')$.

3.3. Proof of Corollary 1.2

Applying Ito's formula, we have

$$\varphi(\bar{X}_t) = \varphi(\bar{X}_0) + \int_0^t \left(\varphi'(\bar{X}_s)b(\bar{X}_s) + \frac{1}{2}\varphi''(\bar{X}_s)\mu_s(f)\right) ds + \int_0^t \varphi'(\bar{X}_s)\sqrt{\mu_s(f)}dW_s$$
$$+ \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\{z\leqslant f(\bar{X}_{s-}\}} \left(\varphi(0) - \varphi(\bar{X}_{s-}\right) \pi(ds, dz, du). \quad (25)$$

Since φ', φ'', b and f are bounded, it follows from Fubini's theorem that

$$\mathbb{E}\left(\int_0^t \left(\varphi'(\bar{X}_s)b(\bar{X}_s) + \frac{1}{2}\varphi''(\bar{X}_s)\mu_s(f)\right)ds|W\right) = \int_0^t E\left(\varphi'(\bar{X}_s)b(\bar{X}_s) + \frac{1}{2}\varphi''(\bar{X}_s)\mu_s(f)|W\right) ds$$
$$= \int_0^t \int_{\mathbb{R}} \left(\varphi'(x)b(x) + \frac{1}{2}\varphi''(x)\mu_s(f)\right) \mu_s(dx)ds.$$

Moreover, by independence of $\bar{X}_0$ and W, $\mathbb{E}(\varphi(\bar{X}_0)|W) = \int_{\mathbb{R}} \varphi(x)\nu_0(dx)$.

To deal with the martingale part in (25), we use an Euler scheme to approximate the stochastic integral $I_t := \int_0^t \varphi'(\bar{X}_s)\sqrt{\mu_s(f)}dW_s$. For that sake, let $t_k^n := k2^{-n}t, 0 \leqslant k \leqslant 2^n$, $n \geqslant 1$, and define

$$I_t^n := \sum_{k=0}^{2^n-1} \varphi'(\bar{X}_{t_k^n})\Delta_k^n,\ \Delta_k^n = \int_{t_k^n}^{t_{k+1}^n} \sqrt{\mu_s(f)}dW_s,$$

then $\mathbb{E}(|I_t - I_t^n|^2) \to 0$ as $n \to \infty$, and therefore $\mathbb{E}(I_t^n|W) \to \mathbb{E}(I_t|W)$ in $L^2(P)$, as $n \to \infty$. But

$$\mathbb{E}(I_t^n|W) = \sum_{k=0}^{2^n-1} \mathbb{E}(\varphi'(\bar{X}_{t_k^n})|W)\Delta_k^n \to \int_0^t \mathbb{E}(\varphi'(\bar{X}_s)|W)\sqrt{\mu_s(f)}dW_s$$

in $L^2(P)$, since the sequence of processes $Y_s^n := \sum_{k=0}^{2^n-1} \mathbb{1}_{]t_k^n, t_{k+1}^n]}(s)\mathbb{E}(\varphi'(\bar{X}_{t_k^n})|W), 0 \leqslant s \leqslant t$, converges in $L^2(\Omega \times [0,t])$ to $\mathbb{E}(\varphi'(\bar{X}_s)|W)$.

We finally deal with the jump part in (25). Since f is bounded, and by independence of W and π, we can rewrite this part in terms of an underlying Poisson process N_t, independent of W and having rate $\|f\|_\infty$, and in terms of i.i.d. variables $(V_n)_{n\geqslant 1}$ uniformly distributed on $[0,1]$, independent of W and of N as follows.

$$\int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \mathbb{1}_{\{z\leqslant f(\bar{X}_{s-}\}} \left(\varphi(0) - \varphi(\bar{X}_{s-}\right) \pi(ds, dz, du) = \sum_{n=1}^{N_t} \mathbb{1}_{\{\|f\|_\infty V_n \leqslant f(\bar{X}_{T_n-})\}}(\varphi(0) - \varphi(\bar{X}_{T_n-})).$$

Taking conditional expectation $E(\cdot|W)$, we obtain

$$\mathbb{E}\left(\sum_{n=1}^{N_t} \mathbb{1}_{\{\|f\|_\infty V_n \leqslant f(\bar{X}_{T_n-})\}}(\varphi(0) - \varphi(\bar{X}_{T_n-}))|W\right) =$$

$$\mathbb{E}\left(\sum_{n=1}^{N_t} \frac{f(\bar{X}_{T_n-})}{\|f\|_\infty}(\varphi(0)-\varphi(\bar{X}_{T_n-}))|W\right) = \int_0^t \mathbb{E}\left(f(\bar{X}_s)(\varphi(0)-\varphi(\bar{X}_s))|W\right) ds,$$

where we have used the independence properties of $(V_n)_n, N_t$ and W and the fact that conditionally on $\{N_t = n\}$, the jump times $(T_1, \dots, T_n)$ are distributed as the order statistics of n i.i.d. times which are uniformly distributed on $[0, t]$. This concludes our proof.

3.4. A priori estimates

In this subsection, we prove useful a priori upper bounds on some moments of the solutions of the SDEs (7) and (14). Most of our previous results were stated under our Assumptions 1 and 2. However our computations hold true under weaker assumptions as shows the following

Lemma 3.1. *If f is subquadratic and b sublinear, if the measures ν and ν_0 admit a second moment and $\int_{\mathbb{R}} u d\nu(u) = 0$, then*

(i) for all $t > 0$, $\sup_{N\in\mathbb{N}^*} \sup_{0\leqslant s\leqslant t} \mathbb{E}\left[\left(X_t^{N,1}\right)^2\right] < +\infty$,

(ii) for all $t > 0$, $\sup_{N\in\mathbb{N}^*} \mathbb{E}\left[\sup_{0\leqslant s\leqslant t}\left|X_t^{N,1}\right|\right] < +\infty$,

(iii) for all $t > 0$, $\sup_{N\in\mathbb{N}^*} \sup_{0\leqslant s\leqslant t} \mathbb{E}\left[\left(\widetilde{X}_t^{N,1}\right)^2\right] < +\infty$.

(iv) for all $t > 0$, $\sup_{N\in\mathbb{N}^*} \mathbb{E}\left[\sup_{0\leqslant s\leqslant t}\left|\widetilde{X}_t^{N,1}\right|\right] < +\infty$.

Proof. We just prove (i) and (ii); (iii) and (iv) follow from similar arguments. By Ito's formula, we have that

$$\left(X_t^{N,1}\right)^2 \leqslant \left(X_0^{N,1}\right)^2 + 2\int_0^t \left(X_s^{N,1}\right) b\left(X_s^{N,1}\right) ds + \sum_{j=2}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} \left[\left(X_{s-}^{N,1} + \frac{u}{\sqrt{N}}\right)^2 - \left(X_{s-}^{N,1}\right)^2\right] \mathbb{1}_{\{z\leqslant f(X_{s-}^{N,j})\}} d\pi^j(s,z,u).$$

As f is subquadratic, b is sublinear, and the $X_s^{N,j}$ are identically distributed,

$$\begin{aligned}\mathbb{E}\left[\left(X_t^{N,1}\right)^2\right] &\leqslant \mathbb{E}\left[\left(X_0^{N,1}\right)^2\right] + Ct + C\int_0^t \mathbb{E}\left[\left(X_s^{N,1}\right)^2\right] ds + \frac{C}{N}\sum_{j=2}^N \int_0^t \mathbb{E}\left[\left(X_s^{N,j}\right)^2\right] ds \\ &\leqslant \mathbb{E}\left[\left(X_0^{N,1}\right)^2\right] + Ct + C\int_0^t \mathbb{E}\left[\left(X_s^{N,1}\right)^2\right] ds,\end{aligned}$$

where the constant C is not the same in the two lines above.

Then, we prove the lemma using Grönwall's lemma, and stopping times $\tau_K^N := \bigwedge_{1\leqslant i\leqslant N} \tau_K^{N,i}$ with $\tau_K^{N,i} := \inf\{t \geqslant 0 \ : \ |X_t^{N,i}| > K\}$.

(ii) We use that

$$\left|X_t^{N,1}\right| \leqslant \left|X_0^{N,1}\right| + Ct + C\int_0^t \left|X_s^{N,1}\right| ds + \frac{1}{\sqrt{N}}|M_t^N|,$$

where M_t^N is the martingale $M_t^N = \sum_{j=2}^N \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u\mathbb{1}_{\{z\leqslant f(X_{s-}^{N,j})\}} d\pi^j(s,z,u)$. Then

$$\sup_{0\leqslant s\leqslant t}\left|X_s^{N,1}\right| \leqslant \left|X_0^{N,1}\right| + Ct + C\int_0^t |X_s^{N,1}|ds + \frac{1}{\sqrt{N}}\sup_{0\leqslant s\leqslant t}|M_s^N|.$$

Now, to conclude the proof, it is sufficient to notice that

$$\frac{1}{\sqrt{N}}\mathbb{E}\left[\sup_{0\leqslant s\leqslant t}|M_s^N|\right] \leqslant \mathbb{E}\left[\frac{1}{N}[M^N]_t\right]^{1/2}$$

is uniformly bounded in N. □

Acknowledgements

This research has been conducted as part of FAPESP projectResearch, Innovation and Dissemination Center for Neuromathematics(grant 2013/07699-0).

Bibliography

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

Brémaud, P. and Massoulié, L. (1996). Stability of Nonlinear Hawkes Processes. *The Annals of Probability* **24** 1563–1588.

Carmona, R., Delarue, F. and Lacker, D. (2016). Mean field games with common noise. *The Annals of Probability* **44** 3740–3803.

Coghi, M. and Flandoli, F. (2016). Propagation of chaos for interacting particles subject to environmental noise. *The Annals of Applied Probability* **26** 1407–1442.

Cormier, Q., Tanré, E. and Veltz, R. (2018). Long time behavior of a mean-field model of interacting neurons. *ArXiv e-prints*.

Daley, D. J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*, Second ed. Springer.

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

Delattre, S., Fournier, N. and Hoffmann, M. (2016). Hawkes processes on large networks. *The Annals of Applied Probability* **26** 216–261.

Dermoune, A. (2003). Propagation and conditional propagation of chaos for pressureless gas equations. *Probab. Theory Related Fields* **126** 459–476.

Ditlevsen, S. and Löcherbach, E. (2017). Multi-class Oscillating Systems of Interacting Neurons. *Stochastic Processes and their Applications* **127** 1840–1869.

Erny, X., Löcherbach, E. and Loukianova, D. (2019). Mean field limits for interacting Hawkes processes in a diffusive regime. *arXiv:1904.06985 [math]*.

Ethier, S. and Kurtz, T. (2005). *Markov Processes. Characterization and Convergence.* Wiley Series In Probability And Statistics.

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.

Galves, A. and Löcherbach, E. (2013). Infinite systems of interacting chains with memory of variable length - a stochastic model for biological neural nets. *Journal of Statistical Physics* **5** 896–921.

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

Hewitt, E. and Savage, L. J. (1955). Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* **80** 470–501.

Ikeda, N. and Watanabe, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, Second ed. North-Holland Publishing Company.

Komlós, J., Major, P. and Tusnády, G. (1976). An approximation of partial sums of independent RV's, and the sample DF. II. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **34** 33–58.

Kurtz, T. G. (1978). Strong approximation theorems for density dependent Markov chains. *Stochastic Processes and their Applications* **8** 223–240.

Rissanen, J. (1983). A universal data compression system. *IEEE Trans. Inform. Theory* **29** 656–664.

Sznitman, A.-S. (1991). 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.