HAL open science

Mean field system of a two-layers neural model in a diffusive regime

Xavier Erny

To cite this version:

Xavier Erny. Mean field system of a two-layers neural model in a diffusive regime. 2021. hal-03255276v1

HAL Id: hal-03255276

https://hal.science/hal-03255276v1

Preprint submitted on 9 Jun 2021 (v1), last revised 6 Apr 2022 (v4)

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.

Mean field system of a two-layers neural model in a diffusive regime

Xavier Erny

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

Abstract: We study a model of interacting neurons. The structure of this neural system is composed of two layers of neurons such that the neurons of the first layer send their spikes to the neurons of the second one: if N is the number of neurons of the first layer, at each spiking time of the first layer, every neuron of both layers receives an amount of potential of the form $U/\sqrt{N}$, where U is a centered random variable. This kind of structure of neurons can model the neural retina. In the model, we study the "averaged effect" of the neurons of the first layer on a single neuron of the second layer. The theoretical model consists in two stochastic processes, one modelling the membrane potential of the neurons of the first layer, and the other the membrane potential of the neuron of the second layer. We prove the convergence of these processes as the number of neurons N goes to infinity and obtain a convergence speed. The proofs rely on similar arguments as those used in Erny, Löcherbach and Loukianova (2019): the convergence speed of the semigroups of the processes is obtained from the convergence speed of their infinitesimal generators using a Trotter-Kato formula, and from the regularity of the limit semigroup. Contrarily to the situation in Erny, Löcherbach and Loukianova (2019), the stochastic flow of the limit process is not continuous, and we need to use a kind of Girsanov's result to recover the regularity of the limit semigroup from the regularity of the stochastic flow of an auxiliary process.

MSC2020 subject classifications: 60J35, 60J60, 60K35, 60G55.
Keywords and phrases: Mean field interaction, Piecewise deterministic Markov processes, Interacting particle systems.

1. Introduction

1.1. Model

The aim of this paper is to study a model of two layers of neurons such that:

- the neurons of the first layer send their spikes to the neurons of both the first and second layers,
- the spikes of the neurons of the second layer do not have effect on these two layers.

This kind of multi-layers structure, where the neurons of a layer send their spikes to the neurons of the next one, can model the neural retina. This kind of interaction with more than two layers has already been studied in Ditlevsen and Löcherbach (2017) and at Section 5 of Erny, Löcherbach and Loukianova (2021).

We note N the number of neurons of the first layer. The idea of our modelling is to study the "averaged effect" of the first layer on a single neuron of the second one. We model the behavior of the first layer in the following way: every neuron of the first layer shares the same membrane potential and this potential does not undergo the effect of the repolarization. In our frame, it is natural not to consider the effect of the repolarization for the first layer since it affects each neuron

individually and we only consider these neurons through an averaged affect on a neuron of the second layer.

Let us note $Z^{N,1}$ the counting process related to the neuron of the second layer, and $Z^{N,i}$ $(2 \leq i \leq N+1)$ the counting processes related to the neurons of the first layer. In the model, $Z_t^{N,1}$ is the number of spikes emitted by the neuron of the second layer, and $Z_t^{N,i}$ $(2 \leq i \leq N+1)$ is the number of spikes emitted by the $i-$th neuron of the first layer before time t. We assume that $Z^{N,1}$ admits a stochastic intensity of the form $(f_2(Y_{t-}^N))_{t\geq 0}$ where Y_t^N is the membrane potential of the neuron of the second layer at time t and f_2 is its spike rate function. Similarly, we assume that every $Z^{N,j}$ $(2 \leq j \leq N+1)$ admits a stochastic intensity of the form $(f_1(X_{t-}^N))_{t\geq 0}$ where X_t^N models the averaged effect of the membrane potentials of the neurons of the first layer at time t and f_1 is the spike rate function of these neurons. We assume that this function f_1 is the same for all the neurons of the first layer, meaning that these neurons have similar characteristics. It is well-known that such point processes can be represented as thinnings of Poisson measures (see e.g. Theorem $II.7.4$ of Ikeda and Watanabe (1989)):

$$
\begin{aligned}
Z_t^{N,1} &= \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\{z\leq f_2(Y_{s-}^N)\}} d\pi^1(t,z),\\
Z_t^{N,i} &= \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\{z\leq f_1(X_{s-}^N)\}} d\pi^i(t,z), \quad 2 \leq i \leq N+1,
\end{aligned}
$$

where π^i $(1 \leq i \leq N+1)$ are i.i.d. Poisson measures on $\mathbb{R}_+^2$ having intensity $dt \cdot dz$. For the moment, let us note J_t^N the jump process that models the effect of the spikes of the neurons of the first layer. Then, the dynamics of the process X^N and Y^N are given by the following SDEs:

$$
\begin{aligned}
dX_t^N &= -\alpha_1 X_t^N dt + dJ_t^N,\\
dY_t^N &= -\alpha_2 Y_t^N dt + dJ_t^N - Y_{t-}^N dZ_t^{N,1}.
\end{aligned} \tag{1}
$$

In the SDEs above, α_1 and α_2 are positive parameters of the model. These drifts model the fact that the potentials of the neurons tend to go to the resting value that we assume to be zero. In this model, we assume that the potentials go to zero at exponential speed with respective rate α_1 for the neurons of the first layer, and α_2 for the neuron of the second layer. In the following, J^N is directed by the point processes $Z^{N,i}$ $(2 \leq i \leq N+1)$, modelling the fact that the spikes of the neurons of the first layers modifies the values of the potentials of the neurons. The third term in the SDE of Y^N above models the repolarization of the potential of the neuron: right after the neuron emits a spike, its potential goes fast (here it jumps immediately) to its resting value that we assume to be zero.

Remark 1.1. *Note that it is not mandatory to assume the parameters α_1 and α_2 to be positive for our results to hold true. This condition is natural for the modelling.*

In order to study the limit of this model as the number of neurons N goes to infinity, we need to assume that the effect of each individual spike of the neurons vanishes as N goes to infinity. To the best of our knowledge, there exists two such kinds of models: when the effect of a spike is of order N^{-1} (which corresponds to a law of large number) and when it is stochastic, centered and of order $N^{-1/2}$ (which corresponds to a central limit theorem). In the first case, the process J^N is of the form

$$
J_t^N = \frac{1}{N} \sum_{j=2}^{N+1} \int_{[0,t]\times\mathbb{R}_+} \mathbb{1}_{\{z\leq f_1(X_{s-}^N)\}} d\pi^j(t,z). \tag{2}
$$

In the second one, to make the height of the jump stochastic, we need to consider Poisson measures π^j $(2 \leq j \leq N+1)$ defined on $\mathbb{R}_+^2 \times \mathbb{R}$ rather than on $\mathbb{R}_+^2$, with intensity $dt \cdot dz \cdot d\nu(u)$, where ν is a centered probability measure. In this case, the process J^N is of the form

$$J_t^N = \frac{1}{\sqrt{N}} \sum_{j=2}^{N+1} \int_{[0,t]\times\mathbb{R}_+\times\mathbb{R}} u \cdot \mathbb{1}_{\{z \leq f_1(X_{s-}^N)\}} d\pi^j(s,z,u). \tag{3}$$

The convergence of the processes X^N, Y^N and $Z^{N,i}$ $(i \geq 1)$ as N goes to infinity in the first case (i.e. when J^N is given by (2)) is classical. We give the related results and references at Section 3.

So in this paper, we focus mainly on the second case (i.e. when J^N is given by (3)). It is a generalization of the model of Erny, Löcherbach and Loukianova (2019) where we have studied the convergence of the process X^N. The interesting points of the model of this paper is that it is more relevant for the modelling of a particular structure of neural network, and that we can still prove similar results.

We are interested in proving the convergence (and obtaining a convergence speed) of the following sequence of two-dimensional Markov processes

$$\begin{aligned} dX_t^N &= -\alpha_1 X_t^N dt + \frac{1}{\sqrt{N}} \sum_{j=2}^{N+1} \int_{\mathbb{R}_+\times\mathbb{R}} u \cdot \mathbb{1}_{\{z \leq f_1(X_{t-}^N)\}} d\pi^j(t,z,u), \\ dY_t^N &= -\alpha_2 Y_t^N dt + \frac{1}{\sqrt{N}} \sum_{j=2}^{N+1} \int_{\mathbb{R}_+\times\mathbb{R}} u \cdot \mathbb{1}_{\{z \leq f_1(X_{t-}^N)\}} d\pi^j(t,z,u) - \int_{\mathbb{R}_+} Y_{t-}^N \mathbb{1}_{\{z \leq f_2(Y_{t-}^N)\}} d\pi^1(t,z). \end{aligned} \tag{4}$$

to the following limit process

$$\begin{aligned} d\bar{X}_t &= -\alpha_1 \bar{X}_t + \sigma\sqrt{f_1(\bar{X}_t)} dW_t, \\ d\bar{Y}_t &= -\alpha_2 \bar{Y}_t dt + \sigma\sqrt{f_1(\bar{X}_t)} dW_t - \int_{\mathbb{R}_+} \bar{Y}_{t-} \mathbb{1}_{\{z \leq f_2(\bar{Y}_{t-})\}} d\bar{\pi}^1(t,z). \end{aligned} \tag{5}$$

where W is a standard one-dimensional Brownian motion and $\bar{\pi}^1$ is a Poisson measure on $\mathbb{R}_+^2$ with Lebesgue intensity.

Note that it can be more practical to simulate the trajectories of the process $(\bar{X}, \bar{Y})$ than those of (X^N, Y^N). Indeed, to simulate (X^N, Y^N) one has to compute the behavior of the N point processes $Z^{N,i}$ $(2 \leq i \leq N+1)$, which can be computationaly hard if N is large. So it can be easier to simulate a single Brownian motion W instead of those point processes. If one wants to approximate the process (X^N, Y^N) with $(\bar{X}, \bar{Y})$, it is interesting to quantify "weak error" between these processes. This is done in Theorem 1.2.

To obtain the convergence speed of the semigroups $(P_t^N)_t$ of the processes $(X_t^N, Y_t^N)_t$ to the semigroup $(\bar{P}_t)_t$ of the limit process $(\bar{X}_t, \bar{Y}_t)_t$, we use similar arguments as in Erny, Löcherbach and Loukianova (2019). We consider the infinitesimal generators A^N and $\bar{A}$ of respectively P^N and $\bar{P}$. We begin by proving the convergence of these generators and by obtaining a convergence speed. Then, we deduce a convergence speed for the semigroups using the following formula:

$$\bar{P}_t g(x,y) - P_t^N g(x,y) = \int_0^t P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(x,y) ds.$$

One of the main step of the proof consists in studying the regularity of the limit semigroup $(\bar{P}_t)_t$. For this purpose, we need to study the regularity of the related stochastic flow. Compared to the proofs of Erny, Löcherbach and Loukianova (2019), there are two additional technical difficulties that we handle at Section 2.2:

- The trajectories of the process $(\bar{X}, \bar{Y})$ are not continuous. Whence, its stochastic flow is not a Brownian flow and Theorem 1.4.1 of Kunita (1986) cannot be used on this flow.
- The jump times of the process $(\bar{X}, \bar{Y})$ occur at the rate $f_2(\bar{Y}_{t-})$. This implies that this rate depends on the initial condition of the process. The regularity of this dependency is hard to study directly.

To overcome the first point, we study the regularity of the stochastic flow between the jump times (see Proposition 2.6). Let us remark that Bally, Goreac and Rabiet (2018) have already used the idea of working between the jump times of the flow to study its regularity and that of its semigroup. For the second point, we introduce an auxiliary limit process $(\tilde{X}, \tilde{Y})$ for which the rate of the jump is constant. Then, we deduce the regularity of the semigroup $(\bar{P}_t)_t$ from the regularity of the flow of $(\tilde{X}, \tilde{Y})$ using a kind of "Girsanov result" giving the explicit Radon-Nikodym derivative between the processes $(\bar{X}, \bar{Y})$ and $(\tilde{X}, \tilde{Y})$ (see Proposition 2.3).

1.2. Notation

Throughout the paper, we use the following notation:

- For $n \in \mathbb{N}^*$, we note $C_b^n(\mathbb{R})$ (resp. $C_b^n(\mathbb{R}^2)$) the set of $\mathbb{R}-$valued functions defined on $\mathbb{R}$ (resp. $\mathbb{R}^2$) that are C^n, bounded and such that all their derivatives (resp. partial derivatives) up to order n are bounded.
- We call $\beta \in \mathbb{N}^2$ a multi-index and we note $|\beta| = \beta_1 + \beta_2$. If $g : \mathbb{R}^2 \to \mathbb{R}$ is C^n and if β is a multi-index such that $|\beta| \leq n$, we write $\partial_\beta g$ the partial derivative of g w.r.t. β. For the first order partial derivatives, we write instead $\partial_1 g$ and $\partial_2 g$, and for the second order derivatives, we write $\partial^2_{i,j} g$ $(i, j \in \{1, 2\})$. When there is no possible confusion, we may use the classical notation $\partial_x g(x, y) = \partial_1 g(x, y)$.
- For $g \in C_b^n(\mathbb{R}^2)$, we note
$$||g||_{n,\infty} = \sum_{|\beta| \leq n} ||\partial_\beta g||_\infty.$$
- Given a process X (resp. a multi-dimensional process (X, Y)) that is solution to some SDE, we note $X^{(x)}$ (resp. $(X^{(x,y)}, Y^{(x,y)})$) the process solution to the SDE w.r.t. the initial condition $X_0 = x$ (resp. $(X_0, Y_0) = (x, y)$). The function $x \mapsto X^{(x)}$ (resp. $(x, y) \mapsto (X^{(x,y)}, Y^{(x,y)})$) is called the stochastic flow of the process X (resp. (X, Y)).
- In all the paper, C denotes an arbitrary constant that is independent of any parameters except the parameters of the model. If a constant depends on some parameter θ, we write C_θ instead. The value of such constants can change from line to line in an equation.

1.3. Assumptions and main result

To begin with, let us state the conditions we need to study the model.

Assumption 1.

(i) The function f_1 is continuous, bounded and positive,
(ii) the function f_2 is non-negative, Lipschitz continuous and belongs to $C_b^3(\mathbb{R})$
(iii) the function $\sqrt{f_1}$ is Lipschitz continuous and belongs to $C_b^4(\mathbb{R})$,
(iv) the probability measure ν is centered and admits a finite third moment.

Under Assumption 1.(i)-(ii), the well-posedness of the previous SDEs is classical. The boundedness of f_1 is not mandatory for our results to hold true but it simplifies some technical details. Without this boundedness, the convergence speed given in Theorem 1.2 would depend on the initial condition. However, the boundedness of f_2 is important to control the Radon-Nikodym derivative between the limit process $(\bar{X}, \bar{Y})$ and the auxiliary process $(\tilde{X}, \tilde{Y})$ that we introduce at Section 2.2 (see Proposition 2.3 for the form of this derivative).

Condition (iii) allows to prove that the stochastic flow of the auxiliary process is C^3 between its jump times (see Proposition 2.5). And condition (iv) allows us to obtain a convergence speed for the generators (see Lemma 2.2).

Under the additional assumption that (X_0^N, Y_0^N) converges in distribution to $(\bar{X}_0, \bar{Y}_0)$, it is classical that the process (X^N, Y^N) converges in distribution to $(\bar{X}, \bar{Y})$ in Skorohod topology (see Theorem IX.4.21 of Jacod and Shiryaev (2003)). The main result of this paper consists in obtaining the convergence speed of the semigroups P^N and $\bar{P}$ of these processes (recalling that the processes (X^N, Y^N) and $(\bar{X}, \bar{Y})$ are defined respectively at (4) and (5)).

Theorem 1.2. *Grant Assumption 1.(i)-(iv). For all $t \geq 0$, $N \in \mathbb{N}^*$, $g \in C_b^3(\mathbb{R}^2)$ and $x, y \in \mathbb{R}$,*

$$\left|\bar{P}_t g(x,y) - P_t^N g(x,y)\right| \leq C_t \cdot N^{-1/2}.$$

Theorem 1.2 is proved at Section 2.3. The following figure presents two simulations of the process Y^N with $N = 100$ and $N = 1000$.

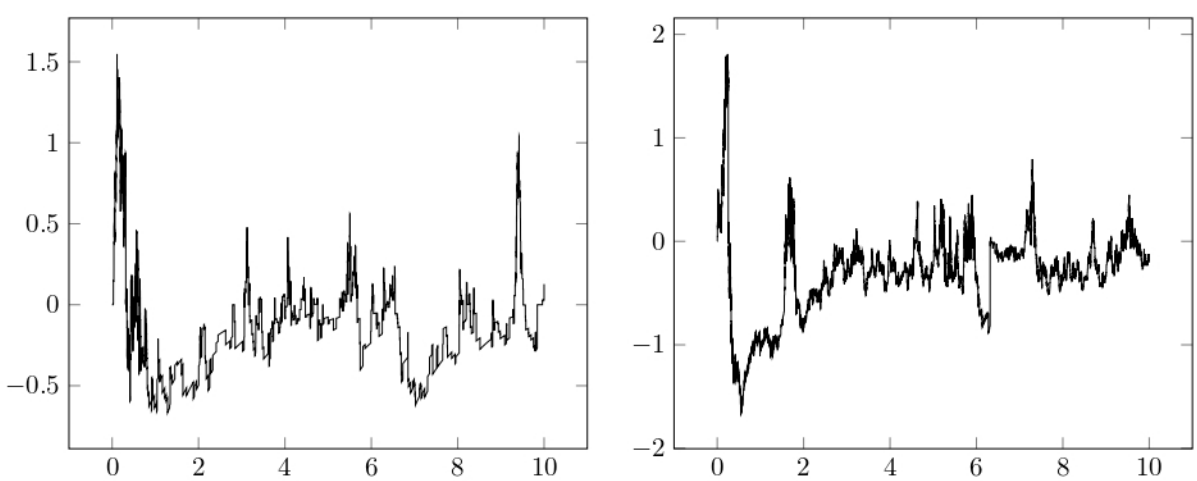

FIGURE 1. *Simulation of trajectories of $(Y_t^N)_{0 \leq t \leq 10}$ with $X_0^N = 0$, $Y_0^N = 0$, $\alpha_1 = \alpha_2 = 1$, $\nu = \mathcal{N}(0,1)$, $f_1(x) = f_2(x) = 2.2 + 1.4 \arctan(10x - 2)$, $N = 100$ (left picture) and $N = 1000$ (right picture).*

2. Proofs

2.1. Infinitesimal generators

One can note that the processes $(\bar{X}, \bar{Y})$ and (X^N, Y^N) $(N \in \mathbb{N}^*)$ are Markov processes. In the following, we note $\bar{P}$ and P^N $(N \in \mathbb{N}^*)$ their respective semigroups, and $\bar{A}$ and A^N $(N \in \mathbb{N}^*)$ their respective infinitesimal generators. We refer to Appendix A of Erny, Löcherbach and Loukianova (2019) for the notion of generator we use here, and for their precise definition. In particular, if X is a Markov process and A its infinitesimal generator (in the sense of Definition $A.1$ of Erny, Löcherbach and Loukianova (2019)), we note $\mathcal{D}'(A)$ the set of functions $g : \mathbb{R}^2 \to \mathbb{R}$ such that, for all $x \in \mathbb{R}^2$,

- $\mathbb{E}_x g(X_t) - g(x) = \mathbb{E}_x \int_0^t Ag(X_s)ds$,
- $\mathbb{E}_x \int_0^t |Ag(X_s)|ds < \infty$.

The following lemma is a direct consequence of Ito's formula and of the form of the SDE (4)

Lemma 2.1. *Grant Assumption 1.(i)-(ii).*
$C_b^2(\mathbb{R}^2) \subseteq \mathcal{D}'(\bar{A})$, *and for* $g \in C_b^2(\mathbb{R}^2)$ *and* $(x, y) \in \mathbb{R}^2$,

$$\bar{A}g(x,y) = -\alpha_1 x \partial_1 g(x,y) - \alpha_2 y \partial_2 g(x,y) + \frac{\sigma^2}{2} f_1(x) \sum_{i,j=1}^{2} \partial_{i,j}^2 g(x,y) + f_2(y)\left[g(x,0) - g(x,y)\right].$$

$C_b^1(\mathbb{R}^2) \subseteq \mathcal{D}'(A^N)$, *and for* $g \in C_b^1(\mathbb{R}^2)$ *and* $(x, y) \in \mathbb{R}^2$,

$$\begin{aligned} A^N g(x,y) = &- \alpha_1 x \partial_1 g(x,y) - \alpha_2 y \partial_2 g(x,y) \\ &+ N f_1(x) \int_{\mathbb{R}} \left[g\left(x + \frac{u}{\sqrt{N}}, y + \frac{u}{\sqrt{N}}\right) - g(x,y) \right] d\nu(u) + f_2(y)\left[g(x,0) - g(x,y)\right]. \end{aligned}$$

We have the following control on the convergence of the generators.

Lemma 2.2. *Grant Assumption 1.(i)-(ii) and Assumption 1.(iv). For* $g \in C_b^2(\mathbb{R}^2)$ *and* $(x, y) \in \mathbb{R}^2$,

$$\left|A^N g(x,y) - \bar{A}g(x,y)\right| \leq \frac{1}{6} f_1(x) N^{-1/2} \left(\int_{\mathbb{R}} |u|^3 d\nu(u) \right) \sum_{|\beta|=3} ||\partial_\beta g||_\infty.$$

Proof. Recalling that ν is a centered probability measure and that σ^2 is its variance,

$$\begin{aligned} &A^N g(x,y) - \bar{A}g(x,y) \\ &= \int_{\mathbb{R}} N f_1(x) \left[g\left(x + \frac{u}{\sqrt{N}}, y + \frac{u}{\sqrt{N}}\right) - g(x,y) \right] d\nu(u) - \frac{\sigma^2}{2} f_1(x) \sum_{i,j=1}^{2} \partial_{i,j}^2 g(x,y) \\ &= N f_1(x) \int_{\mathbb{R}} \left[g\left(x + \frac{u}{\sqrt{N}}, y + \frac{u}{\sqrt{N}}\right) - g(x,y) - \frac{u}{\sqrt{N}} \sum_{i=1}^{2} \partial_i g(x,y) - \frac{u^2}{2N} \sum_{i,j=1}^{2} \partial_{i,j}^2 g(x,y) \right] d\nu(u). \end{aligned}$$

As a consequence

$$\begin{aligned}&\left|A^N g(x,y) - \bar{A}g(x,y)\right| \\ &\leq N f_1(x) \int_{\mathbb{R}} \left[\left| g\left(x + \frac{u}{\sqrt{N}}, y + \frac{u}{\sqrt{N}}\right) - g(x,y) - \frac{u}{\sqrt{N}} \sum_{i=1}^{2} \partial_i g(x,y) - \frac{u^2}{2N} \sum_{i,j=1}^{2} \partial^2_{i,j} g(x,y)\right|\right] d\nu(u).\end{aligned}$$

Then Taylor-Lagrange's inequality gives the result. □

2.2. Semigroups and stochastic flows

In this section, we prove the regularity of the semigroup of the limit process $(\bar{X}, \bar{Y})$ w.r.t. its initial condition and a control of its derivatives. However there are additional technical difficulties in this proof compared to Erny, Löcherbach and Loukianova (2019). The main difficulty is that the trajectories of the process $(\bar{X}, \bar{Y})$ are not continuous. Whence the associated stochastic flow is not a Brownian flow and Theorem 1.4.1 of Kunita (1986) cannot be applied directly to study the differentiability of the flow. So, we study the stochastic flow between its jump times. Then we face another difficulty: the jump times of the stochastic flow depends on the initial condition, and we do not want to study the regularity of this dependency.

This is why we introduce the following auxiliary limit process $(\tilde{X}, \tilde{Y})$. The only difference with the limit process $(\bar{X}, \bar{Y})$ is that the stochastic intensity of the jump term in the SDE of $\tilde{Y}$ is constant. In other words, the occurrences of the jump times of the process $\tilde{Y}$ are distributed as a homogeneous Poisson process.

$$\begin{aligned}d\tilde{X}_t &= -\alpha_1 \tilde{X}_t + \sigma\sqrt{f_1(\tilde{X}_t)} dW_t, \\ d\tilde{Y}_t &= -\alpha_2 \tilde{Y}_t dt + \sigma\sqrt{f_1(\tilde{X}_t)} dW_t - \int_{[0,||f_2||_\infty]} \tilde{Y}_{t-} d\tilde{\pi}^1(t,z),\end{aligned}$$

where $\tilde{\pi}^1$ is a Poisson measure on $\mathbb{R}^2_+$ with Lebesgue intensity.

Let us note $N_t = \tilde{\pi}^1([0,t] \times ||f_2||_\infty)$ the number of jumps of the process $\tilde{Y}$ before time t. By definition of $\tilde{\pi}^1$, $(N_t)_{t\geq 0}$ is a homogeneous Poisson process with rate $||f_2||_\infty$. We also note T_n $(n \geq 1)$ the jump times of the process $\tilde{Y}$.

The next proposition can be seen as a Girsanov's theorem for jump processes. It gives the Radon-Nikodym derivative between the distribution of $(\bar{X}_t, \bar{Y}_t)$ and that of $(\tilde{X}_t, \tilde{Y}_t)$, for any $t \geq 0$. This result is a straightforward application of Theorem 3.5 of Löcherbach (2002) (see also Theorem $VI.T3$ of Brémaud (1981)), its proof is therefore omitted.

Proposition 2.3. *Grant assumption 1.(i)-(ii). For all $g \in C_b(\mathbb{R}^2)$, and, $x, y \in \mathbb{R}$,*

$$\mathbb{E}\left[g\left(\bar{X}_t^{(x)}, \bar{Y}_t^{(x,y)}\right)\right] = \mathbb{E}\left[g\left(\tilde{X}_t^{(x)}, \tilde{Y}_t^{(x,y)}\right) \prod_{n=1}^{N_t} \frac{f_2\left(\tilde{Y}_{T_n}^{(x,y)}\right)}{||f_2||_\infty} \exp\left[-\int_0^t \left(f_2\left(\tilde{Y}_s^{(x,y)}\right) - ||f_2||_\infty\right) ds\right]\right],$$

with $N_t := \tilde{\pi}^1([0,t] \times [0, ||f_2||_\infty])$.

In order to obtain a control on the regularity of the semigroup $\bar{P}$ of the limit process $(\bar{X}, \bar{Y})$, we rely on the previous proposition and on a control of the derivatives of the stochastic flow of the auxiliary process $(\tilde{X}, \tilde{Y})$. The next step is to prove this control.

Let us note $\varphi_t(x, y)$ $(t \in \mathbb{R}_+, (x, y) \in \mathbb{R}^2)$ the stochastic flow related to the following $2-$dimensional SDE:

$$\begin{aligned} dX_t &= -\alpha_1 X_t dt + \sigma\sqrt{f_1(X_t)}dW_t, \\ dY_t &= -\alpha_2 Y_t dt + \sigma\sqrt{f_1(X_t)}dW_t. \end{aligned}$$

One can note that $\varphi_t(x, y)$ is the flow of the process $(\tilde{X}, \tilde{Y})$ between its jump times. We admit that $\varphi_t(x, y)$ is a Brownian flow in the sense of Kunita (1986).

Remark 2.4. *Contrarily to the more general framework of Kunita (1986), the stochastic flow $\varphi_t(x, y)$ does only depend on one time parameter, because the process is time-homogeneous.*

Proposition 2.5. *Grant Assumption 1.(i), Assumption 1.(iii) and assume that f_2 is bounded. Almost surely, for all $t \geq 0$, $(x, y) \mapsto \varphi_t(x, y)$ is a C^3-diffeomorphism.*

Proof. We rely on Theorem 1.4.1 of Kunita (1986) and use the notation therein. One can note that the local characteristics (a, b) of the flow $\varphi_t(x, y)$ are

$$b(x, y) = \begin{pmatrix} -\alpha_1 x \\ -\alpha_2 y \end{pmatrix} \text{ and } a((x, y), (x', y')) = \sigma^2\sqrt{f_1(x)f_1(x')}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

One can check that the Assumptions 1, 2 and 3 of Kunita (1986) are satisfied. Then, Theorem 1.4.1 of Kunita (1986) allows to conclude the proof. □

Proposition 2.6. *Grant Assumption 1.(i)-(iii). Almost surely, for all $t \geq 0$, $(x, y) \mapsto (\tilde{X}_t^{(x)}, \tilde{Y}_t^{(x,y)})$ is C^3 and, for all even $p \in \mathbb{N}^*$, for all multi-index $\beta \in \mathbb{N}^2$ such that $1 \leq |\beta| \leq 3$,*

$$\sup_{x\in\mathbb{R}} \mathbb{E}\left[\sup_{0\leq s\leq t}\left(\partial_\beta \tilde{X}_s^{(x)}\right)^p\right] + \sup_{x,y\in\mathbb{R}} \mathbb{E}\left[\sup_{0\leq s\leq t}\left(\partial_\beta \tilde{Y}_s^{(x,y)}\right)^p\right] < \infty.$$

Proof. We already know that, for all $t \geq 0$, $x \mapsto \tilde{X}_t^{(x)}$ is C^3 by Proposition 2.5. Indeed, by definition $\tilde{X}_t^{(x)}$ is the first coordinate of $\varphi_t(x, y)$.

Now, let us recall that T_n $(n \in \mathbb{N}^*)$ denote the jump times of the process $(\tilde{X}^{(x)}, \tilde{Y}^{(x,y)})$ and that these jump times do not depend on the initial condition (x, y). For $t < T_1$,

$$(\tilde{X}_t^{(x)}, \tilde{Y}_t^{(x,y)}) = \varphi_t(x, y)$$

is C^3 w.r.t. (x, y) by Proposition 2.5. And, for $t \in [T_n, T_{n+1}[$ $(n \in \mathbb{N}^*)$,

$$(\tilde{X}_t^{(x)}, \tilde{Y}_t^{(x,y)}) = \varphi_{t-T_n}(\tilde{X}_{T_n}^{(x)}, 0),$$

which is also C^3 w.r.t. (x, y) as a composition of the function $(x, y) \in \mathbb{R}^2 \mapsto \varphi_{t-T_n}(x, y)$ and the function $(x, y) \in \mathbb{R}^2 \mapsto (\tilde{X}_{T_n}^{(x)}, 0)$. This proves that, almost surely, for all $t \geq 0$, $(x, y) \mapsto (\tilde{X}_t^{(x)}, \tilde{Y}_t^{(x,y)})$ is C^3. Now, we prove the second part of the statement.

By definition, for all $t \geq 0$,

$$\tilde{X}_t^{(x)} = x - \alpha_1 \int_0^t \tilde{X}_s^{(x)} ds + \sigma \int_0^t \sqrt{f_1\left(\tilde{X}_s^{(x)}\right)} dW_s.$$

This implies

$$\partial_x \tilde{X}_t^{(x)} = 1 - \alpha_1 \int_0^t \partial_x \tilde{X}_s^{(x)} ds + \sigma \int_0^t \partial_x \tilde{X}_s^{(x)} \left(\sqrt{f_1}\right)' \left(\tilde{X}_s^{(x)}\right) dW_s.$$

We also know that

$$\partial_x \tilde{Y}_t^{(x,y)} = -\alpha_2 \int_0^t \partial_x \tilde{Y}_s^{(x,y)} ds + \sigma \int_0^t \partial_x \tilde{X}_s^{(x)} \cdot \sqrt{f_1\left(\tilde{X}_s^{(x)}\right)} dW_s - \int_0^t \int_{[0,||f_2||_\infty]} \partial_x \tilde{Y}_s^{(x,y)} d\tilde{\pi}(s,z).$$

Noting that the initial condition of the process $(\partial_x \tilde{X}_t^{(x)}, \partial_x \tilde{Y}_t^{(x,y)})$ is $(1,0)$, Lemma 4.2 allows to prove the result of the lemma for the first order partial derivative w.r.t. x. The results for the other partial derivatives follow from the same reasoning. □

Proposition 2.7. *Grant Assumption 1.(i)-(iii). For all $t \geq 0$ and $g \in C_b^3(\mathbb{R}^2)$, the function $(x,y) \in \mathbb{R}^2 \mapsto \bar{P}_t g(x,y)$ is C^3 and for all multi-index β such that $|\beta| = 3$,*

$$\sup_{0 \leq s \leq t} \left|\left|\partial_\beta \bar{P}_s g\right|\right|_\infty \leq C_t ||g||_{3,\infty},$$

where $C_t > 0$ does not depend on g.

Proof. For the sake of readability, we do not prove the result for the multi-indexes β such that $|\beta| = 3$ but for those such that $|\beta| = 1$. This case is easier to prove and the technics are exactly the same.

Recalling Proposition 2.3, we have

$$\bar{P}_t g(x,y) = \mathbb{E}\left[g\left(\tilde{X}_t^{(x)}, \tilde{Y}_t^{(x,y)}\right) \prod_{n=1}^{N_t} \frac{f_2\left(\tilde{Y}_{T_n}^{(x,y)}\right)}{||f_2||_\infty} \exp\left[-\int_0^t \left(f_2\left(\tilde{Y}_s^{(x,y)}\right) - ||f_2||_\infty\right) ds\right]\right]. \tag{6}$$

To prove that the function $\bar{P}_t g$ is C^1, we rely on Lemma 4.1 which is a generalization of the classical result about the "differentiation under the integral sign" that uses Vitali theorem instead of the dominated convergence theorem.

By Proposition 2.6, we know that the expression in the expectation of the right-hand side of (6) is C^1 almost surely. In addition, one can control the expectation of the derivatives of this expression thanks to Proposition 2.6 (see below for the explicit expression of the derivatives).

So Lemma 4.1 (applied twice) implies that $\bar{P}_t g$ is C^1. In addition, denoting by

$$L_t(x,y) := \prod_{n=1}^{N_t} \frac{f_2\left(\tilde{Y}_{T_n}^{(x,y)}\right)}{||f_2||_\infty} \exp\left[-\int_0^t \left(f_2\left(\tilde{Y}_s^{(x,y)}\right) - ||f_2||_\infty\right) ds\right],$$

we have

$$\partial_2\left(\bar{P}_t g\right)(x,y) = \mathbb{E}\left[\partial_y \tilde{Y}_t^{(x,y)} \cdot \partial_2 g\left(\tilde{X}_t^{(x)}, \tilde{Y}_t^{(x,y)}\right) L_t(x,y)\right]$$
$$+\mathbb{E}\left[g\left(\tilde{X}_t^{(x)}, \tilde{Y}_t^{(x,y)}\right)\sum_{n=1}^{N_t} \partial_y \tilde{Y}_{T_n}^{(x,y)} \frac{f_2'\left(\tilde{Y}_{T_n}^{(x,y)}\right)}{||f_2||_\infty} \prod_{k\neq n} \frac{f_2\left(\tilde{Y}_{T_k}^{(x,y)}\right)}{||f_2||_\infty} \exp\left(\int_0^t (||f_2||_\infty - f_2(\tilde{Y}_s^{(x,y)}))ds\right)\right]$$
$$-\mathbb{E}\left[\int_0^t \partial_y \tilde{Y}_s^{(x,y)} f_2'\left(\tilde{Y}_s^{(x,y)}\right) ds \cdot g\left(\tilde{X}_t^{(x)}, \tilde{Y}_t^{(x,y)}\right) L_t(x,y)\right].$$

As a consequence,

$$\begin{aligned}\left|\partial_2\left(\bar{P}_t g\right)(x,y)\right| \leq & ||\partial_2 g||_\infty e^{t||f_2||_\infty}\mathbb{E}\left[\left|\partial_y \tilde{Y}_t^{(x,y)}\right|\right] \\ & + ||g||_\infty \frac{||f_2'||_\infty}{||f_2||_\infty} e^{||f_2||_\infty t}\mathbb{E}\left[N_t \sup_{0\leq s\leq t}\left|\partial_y \tilde{Y}_s^{(x,y)}\right|\right] \\ & + ||f_2'||_\infty \cdot ||g||_\infty e^{t||f_2||_\infty} t\mathbb{E}\left[\sup_{0\leq s\leq t}\left|\partial_y \tilde{Y}_s^{(x,y)}\right|\right].\end{aligned}$$

The expressions of the first line and the third line above can be bounded uniformly in (x,y) thanks to Proposition 2.6. For the expression of the second line, we can use Cauchy-Schwarz's inequality for example (we can recall that N_t follows the Poisson distribution with parameter $t||f_2||_\infty$).

This implies that

$$\sup_{x,y\in\mathbb{R}} \sup_{0\leq s\leq t} \left|\partial_2\left(\bar{P}_s g\right)(x,y)\right| \leq C_t ||g||_{1,\infty}.$$

With similar computation, we can prove that

$$\sup_{x,y\in\mathbb{R}} \sup_{0\leq s\leq t} \left|\partial_1\left(\bar{P}_s g\right)(x,y)\right| \leq C_t ||g||_{1,\infty}.$$

So we have proved the result for the two multi-indexes β such that $|\beta| = 1$. Note that in the case for the multi-indexes β such that $|\beta| = 3$, we would need a control on higher moments of the processes $\tilde{X}$, $\tilde{Y}$ and their derivatives to control expectations of product of these processes. Indeed, the partial derivative w.r.t. the multi-index $(1,2)$ would involve terms as

$$\mathbb{E}\left[\left(\sup_{0\leq s\leq t}|\partial_x \tilde{Y}_s^{(x,y)}|\right)\left(\sup_{0\leq s\leq t}|\partial_y \tilde{Y}_s^{(x,y)}|\right)^2\right]$$

which can be controlled using Cauchy-Schwarz's inequality, recalling that Proposition 2.6 allows to bound every polynomial moment of the partial derivatives of the process $(\tilde{X}, \tilde{Y})$ w.r.t. its initial condition. □

2.3. Proof of Theorem 1.2

With a similar proof as the one of Proposition 5.6 of Erny, Löcherbach and Loukianova (2019), we can prove that, for all $t \geq 0$, $N \in \mathbb{N}^*$, $g \in C_b^3(\mathbb{R}^2)$, $x, y \in \mathbb{R}$,

$$\left(\bar{P}_t - P_t^N\right) g(x,y) = \int_0^t P_{t-s}^N \left(\bar{A} - A^N\right) \bar{P}_s g(x,y) ds.$$

Consequently,

$$\begin{aligned}\left|\left(\bar{P}_t - P_t^N\right) g(x,y)\right| \leq & \int_0^t \mathbb{E}_{(x,y)}\left[\left|\left(\bar{A} - A^N\right)\bar{P}_s g(X_{t-s}^N, Y_{t-s}^N)\right|\right] ds \\ \leq & C \cdot N^{-1/2} \int_0^t \mathbb{E}_{(x,y)}\left[f_1(X_{t-s}^N) \sum_{|\beta|=3} ||\partial_\beta\left(\bar{P}_s g\right)||_\infty\right] ds \\ \leq & C_t \cdot N^{-1/2} \cdot ||f_1||_\infty ||g||_{3,\infty},\end{aligned}$$

where we have used Lemma 2.2 to obtain the second line above, and Proposition 2.7 to obtain the third one.

3. The model in normalization N^{-1}

The aim of this section is to state the convergence of the process (X^N, Y^N) given in (1) with the process J^N defined by (2) instead of (3). In this section, the process (X^N, Y^N) is solution to

$$\begin{aligned}dX_t^N =& -\alpha_1 X_t^N dt + \frac{1}{N}\sum_{j=2}^{N+1}\int_{\mathbb{R}_+} \mathbb{1}_{\{z \leq f_1(X_{t-}^N)\}} d\pi^j(t,z), \\ dY_t^N =& -\alpha_2 Y_t^N dt + \frac{1}{N}\sum_{j=2}^{N+1}\int_{\mathbb{R}_+} \mathbb{1}_{\{z \leq f_1(X_{t-}^N)\}} d\pi^j(t,z) - \int_{\mathbb{R}_+} Y_{t-}^N \mathbb{1}_{\{z \leq f_2(Y_{t-}^N)\}} d\bar{\pi}^1(t,z),\end{aligned}$$

where π^j $(1 \leq j \leq N+1)$ are i.i.d. Poisson measures on $\mathbb{R}_+^2$ with Lebesgue intensity.

In this framework the limit process $(\bar{X}, \bar{Y})$ is solution to the following SDE:

$$\begin{aligned}d\bar{X}_t =& -\alpha_1 \bar{X}_t dt + \mathbb{E}\left[f_1(\bar{X}_t)\right] dt, \\ d\bar{Y}_t =& -\alpha_2 \bar{Y}_t dt + \mathbb{E}\left[f_1(\bar{X}_t)\right] dt - \int_{\mathbb{R}_+} \bar{Y}_{t-} \mathbb{1}_{\{z \leq f_2(\bar{Y}_{t-})\}} d\bar{\pi}^1(t,z),\end{aligned}$$

where $\bar{\pi}^1$ is a Poisson measure with Lebesgue intensity. Note that the equation of $\bar{X}$ above is deterministic, as long as the initial condition $\bar{X}_0$ is.

With classical computation (see e.g. Delattre, Fournier and Hoffmann (2016), Fournier and Löcherbach (2016) and Ditlevsen and Löcherbach (2017)) and under the assumption that the coefficients of the SDEs are Lipschitz continuous, we can prove that for all $t \geq 0$,

$$\mathbb{E}\left[\sup_{0 \leq s \leq t}\left|X_s^N - \bar{X}_s\right|\right] + \mathbb{E}\left[\sup_{0 \leq s \leq t}\left|Y_s^N - \bar{Y}_s\right|\right] \leq C_t\left(N^{-1/2} + \mathbb{E}\left[|X_0^N - \bar{X}_0|\right] + \mathbb{E}\left[|Y_0^N - \bar{Y}_0|\right]\right).$$

Figure 2 presents two simulations of the process Y^N with $N = 100$ and $N = 1000$.

Note that in this model, a result similar to the one above above holds true even without "averaging" the effect of the first population. More precisely, let us introduce $X^{N,i}$ the membrane potential of the $i-$th neuron of the first population $(2 \leq i \leq N+1)$, and Y^N that of the neuron of the second

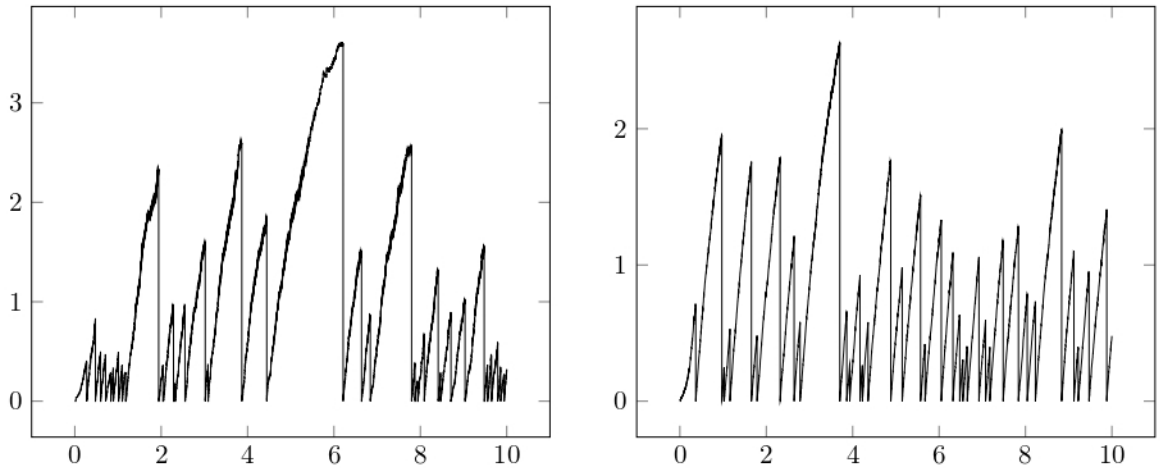

FIGURE 2. *Simulation of trajectories of* $(Y_t^N)_{0\le t\le 10}$ *with* $X_0^N = 0$, $Y_0^N = 0$, $\alpha_1 = \alpha_2 = 1$, $\nu = \mathcal{N}(0,1)$, $f_1(x) = f_2(x) = 2.2 + 1.4\arctan(10x-2)$, $N = 100$ *(left picture) and* $N = 1000$ *(right picture).*

one. Assume that these processes are solutions to

$$dX_t^{N,i} = -\alpha_1 X_t^{N,i}dt + \frac{1}{N}\sum_{\substack{j=2\\ j\neq i}}^{N+1}\int_{\mathbb{R}_+}\mathbb{1}_{\left\{z\le f_1(X_{t-}^{N,j})\right\}}d\pi^j(t,z) - X_{t-}^{N,i}\int_{\mathbb{R}_+}\mathbb{1}_{\left\{z\le f_1(X_{t-}^{N,i})\right\}}d\pi^i(t,z),$$

$$dY_t^N = -\alpha_2 Y_t^N dt + \frac{1}{N}\sum_{j=2}^{N+1}\int_{\mathbb{R}_+}\mathbb{1}_{\left\{z\le f_1(X_{t-}^{N,j})\right\}}d\pi^j(t,z) - Y_{t-}^N\int_{\mathbb{R}_+}\mathbb{1}_{\left\{z\le f_2(Y_{t-}^N)\right\}}d\pi^1(t,z),$$

where $X_0^{N,i}$ $(2\le i\le N+1)$ are exchangeable.

Then one can show that these processes converge respectively to $\bar{X}^i$ $(2\le i\le N+1)$ and $\bar{Y}$:

$$d\bar{X}_t^i = -\alpha_1\bar{X}_t^i + \mathbb{E}\left[f_1(\bar{X}_t^i)\right]dt - \int_{\mathbb{R}_+}\bar{X}_{t-}^i\mathbb{1}_{\left\{z\le f_1(\bar{X}_{t-}^i)\right\}}d\pi^i(t,z), \quad 2\le i\le N+1,$$

$$d\bar{Y}_t = -\alpha_2\bar{Y}_t + \mathbb{E}\left[f_1(\bar{X}_t^1)\right]dt - \int_{\mathbb{R}_+}\bar{Y}_{t-}\mathbb{1}_{\left\{z\le f_2(\bar{Y}_{t-})\right\}}d\bar{\pi}^1(t,z)$$

in the following sense: for all $t\ge 0$,

$$\frac{1}{N}\sum_{i=2}^{N+1}\mathbb{E}\left[\sup_{0\le s\le t}\left|\bar{X}_s^i - X_s^{N,i}\right|\right] + \mathbb{E}\left[\sup_{0\le s\le t}\left|\bar{Y}_s - Y_s^N\right|\right]$$
$$\le C_t\left(N^{-1/2} + \frac{1}{N}\sum_{i=2}^{N+1}\mathbb{E}\left[|\bar{X}_0^i - X_0^{N,i}|\right] + \mathbb{E}\left[\bar{Y}_0 - Y_0^N\right]\right).$$

4. Technical lemmas

The first lemma allows to differentiate under an integral sign. Contrary to the classical result that relies on the dominated convergence theorem, this result relies on Vitali convergence theorem. In other words, we use a uniform integrability criterion instead of a domination condition.

Lemma 4.1. *Let $F : (x, \omega) \in \mathbb{R} \times \Omega \mapsto \mathbb{R}_+$ be a measurable function, where $(\Omega, \mathbb{P})$ is a probability space. Assume that:*

- *$\mathbb{P}-$almost surely, $x \mapsto F(x, \omega)$ is C^1,*
- *there exists $p > 1$ such that*

$$\sup_{x\in\mathbb{R}} \int_\Omega |\partial_1 F(x, \omega)|^p \, d\mathbb{P}(\omega) < \infty.$$

Then, $\mathbb{P}-$almost surely, $x \mapsto \int_\Omega F(x, \omega) d\mathbb{P}(\omega)$ is C^1 and its derivative is

$$\int_\Omega \partial_1 F(x, \omega) d\mathbb{P}(\omega).$$

Proof. Let $(x_n)_n$ be a sequence of real numbers that converges to some $x \in \mathbb{R}$. Let us write

$$G_n(\omega) := \frac{F(x, \omega) - F(x_n, \omega)}{x - x_n}.$$

We have

$$|G_n(\omega)| = \left| \frac{1}{x - x_n} \int_{x_n}^x \partial_1 F(y, \omega) dy \right| \leq \frac{1}{|x - x_n|} \int_{x_n}^x |\partial_1 F(y, \omega)| dy.$$

Then, by Jensen's inequality and Fubini-Tonelli's theorem,

$$\begin{aligned}\int_\Omega |G_n(\omega)|^p \, d\mathbb{P}(\omega) &\leq \frac{1}{|x - x_n|} \int_\Omega \int_{x_n}^x |\partial_1 F(y, \omega)|^p \, dy d\mathbb{P}(\omega) \\ &\leq \frac{1}{|x - x_n|} \int_{x_n}^x \int_\Omega |\partial_1 F(y, \omega)|^p \, d\mathbb{P}(\omega) dy \leq \sup_{y\in\mathbb{R}} \int_\Omega |\partial_1 F(y, \omega)|^p \, d\mathbb{P}(\omega).\end{aligned}$$

As the previous bound is uniform in n, this implies that the family of random variables $\{G_n : n \in \mathbb{N}\}$ is uniformly integrable. Since $\mathbb{P}-$almost surely, $G_n(\omega)$ converges to $\partial_1 F(x, \omega)$ as n goes to infinity, Vitali convergence theorem allows to prove that $\int_\Omega G_n(\omega) d\mathbb{P}(\omega)$ converges to $\int_\Omega \partial_1 F(x, \omega) d\mathbb{P}(\omega)$ as N goes to infinity. □

The next lemma allows to obtain a priori estimates on solutions of certain SDEs. This result is not written to be general, because some specific properties of our SDEs allows to do some simplifications to obtain the result.

Lemma 4.2. *Let $(X_t, Y_t)_t$ be a solution of the following two-dimensional SDE*

$$\begin{aligned}dX_t =& b_1(X_t, Y_t)dt + \varsigma_1(X_t, Y_t)dW_t, \\ dY_t =& b_2(X_t, Y_t)dt + \varsigma_2(X_t, Y_t)dW_t - \int_{[0,+\infty[} Y_{t-} \mathbb{1}_{\{z \leq f(Y_{t-})\}} d\pi(z, t),\end{aligned}$$

where $(W_t)_t$ is one-dimensional standard Brownian motion, π a Poisson measure on $\mathbb{R}_+^2$ with Lebesgue intensity and f a non-negative measurable function. Assume that the functions b_1, b_2, ς_1 and ς_2 are sublinear: for all $x, y \in \mathbb{R}$,

$$|b_1(x,y)| + |b_2(x,y)| + |\varsigma_1(x,y)| + |\varsigma_2(x,y)| \leq C(1+|x|+|y|).$$

Then:

(i) for any even $p \in \mathbb{N}$ such that X_0 and Y_0 both have a finite $p-$order moment, for all $t \geq 0$,

$$\mathbb{E}\left[X_t^p\right] + \mathbb{E}\left[Y_t^p\right] \leq C_{t,p}(1 + \mathbb{E}\left[X_0^p\right] + \mathbb{E}\left[Y_0^p\right]),$$

with $C_{t,p} > 0$ some constant,

(ii) if X_0 and Y_0 both have a finite $2p-$th moment (for some even $p \in \mathbb{N}^*$), for all $t \geq 0$,

$$\mathbb{E}\left[\sup_{0\leq s\leq t} X_s^p\right] + \mathbb{E}\left[\sup_{0\leq s\leq t} Y_s^p\right] \leq C_{t,p}\left(1 + \mathbb{E}\left[X_0^{2p}\right] + \mathbb{E}\left[Y_0^{2p}\right]\right),$$

with $C_{t,p} > 0$ another constant.

Proof. Step 1. Let us prove the point (i). By Ito's formula, for all $t \geq 0$,

$$X_t^p = X_0^p + p\int_0^t X_s^{p-1} b_1(X_s,Y_s)ds + \frac{p(p-1)}{2}\int_0^t X_s^{p-2}\varsigma_1(X_s,Y_s)^2 ds + p\int_0^t X_s^{p-1}\varsigma_1(X_s,Y_s)dW_s, \tag{7}$$

$$\begin{aligned} Y_t^p =& Y_0^p + p\int_0^t Y_s^{p-1} b_2(X_s,Y_s)ds + \frac{p(p-1)}{2}\int_0^t Y_s^{p-2}\varsigma_2(X_s,Y_s)^2 ds + p\int_0^t Y_s^{p-1}\varsigma_2(X_s,Y_s)dW_s \\ &- \int_{[0,t]\times[0,+\infty[} Y_{s-}^p \mathbb{1}_{\{z\leq f(Y_{s-})\}} d\pi(s,z) \\ \leq& Y_0^p + p\int_0^t Y_s^{p-1} b_2(X_s,Y_s)ds + \frac{p(p-1)}{2}\int_0^t Y_s^{p-2}\varsigma_2(X_s,Y_s)^2 ds + p\int_0^t Y_s^{p-1}\varsigma_2(X_s,Y_s)dW_s. \end{aligned} \tag{8}$$

Now for $M > 0$, let us introduce the following stopping time

$$\tau_M := \inf\{t > 0 \ : \ |X_t| > M \text{ or } |Y_t| > M\},$$

and the function

$$u_{M,p}(t) := \mathbb{E}\left[(X_{t\wedge\tau_M})^p + (Y_{t\wedge\tau_M})^p\right].$$

Thanks to (7), we have for all $t \geq 0$,

$$\begin{aligned} &\mathbb{E}\left[(X_{t\wedge\tau_M})^p\right] = \mathbb{E}\left[X_0^p\right] + \mathbb{E}\left[\int_0^{t\wedge\tau_M}\left(pX_s^{p-1}b_1(X_s,Y_s) + \frac{p(p-1)}{2}X_s^{p-2}\varsigma_1(X_s,Y_s)^2\right)ds\right] \\ &= \mathbb{E}\left[X_0^p\right] + \mathbb{E}\left[\int_0^{t\wedge\tau_M}\left(pX_{s\wedge\tau_M}^{p-1}b_1(X_{s\wedge\tau_M},Y_{s\wedge\tau_M}) + \frac{p(p-1)}{2}X_s^{p-2}\varsigma_1(X_{s\wedge\tau_M},Y_{s\wedge\tau_M})^2\right)ds\right] \\ &\leq \mathbb{E}\left[X_0^p\right] + C\int_0^t \mathbb{E}\left[|X_{s\wedge\tau_M}|^{p-1}\cdot|b_1(X_{s\wedge\tau_M},Y_{s\wedge\tau_M})| + X_{s\wedge\tau_M}^{p-2}\varsigma_1(X_{s\wedge\tau_M},Y_{s\wedge\tau_M})^2\right]ds \\ &\leq \mathbb{E}\left[X_0^p\right] + C\int_0^t (1 + u_{M,p}(s))\,ds, \end{aligned}$$

where we have used the sublinearity of the functions b_1 and ς_1 to obtain the last line. With the same reasoning, we also have using (8) for all $t \geq 0$,

$$\mathbb{E}\left[(Y_{t\wedge\tau_M})^p\right] \leq \mathbb{E}\left[Y_0^p\right] + C\int_0^t \left(1 + u_{M,p}(s)\right) ds.$$

So, for all $t \geq 0$,

$$u_{M,p}(t) \leq \mathbb{E}\left[X_0^p\right] + \mathbb{E}\left[Y_0^p\right] + Ct + C\int_0^t u_{M,p}(s)ds.$$

Then, as the function $u_{M,p}$ is finite valued (thanks to the stopping time τ_M), Grönwall's lemma implies that for all $t \geq 0$,

$$u_{M,p}(t) \leq \left(\mathbb{E}\left[X_0^p\right] + \mathbb{E}\left[Y_0^p\right] + Ct\right) e^{Ct} \leq C_t(1 + \mathbb{E}\left[X_0^p\right] + \mathbb{E}\left[Y_0^p\right]), \tag{9}$$

where $C_t > 0$ does not depend on M.

Now let us prove that τ_M goes to infinity almost surely as M goes to infinity. Firstly, as $(\tau_M)_M$ is non-decreasing w.r.t. M, we know that τ_M has an almost sure limit τ as M goes to infinity. We have, for all $t > 0$, by Markov's inequality and thanks to (9),

$$\mathbb{P}\left(\tau_M \leq t\right) \leq \mathbb{P}\left(X^p_{t\wedge\tau_M} + Y^p_{t\wedge\tau_M} \geq M^p\right) \leq \frac{1}{M^p} u_{M,p}(t) \underset{M\to\infty}{\longrightarrow} 0.$$

Consequently, for all $t > 0$,

$$\mathbb{P}\left(\tau \leq t\right) \leq \mathbb{P}\left(\bigcap_{M\in\mathbb{N}^*} \{\tau_M \leq t\}\right) = \lim_{M\to\infty} \mathbb{P}\left(\tau_M \leq t\right) = 0.$$

Whence

$$\mathbb{P}\left(\tau < \infty\right) \leq \mathbb{P}\left(\bigcup_{t\in\mathbb{N}^*} \{\tau \leq t\}\right) = 0.$$

This means that τ_M goes to infinity almost surely as M goes to infinity.

Finally, applying Fatou's lemma to (9) gives that for all $t \geq 0$,

$$\mathbb{E}\left[X_t^p + Y_t^p\right] \leq C_t(1 + \mathbb{E}\left[X_0^p\right] + \mathbb{E}\left[Y_0^p\right]).$$

Step 2. Let us prove the point (ii). By (7) and Burkhölder-Davis-Gundy's inequality,

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\leq s\leq t} X_s^p\right] \leq& \mathbb{E}\left[X_0\right]^p + p\int_0^t \mathbb{E}\left[|X_s^{p-1} b_1(X_s, Y_s)|\right] ds + \frac{p(p-1)}{2}\int_0^t \mathbb{E}\left[X_s^{p-2}\varsigma_1(X_s, Y_s)^2\right] ds \\
&+ Cp\left(\int_0^t \mathbb{E}\left[X_s^{2p-2}\varsigma_1(X_s, Y_s)^2\right] ds\right)^{1/2} \\
\leq& \mathbb{E}\left[X_0^2\right] + C\int_0^t \left(1 + \mathbb{E}\left[X_s^p\right] + \mathbb{E}\left[Y_s^p\right]\right) ds + C\left(\int_0^t \left(1 + \mathbb{E}\left[X_s^{2p}\right] + \mathbb{E}\left[Y_s^{2p}\right]\right) ds\right)^{1/2} \\
\leq& \mathbb{E}\left[X_0^p\right] + C\int_0^t \left(1 + \mathbb{E}\left[X_s^{2p}\right] + \mathbb{E}\left[Y_s^{2p}\right]\right) ds \\
\leq& C_t\left(1 + \mathbb{E}\left[X_0^{2p}\right] + \mathbb{E}\left[Y_0^{2p}\right]\right).
\end{aligned}$$

The result for the process Y follows from the same computation using (8) instead of (7). □

References

Bally, V., Goreac, D. and Rabiet, V. (2018). Regularity and Stability for the Semigroup of Jump Diffusions with State-Dependent Intensity. *Annals of Applied Probability* **28** 3028 – 3074.

Brémaud, P. (1981). *Point Processes and Queues, Martingale Dynamics*. Springer-Verlag.

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

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]*.

Erny, X., Löcherbach, E. and Loukianova, D. (2021). White-noise driven conditional McKean-Vlasov limits for systems of particles with simultaneous and random jumps. *arXiv:2103.04100 [math]*. arXiv: 2103.04100.

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.

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

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

Kunita, H. (1986). *Lectures on Stochastic Flows and Applications for the Indian Institute Of Science Bangalore*. Tata Institute Of Fundamental Research Springer-Verlag.

Löcherbach, E. (2002). Likelihood Ratio Processes for Markovian Particle Systems with Killing and Jumps. *Statistical Inference for Stochastic Processes* **5** 153–177.