



HAL
open science

Mean field limits for interacting Hawkes processes in a diffusive regime

Xavier Erny, Eva Löcherbach, Dasha Loukianova

► **To cite this version:**

Xavier Erny, Eva Löcherbach, Dasha Loukianova. Mean field limits for interacting Hawkes processes in a diffusive regime. 2019. hal-02096662v1

HAL Id: hal-02096662

<https://hal.science/hal-02096662v1>

Preprint submitted on 12 Apr 2019 (v1), last revised 23 Nov 2020 (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 limits for interacting Hawkes processes in a diffusive regime

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 consider a sequence of systems of Hawkes processes having mean field interactions in a diffusive regime. The stochastic intensity of each process is a solution of a stochastic differential equation driven by N independent Poisson random measures. We show that, as the number of interacting components N tends to infinity, this intensity converges in distribution in Skorohod space to a CIR-type diffusion. Moreover, we prove the convergence in distribution of the Hawkes processes to the limit point process having the limit diffusion as intensity. To prove the convergence results, we use analytical technics based on the convergence of the associated infinitesimal generators and Markovian semigroups.

MSC 2010 subject classifications: 60K35, 60G55, 60J35.

Keywords and phrases: Multivariate nonlinear Hawkes processes, Mean field interaction, Piecewise deterministic Markov processes.

Introduction

Hawkes processes were originally introduced by Hawkes (1971) to model the appearance of earthquakes in Japan. Since then these processes have been successfully used in many fields to model various physical, biological or economical phenomena exhibiting self-excitation or -inhibition and interaction, such as seismology (Helmstetter and Sornette (2002), Y. Kagan (2009), Ogata (1999), Bacry and Muzy (2016)), financial contagion (Aït-Sahalia, Cacho-Diaz and Laeven (2015)), high frequency financial order books arrivals (Lu and Abergel (2018), Bauwens and Hautsch (2009), Hewlett (2006)), genome analysis (Reynaud-Bouret and Schbath (2010)) and interactions in social networks (Zhou, Zha and Song (2013)). In particular, multivariate Hawkes processes are extensively used in neurosciences to model temporal arrival of spikes in neural network (Grün, Diedsmann and Aertsen (2010), Okatan, A Wilson and N Brown (2005), Pillow, Wilson and Brown (2008), Reynaud-Bouret et al. (2014)) since they provide good models to describe the typical temporal decorrelations present in spike trains of the neurons as well as the functional connectivity in neural nets.

In this paper, we consider a sequence of multivariate Hawkes processes $(Z^N)_{N \in \mathbb{N}^*}$ of the form $Z^N = (Z_t^{N,1}, \dots, Z_t^{N,N})_{t \geq 0}$. Each Z^N is designed to describe the behaviour of some interacting system with N components, as for example a neural network of N neurons. This is a multivariate counting process where each $Z^{N,i}$ records the number of events related to the i -th component, as for example the number of spikes of the i -th neuron. These counting processes are interacting, that is, any event of type i is able to trigger or to inhibit future events of all other types j . The process $(Z^{N,1}, \dots, Z^{N,N})$ is informally defined via its stochastic intensity process $\lambda^N = (\lambda^{N,1}(t), \dots, \lambda^{N,N}(t))_{t \geq 0}$ through the relation

$$\mathbb{P}(Z^{N,i} \text{ has a jump in }]t, t + dt] | \mathcal{F}_t) = \lambda^{N,i}(t)dt, \quad 1 \leq i \leq N,$$

where $\mathcal{F}_t = \sigma(Z_s^N : 0 \leq s \leq t)$. The stochastic intensity of a Hawkes process is given by

$$\lambda^{N,i}(t) = f_i^N \left(\sum_{j=1}^N \int_{-\infty}^t h_{ij}^N(t-s) dZ^{N,j}(s) \right). \quad (1)$$

Here, h_{ij}^N models the action or the influence of events of type j on those of type i , and how this influence decreases as time goes by. The function f_i^N is called the jump rate function of $Z^{N,i}$.

Since the founding works of Hawkes (1971) and Hawkes and Oakes (1974), many probabilistic properties of Hawkes processes have been well-understood, such as ergodicity, stationarity and long time behaviour (see Brémaud and Massoulié (1996), Daley and Vere-Jones (2003) and Costa et al. (2018)). A number of authors studied the statistical inference for Hawkes processes (Ogata (1978) and Reynaud-Bouret and Schbath (2010)). Another field of study, really active nowadays, concerns the behaviour of the Hawkes process when the number of components N goes to infinity. During the last decade, large population limits of systems of interacting Hawkes processes have been studied both in discrete and continuous time (Fournier and Löcherbach (2016), Delattre, Fournier and Hoffmann (2016), Ditlevsen and Löcherbach (2017)).

Delattre, Fournier and Hoffmann (2016) consider a general class of Hawkes processes whose interactions are given by a graph. In the case where the interactions are of mean field type and scaled in N^{-1} , namely $h_{ij}^N = N^{-1}h$ and $f_i^N = f$ in (1), they show that the Hawkes processes can be approximated by an i.i.d. family of inhomogeneous Poisson processes. They observe that for each fixed integer k , the joint law of k components converges to a product law as N tends to infinity, which is commonly referred to as the propagation of chaos. Ditlevsen and Löcherbach (2017) generalize this result to a multi-population frame and show how oscillations emerge in the large population limit. Note again that the interactions in both papers are scaled in N^{-1} , which leads to limit point processes with deterministic intensity.

The purpose of this paper is to study the large population limit (when N goes to infinity) of the multivariate Hawkes processes $(Z^{N,1}, \dots, Z^{N,N})$ with mean field interactions scaled in $N^{-1/2}$. Contrarily to the situation considered in Delattre, Fournier and Hoffmann (2016) and Ditlevsen and Löcherbach (2017), this scaling leads to a non-chaotic limiting process with stochastic intensity. As we consider interactions scaled in $N^{-1/2}$, we have to center the terms of the sum in (1) to make the intensity process converge according to some kind of central limit theorem. To this end, we consider intensities with stochastic jump heights of the form

$$\lambda^{N,i}(t) = \lambda^N(t) = f \left(\frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{-\infty}^t h(t-s) U_j(s) dZ_s^{N,j} \right),$$

where the variables $U_j(s)$ are i.i.d. and centered.

Moreover we consider functions h of the form $h(t) = e^{-\alpha t}$ so that the process $(X_t^N)_t$ defined by

$$X_t^N := \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{-\infty}^t e^{-\alpha(t-s)} U_j(s) dZ^{N,j}(s)$$

is a piecewise deterministic Markov process. In the framework of neurosciences, X_t^N represents the membrane potential of the neurons at time t , the variables $U_j(s)$ model random synaptic weights and the jumps of $Z^{N,j}$ represent the spike times of neuron j . If neuron j spikes at time t , an

additional random potential height $U_j(t)/\sqrt{N}$ is given to all other neurons in the system. As a consequence, the process X^N has the following dynamics

$$dX_t^N = -\alpha X_t^N dt + \frac{1}{\sqrt{N}} \sum_{i=1}^N U_i(t) dZ^{N,i}(t),$$

and the infinitesimal generator of X^N is given by

$$A^N g(x) = -\alpha x g'(x) + N f(x) \int \left[g\left(x + \frac{u}{\sqrt{N}}\right) - g(x) \right] \mu(du),$$

for sufficiently smooth functions g , where μ is the common distribution of the variables $U_j(s)$.

As N goes to infinity, the above expression converges to

$$\bar{A}g(x) = -\alpha x g'(x) + \frac{1}{2} f(x) g''(x),$$

which is the generator of a CIR-type diffusion solution of the SDE

$$d\bar{X}_t = -\alpha \bar{X}_t dt + \sqrt{f(\bar{X}_t)} dB_t. \quad (2)$$

We show that the convergence of the generators implies the convergence of X^N to \bar{X} in distribution in Skorohod space, together with a control on the speed of convergence. Moreover we establish for each i , the convergence in distribution in Skorohod space of the associated counting process $Z^{N,i}$ to the limit counting process \bar{Z}^i which has intensity $(f(\bar{X}_t))_t$. Conditionally on \bar{X} , the $\bar{Z}^i, i \geq 1$, are independent. This property can be viewed as a conditional propagation of chaos-property, which has to be compared to [Delattre, Fournier and Hoffmann \(2016\)](#) and [Ditlevsen and Löcherbach \(2017\)](#) where the intensity of the limit process is deterministic and its components are truly independent, and to [Carmona, Delarue and Lacker \(2016\)](#) where all interacting components are subject to common noise. In our case, the common noise, that is, the Brownian motion B of (2), emerges in the limit as a consequence of the central limit theorem.

To the best of our knowledge, this is the first result of diffusion limit type for multivariate Hawkes processes.

The convergence in distribution of X^N to \bar{X} ([Theorem 2.1](#)) is obtained by showing first the tightness of the sequence $(X^N)_N$ on Skorohod space, and then the convergence in finite-dimensional distribution. To prove the finite-dimensional convergence we use analytical methods showing first the convergence of the generators from which we deduce the convergence of the semigroups via the formula

$$\bar{P}_t g(x) - P_t^N g(x) = \int_0^t P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x) ds. \quad (3)$$

Here $\bar{P}_t g(x) = \mathbb{E}_x [g(\bar{X}_t)]$ and $P_t^N g(x) = \mathbb{E}_x^N [g(X_t^N)]$ denote the Markovian semigroups of \bar{X} and X^N . This formula is well-known in the classical semigroup theory setting where the generators are strong derivatives of semigroups in the Banach space of continuous bounded functions (see Lemma 1.6.2 of [Ethier and Kurtz \(2005\)](#)). In our case, we have to consider extended generators (see [Davis \(1993\)](#) or [Meyn and Tweedie \(1993\)](#)), i.e. $A^N g(x)$ is the point-wise derivative of $t \mapsto P_t^N g(x)$. The version of formula (3) for extended generators is stated and proved in [Appendix \(Proposition 6.3\)](#).

It is well-known that under suitable assumptions on f , the solution of (2) admits a unique invariant measure π whose density is explicitly known. Thus, a natural question is to consider the limit of the law of X_t^N when t and N go simultaneously to infinity. We prove that under appropriate conditions on the way N and t tend jointly to infinity, this limit is π , and we provide a control of the error (Theorem 2.3). This result can be viewed as an approximation result of the finite size and finite time particle system by the invariant measure π , that is, a simulation algorithm to simulate the law of X_t^N from the invariant law π .

The paper is organized as follow: in Section 1, we introduce the model rigorously and state the assumptions. In Section 2, we formulate the main results. Section 3 is devoted to the proof of the convergence of X^N to \bar{X} (Theorem 2.1), and Section 4 to the study of the simultaneous t and N limit of the law of X_t^N (Theorem 2.3). In Section 5, we prove the convergence of the point processes $Z^{N,i}$ to \bar{Z}^i (Theorem 2.5). Finally in Appendix, we prove some results on the extended generators, and some other technical results that we use throughout the paper.

1. Notation, model and assumptions

1.1. Notation

The following notation are used throughout the paper:

- If X is a random variable, we note $\mathcal{D}(X)$ its distribution.
- If g is a real-valued function which is n times differentiable, we note $\|g\|_{n,\infty} = \sum_{k=0}^n \|g^{(k)}\|_{\infty}$.
- We write $C_b^n(\mathbb{R})$ for the set of the functions g which are n times continuously differentiable such that $\|g\|_{n,\infty} < +\infty$, and we write $C_b^0(\mathbb{R})$ for $C_b^0(\mathbb{R})$.
- If g is a real-valued function and I is an interval, we note $\|g\|_{\infty,I} = \sup_{x \in I} |g(x)|$.
- We write $C_c^n(\mathbb{R})$ for the set of functions that are n times continuously differentiable and that have a compact support.
- We write $D(\mathbb{R}_+, \mathbb{R})$ for the Skorohod space of càdlàg functions from \mathbb{R}_+ to \mathbb{R} , endowed with Skorohod metric (see Chapter 3 Section 16 of Billingsley (1999)). Moreover, $D(\mathbb{R}_+, \mathbb{R}^k)$ denotes the space of \mathbb{R}^k -valued càdlàg functions endowed with the topology that generalizes naturally the topology of $D(\mathbb{R}_+, \mathbb{R})$ (see e.g. Section 3.5 of Ethier and Kurtz (2005)).
- If E is a Polish space, $\mathcal{M}^\#(E)$ denotes the space of locally finite measures on E endowed with the topology of the weak convergence. With this topology, $\mathcal{M}^\#(E)$ is a Polish space (see Theorem A2.6.III of Daley and Vere-Jones (2003)). In this paper, we consider either $E = \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ or $E = \mathbb{R}_+ \times \mathbb{R}_+$, and we write $\mathcal{M}^\#$ for $\mathcal{M}^\#(E)$. It will always be clear which space E we consider.
- W_p denotes the Wasserstein metric of order p , that is, $W_p(\nu_1, \nu_2) = \inf_{X \sim \nu_1, Y \sim \nu_2} \mathbb{E}[|X - Y|^p]^{1/p}$.
- α is a positive constant, L, A, B, σ are fixed parameters defined in Assumptions 1 and 2 below, m_k ($1 \leq k \leq 4$) are fixed parameters introduced in Assumption 3 below, $C_t, D_t, K_t, Q_t^{(k)}$ are constants that depend on t and the previous parameters, which are defined in Lemma 3.1, Theorem 2.2 and Proposition 3.6. Finally, we note Γ any arbitrary constant, so the value of Γ can change from line to line in an equation. Moreover, if Γ depends on some non-fixed parameter θ , we write Γ_θ .

1.2. The model

We consider a sequence of multivariate Hawkes processes $(Z^{N,i})_{1 \leq i \leq N, N \in \mathbb{N}^*}$ of the form

$$Z_t^{N,i} = \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} \mathbb{1}_{\{z \leq f(X_{s-}^N)\}} d\pi_i(s, z, u), \quad 1 \leq i \leq N, \quad (4)$$

where $(\pi_i)_{i \in \mathbb{N}^*}$ are i.i.d. Poisson random measures on $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ of intensity $dt dz d\mu(u)$, μ is a centered probability measure on \mathbb{R} , and $(X_t^N)_{t \in \mathbb{R}_+}$ is given by

$$\begin{cases} X_t^N = X_0^N e^{-\alpha t} + \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} u e^{-\alpha(t-s)} \mathbb{1}_{\{z \leq f(X_{s-}^N)\}} d\pi_j(s, z, u), \\ X_0^N \sim \nu_0^N, \end{cases} \quad (5)$$

where ν_0^N is a probability measure on \mathbb{R} . Notice that X^N is solution of the following SDE

$$dX_t^N = -\alpha X_t^N dt + \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{(z,u) \in \mathbb{R}_+ \times \mathbb{R}} u \mathbb{1}_{\{z \leq f(X_{t-}^N)\}} d\pi_j(t, z, u). \quad (6)$$

Under natural assumptions on f , this SDE (6) admits a unique non-exploding strong solution. This will be proved in Proposition 6.6. In particular, X^N is a piecewise deterministic Markov process.

The aim of this paper is to show that X^N converges in Skorokhod space to the limit process $(\bar{X}_t)_{t \in \mathbb{R}_+}$ which is solution to the SDE

$$\begin{cases} d\bar{X}_t = -\alpha \bar{X}_t dt + \sigma \sqrt{f(\bar{X}_t)} dB_t, \\ \bar{X}_0 \sim \bar{\nu}_0, \end{cases} \quad (7)$$

where σ^2 is the variance of μ , $(B_t)_{t \in \mathbb{R}_+}$ is a one-dimensional standard Brownian motion, and $\bar{\nu}_0$ a suitable probability measure on \mathbb{R} .

In the sequel, we will prove the convergence of X^N to \bar{X} , and we will derive some consequences of this convergence.

To prove our results, we need to introduce the following assumptions.

Assumption 1. \sqrt{f} is a positive and Lipschitz continuous function, having Lipschitz constant L . In particular there exist some constants A and B such that for all $x \in \mathbb{R}$, $f(x) \leq Ax^2 + B$.

Remark 1.1. Obviously, we have $A \leq 2L^2$ and $B \leq 2f(0)^2$, so we could fix $A = 2L^2$ and $B = 2f(0)^2$. However, these choices for A and B are not optimal in general, and using generic constants A and B makes the proofs more readable.

Under Assumption 1, it is classical that the SDE (7) admits a unique non-exploding strong solution (see remark IV.2.1, Theorems IV.2.3, IV.2.4 and IV.3.1 of Ikeda and Watanabe (1989)).

Assumption 2.

- $\int_{\mathbb{R}} x^4 d\bar{\nu}_0(x) < \infty$ and $\sup_{N \in \mathbb{N}^*} \int_{\mathbb{R}} x^4 d\nu_0^N(x) < \infty$.
- μ is a centered probability measure having a fourth moment, we note σ^2 its variance.

Assumption 2 allows us to control the moments up to order four of the processes $(X_t^N)_t$ and $(\bar{X}_t)_t$ (see Lemma 3.1) and to prove the convergence of the generators of the processes $(X_t^N)_t$ (see Proposition 3.5).

Assumption 3. We assume that f is C^4 and for each $1 \leq k \leq 4$, $(\sqrt{f})^{(k)}$ is bounded by some constant m_k .

Assumption 3 guarantees that the stochastic flow associated to (7) has regularity properties with respect to the initial condition $\bar{X}_0 = x$. This will be the main tool to obtain uniform in time estimates of the limit semigroup, see Proposition 3.6.

Example 1.2. The functions $f(x) = 1 + x^2$ and $f(x) = \sqrt{1 + x^2}$ satisfy Assumptions 1 and 3.

Assumption 4. $W_2(\bar{\nu}_0, \nu_0^N)$ vanishes when N goes to infinity.

The convergence of X_0^N to \bar{X}_0 in distribution is a necessary condition for the convergence of the process X^N to \bar{X} . In Proposition 3.9 below establishing the finite dimensional convergence of X^N to \bar{X} , we rely on Assumption 4 which is a bit stronger. Actually if we assume that the first part of Assumption 2 holds, then Assumption 4 is equivalent to the convergence in distribution of X_0^N to \bar{X}_0 .

2. Main results

Our first main result is the convergence of the process X^N to \bar{X} in distribution in Skorohod space.

Theorem 2.1. If Assumptions 1, 2, 3 and 4 hold, then the sequence $(X^N)_N$ converges in distribution to \bar{X} in $D(\mathbb{R}_+, \mathbb{R})$.

Theorem 2.1 is proved in the end of Subsection 3.4. Below we give some simulations of the trajectories of the process $(X_t^N)_{t \geq 0}$ in Figure 1.

Actually, we have more details than just the convergence of X^N to \bar{X} in Skorohod space. Indeed, we are able to establish the rate of convergence of $P_t^N g(x)$ to $\bar{P}_t g(x)$, uniformly in time for $t \in [0, T]$, for sufficiently smooth test-functions g .

Theorem 2.2. If Assumptions 1, 2 and 3 hold, then for all $T \geq 0$, there exists a positive constant K_T such that for each $g \in C_b^3(\mathbb{R})$,

$$\sup_{0 \leq t \leq T} |P_t^N g(x) - \bar{P}_t g(x)| \leq (1 + x^2) K_T \|g\|_{3, \infty} \frac{1}{\sqrt{N}}.$$

The constant K_T can be chosen of the form $K_T = O\left(T^2 \left(1 + e^{T(\sigma^2 A - 2\alpha)}\right) \left(1 + e^{T\beta^{(3)}}\right)\right)$ with $\beta^{(3)} = 12\sigma^2 m_1^2 + 3\sigma^2 m_2^2 - \alpha$.

We refer to Proposition 3.6 for the form of $\beta^{(3)}$. Theorem 2.2 is proved in the end of Subsection 3.3.

If the limit process \bar{X} is sufficiently ergodic (that is, if α is sufficiently large), having invariant probability measure π , then we can even control the speed of convergence of $P_t^N g(x)$ to $\pi(g)$, as t goes to infinity, for suitable choices of $N_t \rightarrow \infty$.

Theorem 2.3. Under Assumptions 1, 2 and 3, let $N : t \in \mathbb{R}_+ \mapsto N_t \in \mathbb{N}^*$ be some function such that one of the conditions below holds:

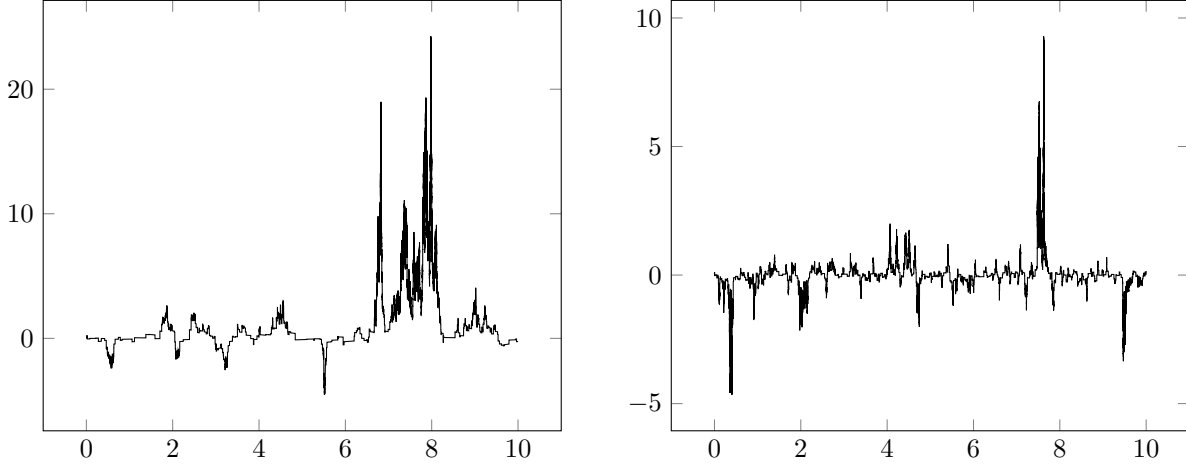


FIGURE 1. Simulation of trajectories of $(X_t^N)_{0 \leq t \leq 10}$ with $X_0^N = 0$, $\alpha = 1$, $\mu = \mathcal{N}(0, 1)$, $N = 10$ (left picture) and $N = 50$ (right picture).

- (a) $\alpha > \sigma^2 \max(L^2/2, A/2, 12m_1^2 + 3m_2^2)$ and $t^6 = o(N_t)$.
(b) $\alpha > \frac{\sigma^2}{2} \max(L^2, A)$ and $t^6 e^{2\beta^{(3)}t} = o(N_t)$.

- (1) Then \bar{X} is uniquely ergodic, having invariant probability measure π .
(2) $P_t^{N_t}(x, \cdot)$ converges weakly to π when t goes to infinity. Besides, for each condition, the speed of convergence for test functions $g \in C_b^3(\mathbb{R})$ is given by:

$$(a) \left| P_t^{N_t}(x, \cdot)g - \pi g \right| \leq \Gamma \|g\|_{3, \infty} \left(\frac{t^2}{\sqrt{N_t}} (1 + x^2 + t) + e^{(\frac{1}{2}\sigma^2 L^2 - \alpha)t} \sqrt{1 + x^2 + t} \right)$$

$$(b) \left| P_t^{N_t}(x, \cdot)g - \pi g \right| \leq \Gamma \|g\|_{3, \infty} \left(\frac{t^2 e^{\beta^{(3)}t}}{\sqrt{N_t}} (1 + x^2 + t) + e^{(\frac{1}{2}\sigma^2 L^2 - \alpha)t} \sqrt{1 + x^2 + t} \right)$$

where Γ is a positive constant.

Theorem 2.3 is proved in the end of Section 4.

Remark 2.4. Formulae (a) and (b) can be seen as a simulation algorithm of the state of the finite particle system of size N_t at time t by the invariant state of the limit process.

Finally, using Theorem 2.1, we show the convergence of the point processes $Z^{N_t, i}$ defined in (4) to limit point processes \bar{Z}^i having stochastic intensity $f(\bar{X}_t)$ at time t . To define the processes \bar{Z}^i ($i \in \mathbb{N}^*$), we fix a Brownian motion $(B_t)_{t \geq 0}$ on some probability space different from the one where the processes X^N ($N \in \mathbb{N}^*$) and the Poisson random measures π_i ($i \in \mathbb{N}^*$) are defined. Then we fix a family of i.i.d. Poisson random measures $\bar{\pi}_i$ ($i \in \mathbb{N}^*$) on the same space as $(B_t)_{t \geq 0}$, independent of $(B_t)_{t \geq 0}$. This independence property is natural (see Proposition 5.2), and it allows us to consider the joint distributions $(\bar{X}, \bar{\pi}_1, \dots, \bar{\pi}_k)$ for each fixed $k \geq 1$, where \bar{X} is defined as the solution of (7) driven by $(B_t)_{t \geq 0}$.

As the Poisson random measures $\bar{\pi}_i$ play the same role as π_i , we shall write π_i instead of $\bar{\pi}_i$ in the rest of the paper. Since π_i and $\bar{\pi}_i$ are not defined on the same space, there will not be any

ambiguity. The limit point processes \bar{Z}^i are then defined by

$$\bar{Z}_t^i = \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} \mathbb{1}_{\{z \leq f(\bar{X}_s)\}} d\pi_i(s, z, u). \quad (8)$$

Theorem 2.5. *Under Assumptions 1, 2, 3 and 4, for all $k \geq 1$, the sequence $(Z^{N,1}, Z^{N,2}, \dots, Z^{N,k})_N$ converges to $(\bar{Z}^1, \bar{Z}^2, \dots, \bar{Z}^k)$ in distribution in $D(\mathbb{R}_+, \mathbb{R}^k)$.*

Let us give a brief interpretation of the above result. Conditionally on \bar{X} , $\bar{Z}^1, \dots, \bar{Z}^k$ are independent. Therefore, the above result can be interpreted as a conditional propagation of chaos property (compare to [Carmona, Delarue and Lacker \(2016\)](#) dealing with the situation where all interacting components are subject to common noise). In our case, the common noise, that is, the Brownian motion B driving the dynamic of \bar{X} , emerges in the limit as a consequence of the central limit theorem. Theorem 2.5 is proved in the end of Section 5.

3. Convergence of $(X_t^N)_N$ in distribution in $D(\mathbb{R}_+, \mathbb{R})$

The goal of this section is to prove Theorem 2.1 and Theorem 2.2. To prove the convergence of the sequence $(X_t^N)_N$, we show in a first time that it is tight, and then the convergence in finite-dimensional distribution. For that purpose we establish the convergence of the generators and then the one of the semigroups.

We start with useful a priori bounds on the moments of X^N and \bar{X} .

Lemma 3.1. *Under Assumptions 1 and 2, for each $T > 0$ there exist some constants $C_T = O(e^{T \max(\sigma^2 A - 2\alpha, 0)})$ and $D_T = O(T e^{T \max(\sigma^2 A - 2\alpha, 0)})$ such that the following holds.*

- (i) for all $N \in \mathbb{N}^*$ and $t \in [0, T]$, $\mathbb{E} \left[(X_t^N)^2 \right] \leq C_T \mathbb{E} \left[(X_0^N)^2 \right] + D_T$,
- (ii) for all $t \in [0, T]$, $\mathbb{E} \left[(\bar{X}_t)^2 \right] \leq C_T \mathbb{E} \left[(\bar{X}_0)^2 \right] + D_T$,
- (iii) $\sup_{N \in \mathbb{N}^*} \mathbb{E} \left[\left(\sup_{0 \leq t \leq T} |X_t^N| \right)^2 \right] < +\infty$,
- (iv) $\mathbb{E} \left[\left(\sup_{0 \leq t \leq T} |\bar{X}_t| \right)^2 \right] < +\infty$,
- (v) for all $t \in [0, T]$ and $n \in \mathbb{N}^*$, $\sup_{0 \leq t \leq T} \mathbb{E} \left[(X_t^N)^4 \right] \leq \Gamma_T \left(1 + \mathbb{E} \left[(X_0^N)^4 \right] \right)$,
- (vi) for all $t \in [0, T]$, $\sup_{0 \leq t \leq T} \mathbb{E} \left[(\bar{X}_t)^4 \right] \leq \Gamma_T \left(1 + \mathbb{E} \left[(\bar{X}_0)^4 \right] \right)$.

We postpone the proof of Lemma 3.1 to Appendix.

3.1. Tightness of $(X_t^N)_N$ in $D(\mathbb{R}_+, \mathbb{R})$

We recall Aldous criterion (see for instance Theorem 16.9 of [Billingsley \(1999\)](#)) for tightness in Skorohod space.

Lemma 3.2. *Let $(Y^N)_N$ be a sequence of processes in $D(\mathbb{R}_+, \mathbb{R})$. We suppose that the two following conditions hold:*

$$(a) \forall T > 0, \forall \varepsilon > 0, \lim_{\delta \searrow 0} \overline{\lim}_{N \rightarrow \infty} \sup_{(S, S') \in A_{\delta, T}} \mathbb{P}(|Y_{S'}^N - Y_S^N| > \varepsilon) = 0,$$

where $A_{\delta, T}$ is the set of all pairs of stopping times (S, S') such that $0 \leq S \leq S' \leq S + \delta \leq T$,

$$(b) \forall T > 0, \lim_{K \nearrow \infty} \sup_N \mathbb{P} \left(\sup_{0 \leq t \leq T} |Y_t^N| \geq K \right) = 0.$$

Then the sequence $(Y^N)_N$ is tight on $D(\mathbb{R}_+, \mathbb{R})$.

Now we prove the tightness of the sequence of processes $(X^N)_N$ using Aldous criterion and Ito's isometry.

Proposition 3.3. *If Assumptions 1 and 2 hold, then $\{(X_t^N)_t : N \in \mathbb{N}^*\}$ is tight on $D(\mathbb{R}_+, \mathbb{R})$.*

Proof. Thanks to Lemma 3.1, the verification of the condition (b) of Aldous criterion is straightforward since,

$$\mathbb{P} \left(\sup_{0 \leq t \leq T} |X_t^N| \geq K \right) \leq \frac{1}{K} \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t^N| \right],$$

which goes to 0, uniformly in N as K goes to infinity.

Now, we check the condition (a). Let S, S' be stopping times such that $0 \leq S \leq S' \leq S + \delta \leq T$. Then

$$X_{S'}^N - X_S^N = -\alpha \int_S^{S'} X_r^N dr + \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{]S, S'] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_j(r, z, u).$$

We have

$$(X_{S'}^N - X_S^N)^2 \leq 2\alpha^2 \left(\int_S^{S'} X_r^N dr \right)^2 + \frac{2}{N} \left(\sum_{j=1}^N \int_{]S, S'] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_j(r, z, u) \right)^2,$$

which equals

$$2\alpha^2 \left(\int_S^{S'} X_r^N dr \right)^2 \tag{9}$$

$$+ \frac{2}{N} \sum_{j=1}^N \left(\int_{]S, S'] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_j(r, z, u) \right)^2 \tag{10}$$

$$+ \frac{4}{N} \sum_{1 \leq i < j \leq N} \left(\int_{]S, S'] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_i(r, z, u) \right) \left(\int_{]S, S'] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_j(r, z, u) \right). \tag{11}$$

In the sequel we will show that the expectation of the expressions in (9), (10) and (11) go to 0 when $\delta \rightarrow 0$, uniformly in N . We check each of these three expressions.

For (9), we have

$$2\alpha^2 \mathbb{E} \left[\left(\int_S^{S'} X_r^N dr \right)^2 \right] \leq 2\alpha^2 \mathbb{E} \left[\left(\sup_{0 \leq t \leq T} |X_t^N| (S' - S) \right)^2 \right] \leq 2\alpha^2 \delta^2 \mathbb{E} \left[\left(\sup_{0 \leq t \leq T} |X_t^N| \right)^2 \right].$$

For (10), taking expectation, we obtain

$$2\mathbb{E} \left[\left(\int_{]0,T] \times \mathbb{R}_+ \times \mathbb{R}} \mathbb{1}_{\{S < r \leq S'\}} u \mathbb{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_1(r, z, u) \right)^2 \right].$$

Using Ito's isometry and the fact that $r \mapsto \mathbb{1}_{\{S < r \leq S'\}} \mathbb{1}_{\{z \leq f(X_{r-}^N)\}}$ is predictable, we obtain

$$\begin{aligned} & \mathbb{E} \left[\left(\int_{]0,T] \times \mathbb{R}_+ \times \mathbb{R}} \mathbb{1}_{\{S < r \leq S'\}} u \mathbb{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_1(r, z, u) \right)^2 \right] \\ &= \mathbb{E} \left[\int_{]0,T] \times \mathbb{R}_+ \times \mathbb{R}} \mathbb{1}_{\{S < r \leq S'\}} u^2 \mathbb{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_1(r, z, u) \right] \\ &= \sigma^2 \mathbb{E} \left[\int_0^T \int_0^\infty \mathbb{1}_{\{S < r \leq S'\}} \mathbb{1}_{\{z \leq f(X_{r-}^N)\}} dz dr \right] = \sigma^2 \mathbb{E} \left[\int_S^{S'} f(X_{r-}^N) dr \right] \\ &\leq \sigma^2 \mathbb{E} \left[\left(A \left(\sup_{0 \leq r \leq T} |X_r^N| \right)^2 + B \right) (S' - S) \right] \\ &\leq \sigma^2 \delta \left(A \mathbb{E} \left[\left(\sup_{0 \leq r \leq T} |X_r^N| \right)^2 \right] + B \right). \end{aligned}$$

Finally, to see that the expectation of (11) is zero, it is sufficient to notice that each term of the sum is zero, using Ito's isometry and the fact that the π_j are independent Poisson random measures. \square

3.2. Convergence of the generators

Throughout this paper, we consider extended generators similar to those used in [Meyn and Tweedie \(1993\)](#) and in [Davis \(1993\)](#), because the classical notion of generator does not suit to our framework (see the beginning of [Section 6.1](#)). As this definition slightly differs from one reference to another, we define explicitly the extended generator in [Definition 6.1](#) and we prove the results on extended generators that we need in this paper. We note A^N the extended generator of X^N and \bar{A} that of \bar{X} . The goal of this section is to prove the convergence of $A^N g(x)$ to $\bar{A}g(x)$ and to establish the rate of convergence for test functions $g \in C_b^3(\mathbb{R})$. Before proving this convergence, we state a lemma which characterizes the generators for some test functions. This lemma is a straightforward consequence of Ito's formula and [Lemma 3.1](#).

Lemma 3.4. $C_b^2(\mathbb{R}) \subseteq \mathcal{D}'(\bar{A})$, and for all $g \in C_b^2(\mathbb{R})$ and $x \in \mathbb{R}$, we have

$$\bar{A}g(x) = -\alpha x g'(x) + \frac{1}{2} \sigma^2 f(x) g''(x).$$

Moreover, $C_b^1(\mathbb{R}) \subseteq \mathcal{D}'(A^N)$, and for all $g \in C_b^1(\mathbb{R})$ and $x \in \mathbb{R}$, we have

$$A^N g(x) = -\alpha x g'(x) + N f(x) \int_{\mathbb{R}} \left[g \left(x + \frac{u}{\sqrt{N}} \right) - g(x) \right] d\mu(u).$$

Now we can prove the main result of this subsection.

Proposition 3.5. *If Assumptions 1 and 2 hold, then for all $g \in C_b^3(\mathbb{R})$,*

$$|\bar{A}g(x) - A^N g(x)| \leq |f(x)| \cdot \|g'''\|_\infty \frac{1}{6\sqrt{N}} \int_{\mathbb{R}} |u|^3 d\mu(u).$$

Proof. For $g \in C_b^3(\mathbb{R})$, if we note U a random variable having distribution μ , we have

$$\begin{aligned} |A^N g(x) - \bar{A}g(x)| &\leq |f(x)| \left| N\mathbb{E} \left[g \left(x + \frac{U}{\sqrt{N}} \right) - g(x) \right] - \frac{1}{2}\sigma^2 g''(x) \right| \\ &= |f(x)| N \left| \mathbb{E} \left[g \left(x + \frac{U}{\sqrt{N}} \right) - g(x) - \frac{U}{\sqrt{N}} g'(x) - \frac{U^2}{2N} g''(x) \right] \right| \\ &\leq |f(x)| N \mathbb{E} \left[\left| g \left(x + \frac{U}{\sqrt{N}} \right) - g(x) - \frac{U}{\sqrt{N}} g'(x) - \frac{U^2}{2N} g''(x) \right| \right]. \end{aligned}$$

Using Taylor-Lagrange's inequality, we obtain the result. \square

3.3. Convergence of the semigroups

Once the convergence $A^N g(x) \rightarrow \bar{A}g(x)$ is established, together with a control of the speed of convergence, our strategy is to rely on formula (16) of Proposition 6.3, stating that

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

under suitable assumptions on X^N and \bar{X} .

Obviously, to be able to apply the above formula, we need to ensure the regularity of $x \mapsto \bar{P}_s g(x)$, together with a control of the associated norms $\|\bar{P}_s g\|_{k,\infty}$, for suitable k . This is done in the next proposition.

Proposition 3.6. *If Assumptions 1, 2 and 3 hold, then for all $t \geq 0$ and for all $g \in C_b^3(\mathbb{R})$, the function $x \mapsto \bar{P}_t g$ is C^3 . Moreover for each $1 \leq k \leq 3$, for all $T \geq 0$ there exists a constant $Q_T^{(k)}$ such that for all $g \in C_b^k(\mathbb{R})$ we have*

$$\sup_{0 \leq t \leq T} \left\| (\bar{P}_t g)^{(k)} \right\|_\infty \leq Q_T^{(k)} \|g\|_{k,\infty}.$$

Moreover, $Q_T^{(k)} = O\left(1 + e^{\beta^{(k)}T}\right)$, where $\beta^{(k)}$ depends on f, σ and α in the following way.

$$\beta^{(1)} = \frac{7}{2}\sigma^2 m_1^2 - \alpha, \beta^{(2)} = 7\sigma^2 m_1^2 + \frac{3}{2}\sigma^2 m_2^2 - \alpha, \beta^{(3)} = 12\sigma^2 m_1^2 + 3\sigma^2 m_2^2 - \alpha.$$

The proof of Proposition 3.6 requires some detailed calculus to obtain the explicit expression for $\beta^{(3)}$, so we postpone it to Appendix.

We shall also need the following bound in the sequel.

Lemma 3.7. For all $g \in C_c^2(\mathbb{R})$ such that $\text{Supp } g \subseteq [-M, M]$, we have

$$\left\| (A^N g)' \right\|_\infty \leq \Gamma \|g\|_{2,\infty} (1 + M^2),$$

for some constant $\Gamma > 0$.

Proof. We have

$$\begin{aligned} (A^N g)'(x) &= -\alpha g(x) - \alpha x g'(x) - N f'(x) g(x) - N f(x) g'(x) \\ &\quad + N f'(x) \mathbb{E} \left[g \left(x + \frac{U}{\sqrt{N}} \right) \right] + N f(x) \mathbb{E} \left[g' \left(x + \frac{U}{\sqrt{N}} \right) \right]. \end{aligned}$$

Then it is clear that for all $x \in \mathbb{R}$, we have

$$\left| (A^N g)'(x) \right| \leq \Gamma \|g\|_{1,\infty} (1 + M^2) + \left| N f'(x) \mathbb{E} \left[g \left(x + \frac{U}{\sqrt{N}} \right) \right] \right| + \left| N f(x) \mathbb{E} \left[g' \left(x + \frac{U}{\sqrt{N}} \right) \right] \right|. \quad (12)$$

We bound the jump terms using the subquadraticity of f and f' (indeed with Assumptions 1 and 3, we know that f' is sublinear, and consequently subquadratic). We can write:

$$\begin{aligned} \mathbb{E} \left[\left| g' \left(x + \frac{U}{\sqrt{N}} \right) \right| \right] &\leq \|g'\|_\infty \mathbb{E} \left[\mathbf{1}_{\{|x+U/\sqrt{N}| \leq M\}} \right] \\ &= \|g'\|_\infty \mathbb{P} \left(\left\{ x + \frac{U}{\sqrt{N}} \geq -M \right\} \cap \left\{ x + \frac{U}{\sqrt{N}} \leq M \right\} \right) \\ &\leq \|g'\|_\infty \mathbb{P} \left(\left\{ x + \frac{|U|}{\sqrt{N}} \geq -M \right\} \cap \left\{ x - \frac{|U|}{\sqrt{N}} \leq M \right\} \right) \\ &= \|g'\|_\infty \mathbb{P} \left(\left\{ |U| \geq -\sqrt{N}(M+x) \right\} \cap \left\{ |U| \geq \sqrt{N}(x-M) \right\} \right). \end{aligned}$$

Then for $x > M + 1$, using that $f(x) \leq \Gamma(1 + x^2)$, and for a constant Γ that may change from line to line,

$$\begin{aligned} \left| f(x) \mathbb{E} \left[g' \left(x + \frac{U}{\sqrt{N}} \right) \right] \right| &\leq \Gamma \|g'\|_\infty (1 + x^2) \mathbb{P} \left(|U| \geq \sqrt{N}(x - M) \right) \\ &\leq \Gamma \frac{1}{N} \mathbb{E} [U^2] \|g'\|_\infty \frac{1 + x^2}{(x - M)^2} \\ &\leq \Gamma \|g'\|_\infty (1 + M^2). \end{aligned}$$

The last inequality comes from the fact that the function $x \in [M + 1, +\infty[\mapsto \frac{1+x^2}{(x-M)^2}$ is bounded by $1 + (M + 1)^2$. With the same reasoning, we know that for all $x < -M - 1$, we have

$$\left| f(x) \mathbb{E} \left[g' \left(x + \frac{U}{\sqrt{N}} \right) \right] \right| \leq \Gamma \|g'\|_\infty (1 + M^2).$$

This concludes the proof. \square

Proof of Theorem 2.2. Step 1. The main part of the proof will be to show that Proposition 6.3 can be applied to $Y^N = X^N$ and $\bar{Y} = \bar{X}$. This will be done in Step 2 below. Indeed, once this is shown, the rest of the proof will be a straightforward consequence of Proposition 3.5, since

$$\begin{aligned}
|\bar{P}_t g(x) - P_t^N g(x)| &= \left| \int_0^t P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x) ds \right| \\
&\leq \int_0^t \mathbb{E}_x^N [|\bar{A} (\bar{P}_s g)(X_{t-s}^N) - A^N (\bar{P}_s g)(X_{t-s}^N)|] ds \\
&\leq \left(\int_{\mathbb{R}} |u|^3 d\mu(u) \right) \frac{1}{6\sqrt{N}} \left(\sup_{0 \leq s \leq t} \|(\bar{P}_s g)'''\|_{\infty} \right) \int_0^t \mathbb{E}_x^N [f(X_{t-s}^N)] ds \\
&\leq \left(\int_{\mathbb{R}} |u|^3 d\mu(u) \right) \frac{1}{6\sqrt{N}} Q_T^{(3)} \|g\|_{3,\infty} \int_0^t [A \mathbb{E}_x^N [(X_{t-s}^N)^2] + B] ds \\
&\leq \frac{1}{\sqrt{N}} K_T \|g\|_{3,\infty} (1 + x^2),
\end{aligned}$$

where we have used Propositions 3.6 and 3.1 to obtain the two last inequalities above.

Step 2. Now we show that X^N and \bar{X} satisfy the hypothesis of Proposition 6.3. To begin with we know that \bar{X} and X^N satisfy the hypothesis (i), (ii) and (iii), using Lemma 3.1. Then the hypothesis (iv) can be proved using Ito's formula for the processes X^N and \bar{X} solving the SDEs (6) and (7), and using Lemma 3.1. We know that \bar{P} satisfy hypothesis (v) thanks to Proposition 3.6.

Besides one can note that \bar{P} satisfy hypothesis (vi) using the calculations of the proof of Proposition 3.6. Then using Lemma 3.4, we see directly that \bar{A} and A^N satisfy the hypothesis (vii) and (ix). In addition (viii) is straightforward for \bar{A} , and it is a consequence of Lemma 3.7 for A^N . The only remaining hypothesis (x) is a straightforward consequence of the following Lemma 3.8. \square

Lemma 3.8. *Let $(g_k)_k$ be a sequence of $C_b^1(\mathbb{R})$ satisfying $\sup_k \|g'_k\|_{\infty} < \infty$, and for all $x \in \mathbb{R}$, $g_k(x) \rightarrow 0$ as $k \rightarrow \infty$.*

Then for all bounded sequences of real numbers $(x_k)_k$, $g_k(x_k) \rightarrow 0$ as $k \rightarrow \infty$.

Proof. Let $(x_k)_k$ be a bounded sequence. In a first time, we suppose that $(x_k)_k$ converges to some $x \in \mathbb{R}$. Then we have $|g_k(x_k)| \leq \|g'_k\|_{\infty} |x - x_k| + |g_k(x)|$ which converges to zero as k goes to infinity. In the general case, we show that for all subsequence of $(g_k(x_k))_k$, there exists a subsequence of the first one that converges to 0 (the second subsequence has to be chosen such that x_k converges). \square

3.4. Convergence in finite-dimensional distribution

Theorem 2.2 and Proposition 4.1 imply the convergence of one dimensional time marginals for functions in $C_b^3(\mathbb{R})$. Using an induction argument we can prove the convergence in finite-dimensional distribution for functions in $C_b^3(\mathbb{R})$, and then, using a classical argument of density of $C_b^3(\mathbb{R})$ in $C_b(\mathbb{R})$, we obtain the following proposition.

Proposition 3.9. *If Assumptions 1, 2, 3 and 4 hold, then for all $n \in \mathbb{N}^*$, $g_1, \dots, g_n \in C_b(\mathbb{R})$, $0 \leq t_1 \leq t_2 \leq \dots \leq t_n, x \in \mathbb{R}$, we have:*

$$\mathbb{E} [g_1(X_{t_1}^N) \dots g_n(X_{t_n}^N)] \xrightarrow{N \rightarrow \infty} \mathbb{E} [g_1(\bar{X}_{t_1}) \dots g_n(\bar{X}_{t_n})].$$

The proof of Proposition 3.9 is given in Appendix.

Now we can prove our main result, Theorem 2.1, which states the convergence in distribution of X^N to \bar{X} in $D(\mathbb{R}_+, \mathbb{R})$.

Proof of Theorem 2.1. Using the tightness of the sequence $(X^N)_N$ on $D(\mathbb{R}_+, \mathbb{R})$ (see Proposition 3.3) and the convergence of X^N to \bar{X} in finite-dimensional distribution (see Proposition 3.9), we know that X^N converges to \bar{X} in distribution in $D(\mathbb{R}_+, \mathbb{R})$ (see Theorems 13.1 and 16.7 of Billingsley (1999)). \square

4. Convergence of the transition semigroups to the invariant measure, as $N \rightarrow \infty$

In this section, we prove Theorem 2.3. In a first time, we prove a stability result for the semigroup \bar{P}_t of the limit process $(\bar{X}_t)_t$ with respect to the initial condition of the process.

Proposition 4.1. *If Assumptions 1 and 2 hold, then for all probability measures having a second moment ν_1, ν_2 ,*

$$W_2(\nu_1 \bar{P}_t, \nu_2 \bar{P}_t)^2 \leq e^{(\sigma^2 L^2 - 2\alpha)t} W_2(\nu_1, \nu_2)^2,$$

where L is a Lipschitz constant for the function \sqrt{f} .

Proof. We consider ν_1, ν_2 probability measures having a second moment. For $\varepsilon > 0$ fixed, let $X_0 \sim \nu_1$ and $Y_0 \sim \nu_2$ such that $\mathbb{E}[(X_0 - Y_0)^2] \leq W_2(\nu_1, \nu_2)^2 + \varepsilon$.

Let $(X_t)_t, (Y_t)_t$ be two solutions of the SDE (7) starting from the initial conditions X_0, Y_0 respectively, driven by the same Brownian motion B . We introduce $\zeta_t = X_t - Y_t$, so we have

$$\zeta_t = \zeta_0 - \alpha \int_0^t \zeta_s ds + \sigma \int_0^t \left(\sqrt{f(X_s)} - \sqrt{f(Y_s)} \right) dB_s.$$

Introducing $Z_t = e^{\alpha t} \zeta_t$,

$$dZ_t = \sigma e^{\alpha t} \left(\sqrt{f(X_t)} - \sqrt{f(Y_t)} \right) dB_t.$$

By Ito's formula,

$$\begin{aligned} \mathbb{E}[(Z_t)^2] &= \mathbb{E}[(Z_0)^2] + \sigma^2 \mathbb{E} \left[\int_0^t e^{2\alpha s} \left(\sqrt{f(X_s)} - \sqrt{f(Y_s)} \right)^2 ds \right] \\ &\leq \mathbb{E}[(Z_0)^2] + \sigma^2 L^2 \mathbb{E} \left[\int_0^t (Z_s)^2 ds \right]. \end{aligned}$$

By Grönwall's lemma,

$$\mathbb{E}[(Z_t)^2] \leq \mathbb{E}[(Z_0)^2] e^{\sigma^2 L^2 t},$$

which implies

$$\mathbb{E}[(X_t - Y_t)^2] \leq \mathbb{E}[(X_0 - Y_0)^2] e^{(\sigma^2 L^2 - 2\alpha)t}.$$

As a consequence,

$$W_2(\nu_1 \bar{P}_t, \nu_2 \bar{P}_t)^2 \leq W_2(\nu_1, \nu_2)^2 e^{(\sigma^2 L^2 - 2\alpha)t} + \varepsilon e^{(\sigma^2 L^2 - 2\alpha)t}.$$

Since the inequality above holds for all $\varepsilon > 0$, the proposition is proved. \square

Now using Proposition 4.1, classical arguments (see e.g. the proof of Theorem 1 of Duarte, Löcherbach and Ost (2018)) imply that \bar{X} possesses an invariant measure π which is unique. In addition, Theorem 4.2 of Meyn and Tweedie (1993) ensures that π admits a second order moment. Therefore the following result holds true.

Proposition 4.2. *If Assumptions 1 and 2 hold, and if we assume $\alpha > \sigma^2 L^2/2$, then the invariant measure π of $(\bar{P}_t)_t$ exists, is unique and admits a second order moment.*

Now we prove Theorem 2.3. We use the Kantorovich-Rubinstein duality for W_1 , that is, for all ν_1, ν_2 probability measures on \mathbb{R} having a first moment, $W_1(\nu_1, \nu_2) = \sup_{\psi} (\int \psi d\nu_1 - \int \psi d\nu_2)$, where ψ ranges over all Lipschitz continuous functions whose Lipschitz constant is smaller or equal than one (see Remark 6.5 of (Villani 2008, p. 107)).

Proof of Theorem 2.3. We fix $0 < \gamma < 1$ and $g \in C_b^3(\mathbb{R})$. Then

$$|P_t^N g(x) - \pi g| \leq |P_t^N g(x) - P_{\gamma t}^N \bar{P}_{t-\gamma t} g(x)| + |P_{\gamma t}^N \bar{P}_{t-\gamma t} g(x) - \pi g|. \quad (13)$$

Moreover,

$$\begin{aligned} |P_t^N g(x) - P_{\gamma t}^N \bar{P}_{t-\gamma t} g(x)| &= |P_{\gamma t}^N (P_{t-\gamma t}^N - \bar{P}_{t-\gamma t}) g(x)| \\ &\leq \mathbb{E}_x^N [|(P_{t-\gamma t}^N - \bar{P}_{t-\gamma t}) g(X_{\gamma t}^N)|] \\ &\leq \frac{1}{\sqrt{N}} K_t \|g\|_{3,\infty} (1 + C_t x^2 + D_t), \end{aligned} \quad (14)$$

where K_t comes from Theorem 2.2, and C_t and D_t from Lemma 3.1.

Furthermore,

$$\begin{aligned} |P_{\gamma t}^N \bar{P}_{t-\gamma t} g(x) - \pi g| &= |P_{\gamma t}^N(x, \cdot) \bar{P}_{t-\gamma t} g - \pi g| \\ &\leq \|g'\|_{\infty} W_1(P_{\gamma t}^N(x, \cdot), \bar{P}_{t-\gamma t}, \pi) \\ &= \|g'\|_{\infty} W_1(P_{\gamma t}^N(x, \cdot), \bar{P}_{t-\gamma t}, \pi \bar{P}_{t-\gamma t}) \\ &\leq \|g'\|_{\infty} W_2(P_{\gamma t}^N(x, \cdot), \bar{P}_{t-\gamma t}, \pi \bar{P}_{t-\gamma t}) \\ &\leq \|g'\|_{\infty} e^{(\frac{1}{2}\sigma^2 L^2 - \alpha)(1-\gamma)t} W_2(P_{\gamma t}^N(x, \cdot), \pi) \\ &\leq \|g'\|_{\infty} e^{(\frac{1}{2}\sigma^2 L^2 - \alpha)(1-\gamma)t} \sqrt{2 \left(C_t x^2 + D_t + \int_{\mathbb{R}} y^2 d\pi(y) \right)}. \end{aligned} \quad (15)$$

Then, replacing (14) and (15) in (13), we obtain that $|P_t^N g(x) - \pi g|$ is upper bounded by

$$\frac{1}{\sqrt{N}} K_t \|g\|_{3,\infty} (1 + C_t x^2 + D_t) + \|g'\|_{\infty} e^{(\frac{1}{2}\sigma^2 L^2 - \alpha)(1-\gamma)t} \sqrt{2 \left(C_t x^2 + D_t + \int_{\mathbb{R}} y^2 d\pi(y) \right)},$$

where C_t and D_t are defined in Lemma 3.1.

If we assume that $\alpha \geq \frac{1}{2}\sigma^2 A$, we know that $C_t = O(1)$ and $D_t = O(t)$. Moreover, if $\alpha \geq 12m_1^2\sigma^2 + 3m_2^2\sigma^2$, then $K_t = O(t^2)$ (see Proposition 3.6). Thus, if α verifies the previous inequalities, then for N_t such that

$$\lim_{t \rightarrow \infty} t^3 / \sqrt{N_t} = 0,$$

we know that $P_t^{N_t}(x, \cdot)$ converges to π for test functions in $C_b^3(\mathbb{R})$ when t goes to infinity. Using the density of $C_b^3(\mathbb{R})$ in $C_b(\mathbb{R})$, we know that this convergence holds for test functions in $C_b(\mathbb{R})$. \square

5. Convergence of $(Z_t^{N,i})_N$ in distribution in $D(\mathbb{R}_+, \mathbb{R})$

In this section we prove Theorem 2.5, that is the convergence in distribution, for each fixed k , of $(Z^{N,1}, \dots, Z^{N,k})$ to $(\bar{Z}^1, \dots, \bar{Z}^k)$ in $D(\mathbb{R}_+, \mathbb{R}^k)$. For each fixed $T > 0$, we consider the usual notion of convergence on $D([0, T], \mathbb{R})$, generalized to $D([0, T], \mathbb{R}^k)$. Namely, a sequence $(g_N)_N$ of $D([0, T], \mathbb{R}^k)$ is said to converge to g in $D([0, T], \mathbb{R}^k)$, if there exists a sequence of increasing continuous functions $\lambda_N : [0, T] \rightarrow \mathbb{R}$ such that $\lambda_N(0) = 0$, $\lambda_N(T) = T$, $\lim_{N \rightarrow \infty} \|Id - \lambda_N\|_{\infty, [0, T]} = 0$ and $\lim_{N \rightarrow \infty} \|g_N - g \circ \lambda_N\|_{\infty, [0, T]} = 0$.

Then, using Theorem 16.2 of Billingsley (1999), we know that a sequence $(g_N)_N$ converges to g in $D(\mathbb{R}_+, \mathbb{R}^k)$ if and only if $(g_N)_N$ converges to g in $D([0, T], \mathbb{R}^k)$ for all $T > 0$ that are continuity points of g . In the following, we only use this convergence criteria when it comes to convergence in $D(\mathbb{R}_+, \mathbb{R}^k)$.

To prove the convergence of $Z^{N,i}$ to \bar{Z}^i ($i \in \mathbb{N}^*$), we start by proving the convergence of their stochastic intensities. This is a straightforward consequence of Theorem 2.1 and the following lemma.

Lemma 5.1. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function. Then the function $\Psi : x \in D(\mathbb{R}_+, \mathbb{R}) \rightarrow f \circ x \in D(\mathbb{R}_+, \mathbb{R})$ is continuous.*

Proof. Let us consider a sequence $(x_n)_n$ of $D(\mathbb{R}_+, \mathbb{R})$ that converges to some x . We fix a $T > 0$ such that $(x_n)_n$ converges to x in $D([0, T], \mathbb{R})$. Then we can consider increasing functions λ_N defined on $[0, T]$ such that $\lambda_N(0) = 0$, $\lambda_N(T) = T$, $\|Id - \lambda_N\|_{\infty, [0, T]}$ vanishes and $\|x_N - x \circ \lambda_N\|_{\infty, [0, T]}$ vanishes as $N \rightarrow \infty$.

For N big enough, we know that $\|x_N\|_{\infty, [0, T]} \leq \|x\|_{\infty, [0, T]} + 1$. Introducing the modulus of continuity w of f restricted to $[0, \|x\|_{\infty, [0, T]} + 1]$, $w : [0, \|x\|_{\infty, [0, T]} + 1] \rightarrow \mathbb{R}_+$, we have

$$\|f \circ x_N - f \circ x \circ \lambda_N\|_{\infty, [0, T]} \leq w(\|x_N - x \circ \lambda_N\|_{\infty, [0, T]}) \rightarrow 0$$

as $N \rightarrow \infty$. □

To prove the convergence of $Z^{N,i}$ to \bar{Z}^i , the convergence of their respective intensities (that is, the convergence of $f(X_t^N)_{t \geq 0}$ to $f(\bar{X}_t)_{t \geq 0}$) is not sufficient, since we also manipulate the Poisson random measure π_i . So we need to prove the convergence of the pair (X^N, π_i) to (\bar{X}, π_i) in distribution. According to our definition of \bar{Z}^i in (8), it is obvious that \bar{X} is independent of any finite subset of $(\pi_j)_{j \in \mathbb{N}^*}$. The goal of Proposition 5.2 is to justify the way we introduced \bar{Z}^i .

Proposition 5.2. *Under Assumptions 1, 2, 3 and 4, for each $k \geq 1$, the sequence $\mathcal{D}(X^N, \pi_1, \dots, \pi_k)$ converges weakly to $\mathcal{D}(\bar{X}) \otimes \mathcal{D}(\pi_1) \otimes \dots \otimes \mathcal{D}(\pi_k)$.*

The proof of the previous proposition consists in applying Theorem II.6.3 of Ikeda and Watanabe (1989), which states that Brownian motion and Poisson random measures defined with respect to the same filtration are necessarily independent. As the proof is technically involved, we postpone it to Appendix.

We now turn to the proof of the convergence of $Z^{N,i}$ to \bar{Z}^i . A first attempt in this direction could be to write

$$Z^{N,i} = \Phi(X^N, \pi_i) \text{ and } \bar{Z}^i = \Phi(f(\bar{X}), \pi)$$

with

$$\Phi : (x, \pi) \in D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M}^\# \mapsto \int_{]0, \cdot] \times \mathbb{R}_+ \times \mathbb{R}} \mathbb{1}_{\{z \leq x(s-)\}} d\pi(s, z, u),$$

and to use the weak convergence of $f(X_t^N)_{t \geq 0}$ to $f(\bar{X}_t)_{t \geq 0}$. The problem is that Φ is not continuous for interesting topologies (see Example 5.3).

Example 5.3. Let us consider the point measure $\pi = \delta_{(1,1)}$ on $\mathbb{R}_+ \times \mathbb{R}_+$ (we omit the third parameter u of the point measure since it is not used here), and the constant function $x : t \in \mathbb{R}_+ \mapsto 1$. In addition, we consider the functions x_N defined as in Figure 2 below. Obviously, $\|x - x_N\|_\infty = 1/N$, but $\Phi(x, \pi)(t) = \mathbf{1}_{\{t \geq 1\}}$ and $\Phi(x_N, \pi) = 0$. In other words, x_N converges strongly to x , but $\Phi(x_N, \pi)$ does not converge to $\Phi(x, \pi)$ for non-trivial topologies.

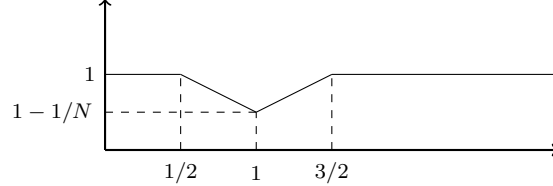


FIGURE 2. Graph of x_N

The reason why the convergence of $Z^{N,i}$ to \bar{Z}^i still holds is the independence between \bar{X} and π_i . This independence entails that the point measure π_i does not charge any point on $\{(t, f(\bar{X}_{t-})) : t \geq 0\}$ (almost surely). To use this property, we use Skorohod's embedding theorem to have an almost sure convergence of a copy of (X^N, π_i) to a copy of (\bar{X}, π_i) .

Proof of Theorem 2.5. In this proof, we note π'_i the Poisson random measure on \mathbb{R}_+^2 defined as $\pi'_i(A \times B) = \pi_i(A \times B \times \mathbb{R})$. As X^N converges in distribution to \bar{X} in Skorohod topology, we know that $f(X_t^N)_t$ converges also to $f(\bar{X}_t)_t$ in distribution (see Lemma 5.1). In particular, the sequence $(f(X_t^N)_t)_N$ is tight on $D(\mathbb{R}_+, \mathbb{R})$. Furthermore, as $\mathcal{M}^\#$ is a Polish space (see Theorem A2.6.III.(i) of Daley and Vere-Jones (2003)), we know that each π'_i is tight on this space. Therefore, the sequence $(f(X_t^N)_t, \pi'_1, \dots, \pi'_k)_N$ is tight on $D(\mathbb{R}_+, \mathbb{R}) \times (\mathcal{M}^\#)^k$.

Let us consider a limit distribution P for the sequence of tuples $(f(X^N), \pi'_1, \dots, \pi'_k)_N$. The marginals of P are respectively the distribution of $f(\bar{X})$ and those of π'_i ($1 \leq i \leq k$). Since the random variables $\bar{X}, \pi'_1, \dots, \pi'_k$ are independent (see Proposition 5.2), the limit distribution is uniquely determined. As a consequence, $(f(X^N), \pi'_1, \dots, \pi'_k)$ converges in distribution to $(f(\bar{X}), \pi'_1, \dots, \pi'_k)$.

Now Skorohod's embedding theorem (see Theorem 6.7 of Billingsley (1999)) implies that there exist random variables \tilde{Y}, \tilde{Y}^N ($N \in \mathbb{N}^*$), $\tilde{\pi}_i$ ($1 \leq i \leq k$), $\tilde{\pi}_i^N$ ($1 \leq i \leq k, N \in \mathbb{N}^*$) defined on some probability space Ω' such that:

- $(\tilde{Y}, \tilde{\pi}_1, \dots, \tilde{\pi}_k)$ has the same distribution as $(f(\bar{X}), \pi'_1, \dots, \pi'_k)$.
- $(\tilde{Y}^N, \tilde{\pi}_1^N, \dots, \tilde{\pi}_k^N)$ has the same distribution as $(f(X^N), \pi'_1, \dots, \pi'_k)$.
- \tilde{Y}^N converges almost surely to \tilde{Y} in Skorohod topology.
- $\tilde{\pi}_i^N$ converges almost surely to $\tilde{\pi}_i$ in $\mathcal{M}^\#$.

Let

$$\tilde{Z}_t^{N,i} = \int_{]0,t] \times \mathbb{R}_+} \mathbf{1}_{\{z \leq Y_{s-}^N\}} d\tilde{\pi}_i^N(s, z).$$

Theorem A2.6.III.(ii) of Daley and Vere-Jones (2003) implies that $\tilde{Z}^N = \left(\tilde{Z}^{N,i}\right)_{1 \leq i \leq k}$ has the same distribution as $Z^N = (Z^{N,i})_{1 \leq i \leq k}$, and \tilde{Z} has the same distribution as \bar{Z} . So to prove the convergence in distribution of Z^N to \bar{Z} , it is sufficient to prove the almost sure convergence of \tilde{Z}^N to \tilde{Z} .

From now on, we fix a ω in Ω' (i.e. the probability space given by Skorohod's embedding theorem) satisfying the following conditions:

- \tilde{Y}^ω is continuous.
- for each $1 \leq i \leq k$, $\tilde{\pi}_i^\omega \left(\left\{ (t, \tilde{Y}_t^\omega) : t \in \mathbb{R}_+ \right\} \right) = 0$.
- for each $1 \leq i \leq k$, $\forall t \geq 0$, $\tilde{\pi}_i^\omega(\{t\} \times \mathbb{R}_+) \leq 1$.
- $\tilde{Y}^{N,\omega}$ converges to \tilde{Y}^ω in Skorohod topology.
- $\tilde{\pi}_i^{N,\omega}$ converges to $\tilde{\pi}_i^\omega$ in $\mathcal{M}^\#$.
- for all $1 \leq i, j \leq k$, if $i \neq j$ then for all $t \in \mathbb{R}_+$, $\tilde{\pi}_i^\omega(\{t\} \times \mathbb{R}_+) \neq 0 \Rightarrow \tilde{\pi}_j^\omega(\{t\} \times \mathbb{R}_+) = 0$.

We emphasize the fact that these properties are satisfied for almost all $\omega \in \Omega'$. Until the last paragraph of the proof, this $\omega \in \Omega'$ is fixed. To lighten the notations, we omit this ω , and we just write \tilde{Y} or $\tilde{\pi}_i$, instead of \tilde{Y}^ω or $\tilde{\pi}_i^\omega$.

We fix $t \geq 0$ such that for each $1 \leq i \leq k$, $\tilde{\pi}_i(\{t\} \times \mathbb{R}_+) = 0$ and for all $N \in \mathbb{N}^*$, $\tilde{\pi}_i^N(\{t\} \times \mathbb{R}_+) = 0$. In particular t is a point of continuity of \tilde{Z} and of each \tilde{Z}^N . We consider $T \in \mathbb{N}^*$ such that $T > \max \left(t, \left\| \tilde{Y} \right\|_{\infty, [0, t]}, \sup_N \left\| \tilde{Y}^N \right\|_{\infty, [0, t]} \right)$ and such that $\tilde{\pi}_i(\{T\} \times [0, T] \cup [0, T] \times \{T\}) = 0$ for each $1 \leq i \leq k$. Let us consider $n_i = \tilde{\pi}_i([0, T]^2)$ and $n_i^N = \tilde{\pi}_i^N([0, T]^2)$. In the rest of the proof, we identify the point measure $\tilde{\pi}_i$ with the related set of points. We write $\tilde{\pi}_i \cap [0, T]^2 = \{(\tau_{i,j}, \zeta_{i,j}) : 1 \leq j \leq n_i\}$ and $\tilde{\pi}_i^N \cap [0, T]^2 = \{(\tau_{i,j}^N, \zeta_{i,j}^N) : 1 \leq j \leq n_i^N\}$, where the pairs are lexicographically ordered.

Firstly as $\tilde{\pi}_i^N$ converges to $\tilde{\pi}_i$ in $\mathcal{M}^\#$ and $\tilde{\pi}_i(\{T\} \times [0, T] \cup [0, T] \times \{T\}) = 0$, we can apply Proposition A2.6.II.(iv) of Daley and Vere-Jones (2003) to show that n_i^N converges to n_i , so we know that $n_i^N = n_i$ for N big enough.

Now we show that for all $1 \leq i \leq k$, for all $1 \leq j \leq n_i$, $\tau_{i,j}^N$ and $\zeta_{i,j}^N$ converge respectively to $\tau_{i,j}$ and $\zeta_{i,j}$. The idea of the proof consists in defining disjoint sets $U_{i,j}^\varepsilon$ of radius ε that contain each $(\tau_{i,j}, \zeta_{i,j})$ and to use Proposition A2.6.II.(iv) of Daley and Vere-Jones (2003) again to show that $(\tau_{i,j}^N, \zeta_{i,j}^N)$ is necessarily in $U_{i,j}^\varepsilon$ for all ε . We fix $1 \leq i \leq k$, some $\varepsilon > 0$ and we consider $\gamma_i = \min_{1 \leq j \leq n_i} \tau_{i,j+1} - \tau_{i,j} > 0$. We can choose $0 < \eta < \varepsilon \wedge \frac{\gamma_i}{3}$ such that for all $1 \leq j_1, j_2 \leq n_i$, if $j_1 \neq j_2$ then $B((\tau_{i,j_1}, \zeta_{i,j_1}), \eta) \cap B((\tau_{i,j_2}, \zeta_{i,j_2}), \eta) = \emptyset$ (where we endow \mathbb{R}_+^2 with $\|\bullet\|_\infty$). Then we know that for all $1 \leq j \leq n_i$, $|\tilde{\pi}_i^N \cap B((\tau_{i,j}, \zeta_{i,j}), \eta)|$ converges to $|\tilde{\pi}_i \cap B((\tau_{i,j}, \zeta_{i,j}), \eta)| = 1$. This means that for all $1 \leq j \leq n_i$, there exists a unique $l_j^N \in \llbracket 1, n_i \rrbracket$ such that $(\tau_{i,l_j^N}^N, \zeta_{i,l_j^N}^N) \in B((\tau_{i,j}, \zeta_{i,j}), \eta)$. We note that for all $1 \leq j \leq n_i - 1$, $\tau_{i,l_j^N}^N < \tau_{i,j} + \frac{\gamma_i}{3} < \tau_{i,j+1} - \frac{\gamma_i}{3} < \tau_{i,l_{j+1}^N}^N$, so this implies that $\tau_{i,l_1^N}^N < \tau_{i,l_2^N}^N < \dots < \tau_{i,l_{n_i}^N}^N$, since we have ordered the pairs lexicographically, this implies $l_j^N = j$. So we just proved that for all j , for all N (big enough), $(\tau_{i,j}^N, \zeta_{i,j}^N) \in B((\tau_{i,j}, \zeta_{i,j}), \eta)$,

$$i.e. \quad |\tau_{i,j} - \tau_{i,j}^N| \vee |\zeta_{i,j} - \zeta_{i,j}^N| < \eta < \varepsilon.$$

Thus, $\tau_{i,j}^N$ and $\zeta_{i,j}^N$ converge respectively to $\tau_{i,j}$ and $\zeta_{i,j}$.

Notice that

$$\tilde{Z}_t^{N,i} = \sum_{j=1}^{n_i} \mathbb{1}_{\left\{ \zeta_{i,j}^N \leq \tilde{Y}_{\tau_{i,j}^N}^N \right\}} \mathbb{1}_{\{\tau_{i,j}^N \leq t\}}.$$

Now we argue that $\mathbb{1}_{\left\{ \zeta_{i,j}^N \leq \tilde{Y}_{\tau_{i,j}^N}^N \right\}}$ converges to $\mathbb{1}_{\{\zeta_{i,j} \leq \tilde{Y}_{\tau_{i,j}}\}}$. Indeed: there are two cases, either $\zeta_{i,j} < \tilde{Y}_{\tau_{i,j}}$, or $\zeta_{i,j} > \tilde{Y}_{\tau_{i,j}}$, in the first case we consider $\varepsilon > 0$ such that $\zeta_{i,j} + \varepsilon < \tilde{Y}_{\tau_{i,j}}$. Then using Lemma 6.7, for N big enough, we have $\zeta_{i,j}^N < \zeta_{i,j} + \frac{\varepsilon}{3} < \tilde{Y}_{\tau_{i,j}} - \frac{\varepsilon}{3} < \tilde{Y}_{\tau_{i,j}^N}^N$, implying the convergence of $\mathbb{1}_{\left\{ \zeta_{i,j}^N \leq \tilde{Y}_{\tau_{i,j}^N}^N \right\}}$. The second case is handled in the same way. For the same reason, $\mathbb{1}_{\{\tau_{i,j}^N \leq t\}}$ converges to $\mathbb{1}_{\{\tau_{i,j} \leq t\}}$ (since we chose t such that $\tilde{\pi}_i(\{t\} \times \mathbb{R}_+) = 0$).

To resume, we have just showed that for all $t \geq 0$ satisfying that for each $1 \leq i \leq k$, $\tilde{\pi}_i(\{t\} \times \mathbb{R}_+) = 0$ and for all $N \in \mathbb{N}^*$, $\tilde{\pi}_i^N(\{t\} \times \mathbb{R}_+) = 0$, \tilde{Z}_t^N converges to \tilde{Z}_t in \mathbb{R}^k . Observing that these points are dense in \mathbb{R}_+ , we can apply Lemma 6.8 to obtain that \tilde{Z}^N converges to \tilde{Z} in $D([0, t], \mathbb{R}^k)$ for all t with the above properties. We observe that such t are points of continuity of \tilde{Z} , and that we can choose an increasing sequence $(t_n)_n$ of such points that tends to infinity. As a consequence, Proposition 16.2 of Billingsley (1999) can still be used to show that \tilde{Z}^N converges to \tilde{Z} in $D(\mathbb{R}_+, \mathbb{R}^k)$.

In the previous paragraph, we have worked with a fixed $\omega \in \Omega'$ satisfying a finite number of almost sure properties. So we just showed the almost sure convergence of \tilde{Z}^N to \tilde{Z} in $D(\mathbb{R}_+, \mathbb{R}^k)$ which implies that $Z^N = (Z^{N,1}, \dots, Z^{N,k})$ converges in distribution to $\bar{Z} = (\bar{Z}^1, \dots, \bar{Z}^k)$ in $D(\mathbb{R}_+, \mathbb{R}^k)$. \square

Corollary 5.4. *If Assumptions 1, 2, 3 and 4 hold, then the sequence $(Z^{N,1}, Z^{N,2}, \dots, Z^{N,k})_N$ converges to $(\bar{Z}^1, \bar{Z}^2, \dots, \bar{Z}^k)$ in distribution in $D(\mathbb{R}_+, \mathbb{R}^k)$, for all fixed $k \geq 1$.*

Proof. We consider a function g that is bounded and continuous in $D(\mathbb{R}_+, \mathbb{R}^k)$. It is sufficient to show that g is continuous in $D(\mathbb{R}_+, \mathbb{R}^k)$. We consider a sequence $(x^{N,1}, \dots, x^{N,k})_N$ that converges to some limit $(\bar{x}^1, \dots, \bar{x}^k)$ in $D(\mathbb{R}_+, \mathbb{R}^k)$. Then using Theorem 16.2 of Billingsley (1999), we know that $(x^{N,1}, \dots, x^{N,k})_N$ converges to $(\bar{x}^1, \dots, \bar{x}^k)$ in $D(\mathbb{R}_+, \mathbb{R}^k)$. So we know that $g(x^{N,1}, \dots, x^{N,k})$ converges to $g(\bar{x}^1, \dots, \bar{x}^k)$. \square

6. Appendix

6.1. Extended generators

In this subsection, we define clearly the notion of generators we use and we prove the results that we use to prove formula (3). In the general theory of semigroups, one defines the generators on some Banach space. In the frame of semigroups related to Markov processes, one generally considers $(C_b(\mathbb{R}), \|\bullet\|_\infty)$. In this context, the generator A of a semigroup $(P_t)_t$ is defined on the set of functions $\mathcal{D}(A) = \left\{ g \in C_b(\mathbb{R}) : \exists h \in C_b(\mathbb{R}), \left\| \frac{1}{t}(P_t g - g) - h \right\|_\infty \xrightarrow{t \rightarrow 0} 0 \right\}$. Then one denotes the previous function h as Ag . If A is the generator of a diffusion, we can only guarantee that $\mathcal{D}(A)$ contains the functions that have a compact support, but to prove Proposition 6.3, we need to apply the generators of the processes $(X_t^N)_t$ and $(\bar{X}_t)_t$ to functions of the type $\bar{P}_s g$, and we cannot guarantee that $\bar{P}_s g$ has compact support even if we assume g to be in $C_c^\infty(\mathbb{R})$.

That is why we consider extended generators (see for instance [Meyn and Tweedie \(1993\)](#) or [Davis \(1993\)](#)) defined by the point-wise convergence on \mathbb{R} instead of the uniform convergence that allows us to define the generator on $C_b^n(\mathbb{R})$ for suitable $n \in \mathbb{N}^*$ and to prove that some properties of the classical theory of semigroups still hold for this larger class of functions.

Definition 6.1. Let $(X_t)_t$ be a Markov process on \mathbb{R} . We define $P_t g(x) = \mathbb{E}_x[g(X_t)]$ for all functions g such that the previous expression is well-defined and finite for $x \in \mathbb{R}$. Then we define $\mathcal{D}'(A)$ to be the set of functions $g \in C_b(\mathbb{R})$ such that for each $x \in \mathbb{R}$, $\frac{1}{t}(P_t g(x) - g(x))$ converge to some limit that we note $Ag(x)$ and such that:

- for all $t \geq 0$, $\int_0^t |Ag(X_s)| ds$ is almost surely finite,
- $g(X_t) - g(X_0) - \int_0^t Ag(X_s) ds$ is a \mathbb{P}_x -martingale for all x .

We note $\mathcal{D}'(A)$ the domain of the extended generator to avoid confusions with $\mathcal{D}(A)$ which is reserved for the domain of A for the uniform convergence.

Now we generalize a classical result for generators defined with respect to the uniform convergence to extended generators.

Lemma 6.2. Let $(X_t)_t$ be a Markov process with semigroup $(P_t)_t$ and extended generator A .

- (1) Let $g \in \mathcal{D}'(A)$ and $x \in \mathbb{R}$ such that for all $t \geq 0$, $\mathbb{E}_x \left[\sup_{0 \leq s \leq t} |P_s Ag(X_t)| \right]$ is finite. Then the function $t \mapsto P_t g(x)$ is right differentiable at every $t \geq 0$, and we have

$$\frac{d^+}{dt} (P_t g(x)) = P_t Ag(x).$$

In addition, if $P_t g \in \mathcal{D}'(A)$, then $AP_t g(x) = P_t Ag(x)$.

- (2) Let $g \in \mathcal{D}'(A)$ and $x \in \mathbb{R}$ such that there exists some non-negative function $M : \mathbb{R} \rightarrow \mathbb{R}_+$ such that for all $t \geq 0$, $\sup_{0 \leq s \leq t} \mathbb{E}_x[M(X_s)]$ is finite and such that for all $0 \leq t \leq 1$ and $y \in \mathbb{R}$, we have $|P_t Ag(y) - Ag(y)| \leq \Gamma M(y) \varepsilon(t)$ for some constant Γ that is allowed t depend on g , where $\varepsilon(t)$ vanishes when t goes to 0. Then the function $t \mapsto P_t g(x)$ is left differentiable at every $t > 0$, and we have

$$\frac{d^-}{dt} (P_t g(x)) = P_t Ag(x).$$

Proof. For the point (1), we know that for all $h > 0$, we have:

$$\left| \frac{1}{h} (P_{t+h} g(x) - P_t g(x)) - P_t Ag(x) \right| \leq \mathbb{E}_x \left[\left| \frac{1}{h} (P_h g(X_t) - g(X_t)) - Ag(X_t) \right| \right].$$

As the expression appearing within the expectation above vanishes almost surely when h goes to 0 (since $g \in \mathcal{D}'(A)$), and as we can bound it by $\sup_{0 \leq s \leq t} |P_s Ag(X_t)| + |Ag(X_t)|$ (using the fact that

$P_h g(y) - g(y) = \int_0^h P_s Ag(y) ds$ since we take $g \in \mathcal{D}'(A)$), we know that this expectation vanishes as h goes to 0 by dominated convergence. This means exactly that $\frac{d^+}{dt} (P_t g(x))$ exists and is $P_t Ag(x)$.

If we suppose in addition that $P_t g \in \mathcal{D}'(A)$, then $AP_t g(x)$ is the limit of $h^{-1} (P_{t+h} g(x) - P_t g(x))$, which is $\frac{d^+}{dt} P_t g(x) = P_t Ag(x)$.

Now we prove the point (2) of the lemma. Let h be some positive number. We know that

$$\left| \frac{1}{-h} (P_{t-h}g(x) - P_tg(x)) - P_tAg(x) \right|$$

is upper bounded by

$$\begin{aligned} & \mathbb{E}_x \left[\left| \frac{1}{h} (P_hg(X_{t-h}) - g(X_{t-h})) - Ag(X_{t-h}) \right| \right] + \mathbb{E}_x [|Ag(X_{t-h}) - P_hAg(X_{t-h})|] \\ & \leq \mathbb{E}_x \left[\sup_{0 \leq s \leq h} |Ag(X_{t-h}) - P_sAg(X_{t-h})| \right] + \mathbb{E}_x [|Ag(X_{t-h}) - P_hAg(X_{t-h})|]. \end{aligned}$$

Then we just have to show that $\mathbb{E}_x \left[\sup_{0 \leq s \leq h} |Ag(X_{t-h}) - P_sAg(X_{t-h})| \right]$ vanishes when h goes to 0.

But this follows from the fact that it is upper bounded by $\Gamma \left(\sup_{0 \leq s \leq h} \varepsilon(s) \right) \left(\sup_{0 \leq r \leq t} \mathbb{E}_x [M(X_r)] \right)$. \square

The goal of the next proposition is to obtain a control of the difference between the semigroups of two Markov processes, provided we dispose already of a control of the distance between the two generators. This proposition is an adaptation of Lemma 1.6.2 from [Ethier and Kurtz \(2005\)](#) to the notion of extended generators defined by the point-wise convergence.

Proposition 6.3. *Let $(Y_t^N)_{t \in \mathbb{R}_+}$ and $(\bar{Y}_t)_{t \in \mathbb{R}_+}$ be Markov processes whose semigroups and (extended) generators are respectively P^N, A^N and \bar{P}, \bar{A} . We suppose that:*

- (i) for all $x \in \mathbb{R}$ and $T > 0$, $\sup_{0 \leq t \leq T} \mathbb{E}_x \left[(\bar{Y}_t)^2 \right] \leq \Gamma_T (1 + x^2)$ and $\sup_{0 \leq t \leq T} \mathbb{E}_x^N \left[(Y_t^N)^2 \right] \leq \Gamma_T (1 + x^2)$ for some $\Gamma_T > 0$ independent of x .
- (ii) for all $x \in \mathbb{R}$ and $T > 0$, $\sup_{0 \leq t \leq T} \mathbb{E}_x \left[(\bar{Y}_t)^4 \right] \leq \Gamma_T (1 + x^4)$ and $\sup_{0 \leq t \leq T} \mathbb{E}_x^N \left[(Y_t^N)^4 \right] \leq \Gamma_T (1 + x^4)$.
- (iii) for all $T > 0$, $\mathbb{E} \left[\left(\sup_{0 \leq t \leq T} |Y_t^N| \right)^2 \right] < +\infty$.
- (iv) for all $0 \leq s, t \leq T$ and $x \in \mathbb{R}$,

$$\mathbb{E}_x \left[(\bar{Y}_t - \bar{Y}_s)^2 \right] \leq \Gamma_{T,x} \varepsilon(|t - s|) \text{ and } \mathbb{E}_x^N \left[(Y_t^N - Y_s^N)^2 \right] \leq \Gamma_{T,x} \varepsilon(|t - s|),$$

where $\varepsilon(h)$ vanishes when h goes to 0, and where $\Gamma_{T,x}$ is some constant that depends only on T and x .

- (v) for all $g \in C_b^3(\mathbb{R})$, $\bar{P}_t g \in C_b^3(\mathbb{R})$, and for all $T > 0$, $\sup_{0 \leq t \leq T} \|\bar{P}_t g\|_{3,\infty} \leq Q_T \|g\|_{3,\infty}$ for some $Q_T > 0$.

- (vi) for all $g \in C_b^3(\mathbb{R})$, $i \in \{0, 1, 2\}$ and $x \in \mathbb{R}$, $s \mapsto (\bar{P}_s g)^{(i)}(x) = \frac{\partial^i}{\partial x^i} (\bar{P}_s g(x))$ is continuous.
- (vii) $C_b^3(\mathbb{R}) \subseteq \mathcal{D}'(A^N) \cap \mathcal{D}'(\bar{A})$. For all $g \in C_b^3(\mathbb{R})$ and $x \in \mathbb{R}$, $|Ag(x)| \leq \Gamma \|g\|_{2,\infty} (1 + x^2)$ and $|A^N g(x)| \leq \Gamma \|g\|_{2,\infty} (1 + x^2)$.
- (viii) for all $g \in C_c^3(\mathbb{R})$ such that $\text{Supp } g \subseteq [-M, M]$, $\|(\bar{A}g)'\|_\infty \leq \Gamma \|g\|_{3,\infty} (1 + M^2)$ and $\|(A^N g)'\|_\infty \leq \Gamma \|g\|_{3,\infty} (1 + M^2)$.

- (ix) there exists some $\Gamma > 0$ such that for all $x, y \in \mathbb{R}$, for all $g \in C_b^3(\mathbb{R})$, $|\bar{A}g(x) - \bar{A}g(y)| \leq \Gamma(1+x^2+y^2)|x-y|$ and $|A^N g(x) - A^N g(y)| \leq \Gamma(1+x^2+y^2)|x-y|$.
- (x) we assume that $\lim_{k \rightarrow \infty} \bar{A}g_k(x_k) = \lim_{k \rightarrow \infty} A^N g_k(x_k) = 0$, for any bounded sequence of real numbers $(x_k)_k$, and for any sequence $(g_k)_k$ of $C_b^3(\mathbb{R})$ satisfying
- (1) $\forall i \in \{0, 1, 2\}, \forall x \in \mathbb{R}, g_k^{(i)}(x) \xrightarrow[k \rightarrow \infty]{} 0$,
 - (2) $\forall i \in \{0, 1, 2, 3\}, \sup_k \|g_k^{(i)}\|_\infty < \infty$.

Then we have for each $g \in C_b^3(\mathbb{R}), x \in \mathbb{R}$ and $t \in \mathbb{R}_+$:

$$(\bar{P}_t - P_t^N)g(x) = \int_0^t P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x) ds. \quad (16)$$

Remark 6.4. Notice that the conditions of Proposition 6.3 are not all symmetric with respect to the processes \bar{Y} and Y^N . Indeed, the regularity hypothesis of the semigroup with respect to the initial condition only concerns \bar{P} (see hypothesis (v) and (vi)). Moreover, hypothesis (iii) provides a stronger control on Y^N than what is needed for \bar{Y} .

Proof. We fix $t \geq 0, N \in \mathbb{N}^*, g \in C_b^3(\mathbb{R}), x \in \mathbb{R}$ in the proof. We note $u(s) = P_{t-s}^N \bar{P}_s g(x)$. Firstly we show that $s \mapsto \bar{P}_s g(x)$ and $s \mapsto P_s^N h(x)$ are differentiable for all $h \in C_b^3(\mathbb{R})$, by showing that \bar{P} and P^N satisfy the hypothesis of Lemma 6.2. The condition of the point (1) of the lemma is a straightforward consequence of hypothesis (i) and (vii), and the conditions of the point (2) are satisfied for $M(x) = \sqrt{1+x^4}$ using hypothesis (i), (ii), (iv) and (ix). As a consequence, and thanks to hypothesis (v), u is differentiable and

$$\begin{aligned} u'(s) &= -\frac{d}{du} (P_u^N \bar{P}_s g(x)) \Big|_{u=t-s} + \frac{d}{du} (P_{t-s}^N \bar{P}_u g(x)) \Big|_{u=s} \\ &= -P_{t-s}^N A^N \bar{P}_s g(x) + P_{t-s}^N \bar{P}_s \bar{A}g(x) \\ &= P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x). \end{aligned}$$

The second equality comes from the fact that \bar{P} satisfy the additional assumption of the point (1) of Lemma 6.2 (see hypothesis (v) and (vii)).

Now we show that u' is continuous. Indeed if it is the case, then we will have

$$u(t) - u(0) = \int_0^t u'(s) ds,$$

which is exactly the assertion. In order to prove the continuity of u' , we consider a sequence $(s_k)_k$ that converges to some $s \in [0, t]$, and we write

$$\begin{aligned} |P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x) - P_{t-s_k}^N (\bar{A} - A^N) \bar{P}_{s_k} g(x)| &\leq |(P_{t-s}^N - P_{t-s_k}^N) (\bar{A} - A^N) g_s(x)| \quad (17) \\ &\quad + |P_{t-s_k}^N (\bar{A} - A^N) (\bar{P}_s - \bar{P}_{s_k}) g(x)|, \quad (18) \end{aligned}$$

where $g_s = \bar{P}_s g \in C_b^3(\mathbb{R})$.

To show that the term (17) vanishes when k goes to infinity, we introduce, for all $M > 0$ the function $\varphi_M(g_s)(y) = g_s(y) \cdot \xi_M(y)$ where $\xi_M : \mathbb{R} \rightarrow [0, 1]$ is C^∞ , and $\forall |y| \leq M, \xi_M(y) = 1$ and $\forall |y| \geq M+1, \xi_M(y) = 0$. We note that the term (17) is bounded by

$$|(P_{t-s}^N - P_{t-s_k}^N) (\bar{A} - A^N) \varphi_M(g_s)(x)| + |(P_{t-s}^N - P_{t-s_k}^N) (\bar{A} - A^N) (g_s - \varphi_M(g_s))(x)| =: A_1 + A_2.$$

If we consider the function $h_{M,s} = (\bar{A} - A^N) \varphi_M(g_s)$, using hypothesis (iv), (v) and (viii), we have

$$\begin{aligned} A_1 &\leq \mathbb{E}_x^N [|h_{M,s}(Y_{t-s}^N) - h_{M,s}(Y_{t-s_k}^N)|] \\ &\leq \|h'_{M,s}\|_\infty \mathbb{E}_x^N [|Y_{t-s}^N - Y_{t-s_k}^N|] \\ &\leq \Gamma (1 + M^2) \|g\|_{3,\infty} \varepsilon (|s - s_k|)^{1/2}. \end{aligned}$$

Choosing $M = M_k = \varepsilon(|s - s_k|)^{-1/5}$, it follows that $\lim_{k \rightarrow \infty} A_1 = 0$. To see that the term A_2 vanishes, it is sufficient to notice that A_2 is bounded by

$$\mathbb{E}_x^N [|(\bar{A} - A^N)(g_s - \varphi_{M_k}(g_s))(Y_{t-s}^N)|] + \mathbb{E}_x^N [|(\bar{A} - A^N)(g_s - \varphi_{M_k}(g_s))(Y_{t-s_k}^N)|].$$

We know that the expressions in the expectations vanish almost surely (using hypothesis (x)), and then we can apply dominated convergence (using hypothesis (iii) and (vii)).

We just proved that the term (17) vanishes. To finish the proof, we need to show that the term (18) vanishes. We note that the term (18) is bounded by:

$$\mathbb{E}_x^N [|\bar{A}g_k(Y_{t-s_k}^N)|] + \mathbb{E}_x^N [|A^N g_k(Y_{t-s_k}^N)|],$$

where $g_k = (\bar{P}_s - \bar{P}_{s_k})g \in C_b^3(\mathbb{R})$.

We have to show that the terms in the sum above vanish as k goes to infinity. Firstly we know that $\bar{A}g_k(Y_{t-s_k}^N)$ and $A^N g_k(Y_{t-s_k}^N)$ vanish almost surely when k goes to infinity (see hypothesis (iii), (v), (vi) and (x)). Dominated convergence, using the hypothesis (i), (iii), (v) and (vii), then implies the result. \square

6.2. Grönwall's lemma

The version of Grönwall's lemma we use in the paper is a particular case of [Grönwall's inequality \(2019\)](#). We state it below.

Lemma 6.5. *Let γ and u be non-negative measurable functions defined on \mathbb{R}_+ , and let α be a non-negative constant. Assume that $u \in L_{loc}^1(dt)$, and that for all $t \geq 0$,*

$$u(t) \leq \gamma(t) + \alpha \int_0^t u(s) ds, \tag{19}$$

then for all $t \geq 0$, we have

$$u(t) \leq \gamma(t) + \alpha \int_0^t \gamma(s) e^{\alpha(t-s)} ds.$$

Moreover, if γ is nondecreasing then, for all $t \geq 0$, we have:

$$u(t) \leq \gamma(t) e^{\alpha t}.$$

An interesting point of Lemma 6.5 is that it does not require any continuity hypothesis on u , contrarily to more common versions of Grönwall's lemma. We reproduce the proof of [Grönwall's inequality \(2019\)](#) for self-containedness.

Proof. We note μ the measure $\mu(dt) = \alpha dt$. Firstly we prove by induction on n that for all $n \in \mathbb{N}$

$$u(t) \leq \gamma(t) + \int_0^t \gamma(s) \sum_{k=0}^{n-1} \mu^{\otimes k}(A_k(s, t)) \mu(ds) + R_n(t), \quad (20)$$

where $R_n(t) = \int_0^t u(s) \mu^{\otimes n}(A_n(s, t)) \mu(ds)$ and $A_n(s, t) = \{(s_1, \dots, s_n) \in]s, t[^n : s < s_1 < \dots < s_n < t\}$.

The case $n = 0$ is inequality (19). To show the induction step, we replace the assumed inequality in the expression of $R_n(t)$ and obtain

$$R_n(t) \leq \int_0^t \gamma(s) \mu^{\otimes n}(A_n(s, t)) \mu(ds) + \tilde{R}_n(t),$$

with $\tilde{R}_n(t) = \int_0^t (\int_0^r u(s) \mu(ds)) \mu^{\otimes n}(A_n(r, t)) \mu(dr)$.

Using Fubini-Tonelli's theorem, we have $\tilde{R}_N(t) = R_{n+1}(t)$. As a consequence, equality (20) is proved for all $n \in \mathbb{N}$.

A straightforward induction gives

$$\mu^{\otimes n}(A_n(s, t)) = \frac{\alpha^n}{n!} (t - s)^n,$$

implying that, for all $n \in \mathbb{N}$,

$$u(t) \leq \gamma(t) + \int_0^t \gamma(s) \sum_{k=0}^{n-1} \frac{\alpha^k}{k!} (t - s)^k \mu(ds) + R_n(t). \quad (21)$$

As $R_n(t) = \frac{\alpha^n}{n!} \int_0^t u(s) (t - s)^n ds \leq \frac{\alpha^n}{n!} t^n \int_0^t u(s) ds$, we know that $R_n(t)$ vanishes when n goes to infinity, since u is locally integrable. Letting n go to infinity in equation (21), we obtain the assertion. \square

6.3. Existence and uniqueness of the process $(X_t^N)_t$

Proposition 6.6. *If assumptions 1 and 2 hold, the equation (6) admits a unique non-exploding strong solution.*

Proof. It is well known that if f is bounded, there is a unique strong solution of (6) (see Theorem IV.9.1 of Ikeda and Watanabe (1989)). In the general case we reason in a similar way as in the proof of Proposition 2 in Fournier and L ocherbach (2016). Consider the solution $(X_t^{N,K})_{t \in \mathbb{R}_+}$ of the equation (6) where f is replaced by $f_K : x \in \mathbb{R} \mapsto f(x) \wedge \sup_{|y| \leq K} f(y)$ for some $K \in \mathbb{N}^*$. Introduce moreover the stopping time

$$\tau_K^N = \inf \left\{ t \geq 0 : |X_t^{N,K}| \geq K \right\}.$$

Since for all $t \in [0, \tau_K^N \wedge \tau_{K+1}^N]$, $X_t^{N,K} = X_t^{N,K+1}$, we know that $\tau_K^N(\omega) \leq \tau_{K+1}^N(\omega)$ for all ω . Then we can define τ^N as the non-decreasing limit of τ_K^N . Since we can bound $\mathbb{E} \left[\left(X_{t \wedge \tau_K^N}^{N,K} \right)^2 \right]$

uniformly in K (see inequality (22) in the proof of Lemma 3.1), we know that τ^N equals infinity almost surely. So we can simply define X_t^N as the limit of $X_t^{N,K}$, as K goes to infinity. Now we show that the trajectories of X^N satisfy equation (6). Consider some $\omega \in \Omega$ and $t > 0$, and choose K such that $\tau_K^N(\omega) > t$. Then we know that for all $s \in [0, t]$, $X_s^N(\omega) = X_s^{N,K}(\omega)$ and $f(X_{s-}^N(\omega)) = f_K(X_{s-}^{N,K}(\omega))$. Moreover, as $X^{N,K}(\omega)$ satisfies the equation (6) with f replaced by f_K , we know that $X^N(\omega)$ verifies the equation (6) on $[0, t]$. This holds for all $t > 0$. As a consequence, we know that X^N satisfies the equation (6). This proves the existence of strong solution. The uniqueness is a consequence of the uniqueness of strong solutions of (6), if we replace f by f_K in (6), and of the fact that any strong solution $(Y_t^N)_t$ equals necessarily $(X_t^{N,K})_t$ on $[0, \tau_K^N]$. \square

6.4. Proof of Lemma 3.1

We just prove the points (i) and (iii) because (ii), (iv), (v) and (vi) are similar. We begin with (i). In a first time we prove the result for the process $(X_{t \wedge \tau_K^N}^{N,K})_t$ introduced in the proof of Proposition 6.6.

Applying Ito's formula to $(X_t^{N,K})_t$ with the function $x \mapsto x^2$, we obtain

$$\begin{aligned} (X_t^{N,K})^2 &= (X_0^{N,K})^2 - 2\alpha \int_0^t (X_s^{N,K})^2 ds \\ &\quad + \sum_{j=1}^N \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} \left[\frac{u^2}{N} + 2X_{s-}^{N,K} \frac{u}{\sqrt{N}} \right] \mathbb{1}_{\{z \leq f_K(X_{s-}^{N,K})\}} d\pi_j(s, z, u). \end{aligned}$$

Let $m_t^{N,K} = \mathbb{E} \left[(X_{t \wedge \tau_K^N}^{N,K})^2 \right]$. As μ is centered, we have

$$\begin{aligned} m_t^{N,K} &= m_0^N - 2\alpha \int_0^{t \wedge \tau_K^N} m_s^{N,K} ds + \sum_{j=1}^N \mathbb{E} \left[\int_0^{t \wedge \tau_K^N} \int_{\mathbb{R}} f_K(X_{s-}^{N,K}) \left[\frac{u^2}{N} + 2X_{s-}^{N,K} \frac{u}{\sqrt{N}} \right] d\mu(u) ds \right] \\ &= m_0^N - 2\alpha \int_0^{t \wedge \tau_K^N} m_s^{N,K} ds + \sigma^2 \int_0^{t \wedge \tau_K^N} \mathbb{E} \left[f_K(X_{s-}^{N,K}) \right] ds \\ &\leq m_0^N - 2\alpha \int_0^t m_s^{N,K} ds + \sigma^2 Bt + \sigma^2 A \int_0^t m_s^{N,K} ds \\ &= m_0^N + \sigma^2 Bt + (\sigma^2 A - 2\alpha) \int_0^t m_s^{N,K} ds. \end{aligned}$$

Then by Grönwall's lemma (see Lemma 6.5),

$$\sup_{0 \leq t \leq T} m_t^{N,K} \leq (m_0^N + \sigma^2 BT) e^{T \max(\sigma^2 A - 2\alpha, 0)}. \quad (22)$$

Letting $K \rightarrow \infty$ implies the result, because $\lim_{K \rightarrow \infty} \tau_K^N = +\infty$, which is a consequence of (22).

Now we prove the point (iii). We note M_t^N the local martingale

$$M_t^N = \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{s-}^N)\}} d\pi_j(s, z, u).$$

Obviously, $|X_t^N| \leq |X_0^N| + |M_t^N| + \alpha \int_0^t |X_s^N| ds$. Applying directly Lemma 6.5 (Grönwall's lemma) to the previous inequality, we obtain for all $t \leq T$,

$$\begin{aligned} |X_t^N| &\leq |X_0^N| + |M_t^N| + \alpha \int_0^t (|X_0^N| + |M_s^N|) e^{\alpha(t-s)} ds \\ &\leq \left(|X_0^N| + \sup_{0 \leq t \leq T} |M_t^N| \right) \left(1 + \alpha e^{\alpha T} \int_0^T e^{-\alpha s} ds \right) \\ &\leq \left(|X_0^N| + \sup_{0 \leq t \leq T} |M_t^N| \right) (1 + e^{\alpha T}). \end{aligned}$$

Thus,

$$\mathbb{E} \left[\left(\sup_{0 \leq t \leq T} |X_t^N| \right)^2 \right] \leq 2(1 + e^{\alpha T})^2 \left(\sup_N \mathbb{E} \left[(X_0^N)^2 \right] + \mathbb{E} \left[\left(\sup_{0 \leq t \leq T} |M_t^N| \right)^2 \right] \right).$$

As $(M_t^N)_t$ is a local martingale, we can apply Burkholder-Davis-Gundy inequality. So we know that there exists a constant C_2 such that

$$\mathbb{E} \left[\left(\sup_{0 \leq t \leq T} |X_t^N| \right)^2 \right] \leq 2(1 + e^{\alpha T})^2 \left(\sup_N \mathbb{E} \left[(X_0^N)^2 \right] + C_2 \mathbb{E} \left[[M^N]_T \right] \right), \quad (23)$$

where

$$\mathbb{E} \left[[M^N]_T \right] = \mathbb{E} \left[\int_{]0,T] \times \mathbb{R}_+ \times \mathbb{R}} u^2 \mathbf{1}_{\{z \leq f(X_{s-}^N)\}} d\pi_1(s, z, u) \right] = \sigma^2 \int_0^T \mathbb{E} \left[f(X_s^N) \right] ds \leq \sigma^2 T (A\Gamma_T + B),$$

with $\Gamma_T = C_T \left(\sup_N \mathbb{E} \left[(X_0^N)^2 \right] \right) + D_T$ using the point (i) of Lemma 3.1. Using the last inequality above in (23), we have:

$$\mathbb{E} \left[\left(\sup_{0 \leq t \leq T} |X_t^N| \right)^2 \right] \leq 2(1 + e^{\alpha T})^2 \left(\sup_N \mathbb{E} \left[(X_0^N)^2 \right] + C_2 \sigma^2 T (A\Gamma_T + B) \right).$$

6.5. Proof of Proposition 3.6

To begin with, we use Theorem 1.4.1 of Kunita (1986) to prove that the flow associated to the SDE (7) admits a modification which is C^3 with respect to the initial condition x (see also Theorem 4.6.5 of Kunita (1990)). Indeed the local characteristics of the flow are given by

$$b(x, t) = -\alpha x \quad \text{and} \quad a(x, y, t) = \sigma^2 \sqrt{f(x)f(y)},$$

and, under Assumptions 1 and 3, they satisfy the conditions of Theorem 1.4.1 of Kunita (1986):

- $\exists \Gamma, \forall x, y, t, |b(x, t)| \leq \Gamma(1 + |x|)$ and $|a(x, y, t)| \leq \Gamma(1 + |x|)(1 + |y|)$.
- $\exists \Gamma, \forall x, y, t, |b(x, t) - b(y, t)| \leq \Gamma|x - y|$ and $|a(x, x, t) + a(y, y, t) - 2a(x, y, t)| \leq \Gamma|x - y|^2$.
- $\forall 1 \leq k, l \leq 4, \frac{\partial^k}{\partial x^k} b(x, t)$ and $\frac{\partial^{k+l}}{\partial x^k \partial y^l} a(x, y, t)$ are bounded.

In the following, we consider the process $(\bar{X}_t^{(x)})_t$ that is a solution of the SDE (7) that satisfies $\bar{X}_0^{(x)} = x$. Then we can consider a modification of the flow $\bar{X}_t^{(x)}$ which is C^3 with the respect to the initial condition $x = \bar{X}_0^{(x)}$. It is then sufficient to control the moment of the derivatives of $\bar{X}_t^{(x)}$ with respect to x , since with those controls we will have

$$\bar{P}_t g(x) = \mathbb{E} \left[g \left(\bar{X}_t^{(x)} \right) \right], \quad (24)$$

$$(\bar{P}_t g)'(x) = \mathbb{E} \left[\frac{\partial \bar{X}_t^{(x)}}{\partial x} g' \left(\bar{X}_t^{(x)} \right) \right], \quad (25)$$

$$(\bar{P}_t g)''(x) = \mathbb{E} \left[\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} g' \left(\bar{X}_t^{(x)} \right) + \left(\frac{\partial \bar{X}_t^{(x)}}{\partial x} \right)^2 g'' \left(\bar{X}_t^{(x)} \right) \right], \quad (26)$$

$$(\bar{P}_t g)'''(x) = \mathbb{E} \left[\frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3} g' \left(\bar{X}_t^{(x)} \right) + 3 \frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} \cdot \frac{\partial \bar{X}_t^{(x)}}{\partial x} g'' \left(\bar{X}_t^{(x)} \right) + \left(\frac{\partial \bar{X}_t^{(x)}}{\partial x} \right)^3 g''' \left(\bar{X}_t^{(x)} \right) \right]. \quad (27)$$

We start with the representation

$$\bar{X}_t^{(x)} = x e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} \sqrt{f \left(\bar{X}_s^{(x)} \right)} dB_s.$$

This implies

$$\frac{\partial \bar{X}_t^{(x)}}{\partial x} = e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} \frac{\partial \bar{X}_s^{(x)}}{\partial x} \left(\sqrt{f} \right)' \left(\bar{X}_s^{(x)} \right) dB_s. \quad (28)$$

Writing $U_t = e^{\alpha t} \frac{\partial \bar{X}_t^{(x)}}{\partial x}$, we obtain

$$U_t = 1 + \sigma \int_0^t U_s \left(\sqrt{f} \right)' \left(\bar{X}_s^{(x)} \right) dB_s,$$

and then, by Ito's formula

$$(U_t)^8 = 1 + 8 \int_0^t (U_s)^7 dU_s + 28 \int_0^t (U_s)^6 d\langle U \rangle_t.$$

Truncating at level K as in the proof of Lemma 3.1 and then letting $K \rightarrow \infty$, we deduce from this

$$\mathbb{E} \left[(U_t)^8 \right] \leq 1 + 28\sigma^2 m_1^2 \int_0^t \mathbb{E} \left[(U_s)^8 \right] ds,$$

where m_1 is the bound of $(\sqrt{f})'$ introduced in Assumption 3.

By Grönwall's lemma, we obtain

$$\mathbb{E} \left[(U_t)^8 \right] \leq e^{28\sigma^2 m_1^2 t}, \quad (29)$$

whence

$$\mathbb{E} \left[\left(\frac{\partial \bar{X}_t^{(x)}}{\partial x} \right)^8 \right] \leq e^{t(28\sigma^2 m_1^2 - 8\alpha)}. \quad (30)$$

Using Hölder's inequality, this implies

$$\mathbb{E} \left[\left| \frac{\partial \bar{X}_t^{(x)}}{\partial x} \right| \right] \leq e^{t(\frac{7}{2}\sigma^2 m_1^2 - \alpha)}.$$

With the notations of the lemma, using (25) we have shown that

$$\beta^{(1)} = \frac{7}{2}\sigma^2 m_1^2 - \alpha.$$

Differentiating (28) with respect to x , we obtain

$$\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} = \sigma \int_0^t e^{-\alpha(t-s)} \left[\frac{\partial^2 \bar{X}_s^{(x)}}{\partial x^2} (\sqrt{f})' (\bar{X}_s^{(x)}) + \left(\frac{\partial \bar{X}_s^{(x)}}{\partial x} \right)^2 (\sqrt{f})^{(2)} (\bar{X}_s^{(x)}) \right] dB_s. \quad (31)$$

We introduce $V_t = \frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} e^{\alpha t}$, and deduce from this that

$$V_t = \sigma \int_0^t \left[V_s (\sqrt{f})' (\bar{X}_s^{(x)}) + e^{-\alpha s} U_s (\sqrt{f})^{(2)} (\bar{X}_s^{(x)}) \right] dB_s,$$

implying

$$(V_t)^4 = 4 \int_0^t (V_s)^3 dV_s + 6 \int_0^t (V_s)^2 d\langle V \rangle_s,$$

and thus

$$\begin{aligned} \mathbb{E} \left[(V_t)^4 \right] &\leq 12\sigma^2 \mathbb{E} \left[\int_0^t \left[(V_s)^4 \left((\sqrt{f})' (\bar{X}_s^{(x)}) \right)^2 + (V_s)^2 (U_s)^2 \left((\sqrt{f})^{(2)} (\bar{X}_s^{(x)}) \right)^2 \right] ds \right] \\ &\leq 12\sigma^2 \left(m_1^2 + \frac{m_2^2}{2} \right) \int_0^t \mathbb{E} \left[(V_s)^4 \right] ds + 6\sigma^2 m_2^2 \int_0^t \mathbb{E} \left[(U_s)^4 \right] ds \\ &\leq 12\sigma^2 \left(m_1^2 + \frac{m_2^2}{2} \right) \int_0^t \mathbb{E} \left[(V_s)^4 \right] ds + 6\sigma^2 m_2^2 e^{14\sigma^2 m_1^2 t}. \end{aligned}$$

In the last inequality above, we have used (29).

By Grönwall's lemma,

$$\mathbb{E} \left[(V_t)^4 \right] \leq 6\sigma^2 m_2^2 e^{14\sigma^2 m_1^2 t} e^{12\sigma^2 \left(m_1^2 + \frac{m_2^2}{2} \right) t}, \quad (32)$$

and thus

$$\mathbb{E} \left[\left(\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} \right)^4 \right] \leq 6\sigma^2 m_2^2 e^{(26\sigma^2 m_1^2 + 6\sigma^2 m_2^2 - 4\alpha)t}. \quad (33)$$

In particular, using Hölder's inequality in (30) and (33) to bound respectively $\mathbb{E} \left[\left(\frac{\partial \bar{X}_t^{(x)}}{\partial x} \right)^2 \right]$ and $\mathbb{E} \left[\left| \frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} \right| \right]$, and inserting these bounds in (26), we know that

$$\begin{aligned} \beta^{(2)} &= \max \left(7\sigma^2 m_1^2 - 2\alpha, \frac{13}{2}\sigma^2 m_1^2 + \frac{3}{2}\sigma^2 m_2^2 - \alpha \right) \\ &\leq 7\sigma^2 m_1^2 + \frac{3}{2}\sigma^2 m_2^2 - \alpha. \end{aligned}$$

Finally, differentiating (31),

$$\begin{aligned} \frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3} &= \sigma \int_0^t e^{-\alpha(t-s)} \left[\frac{\partial^3 \bar{X}_s^{(x)}}{\partial x^3} (\sqrt{f})' (\bar{X}_s^{(x)}) + 3 \frac{\partial^2 \bar{X}_s^{(x)}}{\partial x^2} \frac{\partial \bar{X}_s^{(x)}}{\partial x} (\sqrt{f})^{(2)} (\bar{X}_s^{(x)}) \right. \\ &\quad \left. + \left(\frac{\partial \bar{X}_s^{(x)}}{\partial x} \right)^3 (\sqrt{f})^{(3)} (\bar{X}_s^{(x)}) \right] dB_s. \end{aligned}$$

With $W_t = e^{\alpha t} \frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3}$, we obtain

$$W_t = \sigma \int_0^t \left[W_s (\sqrt{f})' (\bar{X}_s^{(x)}) + 3e^{-\alpha s} U_s V_s (\sqrt{f})^{(2)} (\bar{X}_s^{(x)}) + e^{-2\alpha s} U_s^3 (\sqrt{f})^{(3)} (\bar{X}_s^{(x)}) \right] dB_s.$$

Thus

$$\mathbb{E} [(W_t)^2] = 2 \int_0^t W_s dW_s + \int_0^t d\langle W \rangle_s$$

and

$$\begin{aligned} \mathbb{E} [(W_t)^2] &\leq 3\sigma^2 m_1^2 \int_0^t \mathbb{E} [(W_s)^2] ds + \frac{9}{2}\sigma^2 m_2^2 \int_0^t \mathbb{E} [(U_s)^4] ds \\ &\quad + \frac{9}{2}\sigma^2 m_2^2 \int_0^t \mathbb{E} [(V_s)^4] ds + 3\sigma^2 m_3^2 \int_0^t \mathbb{E} [(U_s)^6] ds. \end{aligned}$$

Grönwall's lemma, (29), (32) and Hölder's inequality imply

$$\mathbb{E} [(W_t)^2] \leq \Gamma e^{(24\sigma^2 m_1^2 + 6\sigma^2 m_2^2)t}.$$

As a consequence,

$$\mathbb{E} \left[\left(\frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3} \right)^2 \right] \leq \Gamma e^{(24\sigma^2 m_1^2 + 6\sigma^2 m_2^2 - 2\alpha)t}. \quad (34)$$

To find $\beta^{(3)}$, we use (27). We bound $\mathbb{E} \left[\left| \frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3} \right| \right]$ using Cauchy-Schwarz's inequality and (34), we also bound $\mathbb{E} \left[\left| \frac{\partial \bar{X}_t^{(x)}}{\partial x} \right|^3 \right]$ using Hölder's inequality and (30), and we bound $\mathbb{E} \left[\left| \frac{\partial \bar{X}_t^{(x)}}{\partial x} \cdot \frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} \right| \right]$ by $\mathbb{E} \left[\left(\frac{\partial \bar{X}_t^{(x)}}{\partial x} \right)^2 \right]^{1/2} \mathbb{E} \left[\left(\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} \right)^2 \right]^{1/2}$, and then apply Hölder's inequality to both terms of the product and use (30) and (33). So we know that $\beta^{(3)}$ has to be bigger than the maximum of the three terms:

- $12\sigma^2 m_1^2 + 3\sigma^2 m_2^2 - \alpha$
- $10\sigma^2 m_1^2 + \frac{3}{2}\sigma^2 m_2^2 - \alpha$
- $\frac{21}{2}\sigma^2 m_1^2 - 3\alpha$

Since the second term and the third one are smaller than the first one, we have

$$\beta^{(3)} = 12\sigma^2 m_1^2 + 3\sigma^2 m_2^2 - \alpha.$$

6.6. Proof of Proposition 3.9

In a first time we prove the result only for functions $g_1, \dots, g_n \in C_b^3(\mathbb{R})$ by induction on n . For $n = 1$ we use Theorem 2.2, Proposition 4.1 and the Kantorovich-Rubinstein duality (see Remark 6.5 of Villani (2008)) to obtain

$$\begin{aligned} |\mathbb{E}[g(X_t^N)] - \mathbb{E}[g(\bar{X}_t)]| &= |\nu_0^N P_t(g) - \bar{\nu}_0 \bar{P}_t(g)| \\ &\leq |\nu_0^N P_t(g) - \nu_0^N \bar{P}_t(g)| + |\nu_0^N \bar{P}_t(g) - \bar{\nu}_0 \bar{P}_t(g)| \\ &\leq \int_{\mathbb{R}} |P_t^N g(x) - \bar{P}_t g(x)| d\nu_0^N(x) + \|g'\|_{\infty} W_1(\nu_0^N \bar{P}_t, \bar{\nu}_0 \bar{P}_t) \\ &\leq \|g\|_{3,\infty} \frac{1}{\sqrt{N}} K_t \left(1 + \mathbb{E}[(X_0^N)^2]\right) + \|g'\|_{\infty} e^{(\frac{1}{2}\sigma^2 L^2 - \alpha)t} W_2(\nu_0^N, \bar{\nu}_0). \end{aligned}$$

We now show the inductive step. We know that $\mathbb{E}[g_1(X_{t_1}^N) \dots g_n(X_{t_n}^N)]$ equals

$$\begin{aligned} &\int d\nu_0^N(x) \int P_{t_1}^N(x, dx_1) g_1(x_1) \int P_{t_2-t_1}^N(x_1, dx_2) g_2(x_2) \dots \int P_{t_n-t_{n-1}}^N(x_{n-1}, dx_n) g_n(x_n) \\ &= \int d\nu_0^N(x) \int P_{t_1}^N(x, dx_1) g_1(x_1) \int P_{t_2-t_1}^N(x_1, dx_2) g_2(x_2) \dots \\ &\quad \int P_{t_{n-1}-t_{n-2}}^N(x_{n-2}, dx_{n-1}) \left(g_{n-1} \cdot P_{t_n-t_{n-1}}^N g_n \right) (x_{n-1}) \\ &= \mathbb{E} \left[g_1(X_{t_1}^N) \dots g_{n-2}(X_{t_{n-2}}^N) \left(g_{n-1} \cdot P_{t_n-t_{n-1}}^N g_n \right) (X_{t_{n-1}}^N) \right]. \end{aligned}$$

Analogously

$$\mathbb{E}[g_1(\bar{X}_{t_1}) \dots g_n(\bar{X}_{t_n})] = \mathbb{E}[g_1(\bar{X}_{t_1}) \dots g_{n-2}(\bar{X}_{t_{n-2}}) (g_{n-1} \cdot \bar{P}_{t_n-t_{n-1}} g_n) (\bar{X}_{t_{n-1}})].$$

Then, by triangle inequality,

$$|\mathbb{E} [g_1(X_{t_1}^N) \cdots g_n(X_{t_n}^N)] - \mathbb{E} [g_1(\bar{X}_{t_1}) \cdots g_n(\bar{X}_{t_n})]| \leq$$

$$\left| \mathbb{E} \left[g_1(X_{t_1}^N) \cdots g_{n-2}(X_{t_{n-2}}^N) \left((g_{n-1} \cdot P_{t_n-t_{n-1}}^N g_n)(X_{t_{n-1}}^N) - (g_{n-1} \cdot \bar{P}_{t_n-t_{n-1}} g_n)(X_{t_{n-1}}^N) \right) \right] \right| \quad (35)$$

$$+ \left| \mathbb{E} \left[g_1(X_{t_1}^N) \cdots g_{n-2}(X_{t_{n-2}}^N) (g_{n-1} \cdot \bar{P}_{t_n-t_{n-1}} g_n)(X_{t_{n-1}}^N) \right] - \mathbb{E} \left[g_1(\bar{X}_{t_1}) \cdots g_{n-2}(\bar{X}_{t_{n-2}}) (g_{n-1} \cdot \bar{P}_{t_n-t_{n-1}} g_n)(\bar{X}_{t_{n-1}}) \right] \right|. \quad (36)$$

As $g_{n-1} \cdot \bar{P}_{t_n-t_{n-1}} g_n$ is in $C_b^3(\mathbb{R})$ (see Proposition 3.6), we know that (36) goes to 0 when N goes to infinity (by induction hypothesis).

Moreover, we can bound (35) by

$$\begin{aligned} & \|g_1\|_\infty \cdots \|g_{n-1}\|_\infty \mathbb{E} \left[\left| (P_{t_n-t_{n-1}}^N - P_{t_n-t_{n-1}}) g_n(X_{t_{n-1}}^N) \right| \right] \leq \\ & \|g_1\|_\infty \cdots \|g_{n-1}\|_\infty \left(1 + \mathbb{E} \left[(X_{t_{n-1}}^N)^2 \right] \right) K_T \|g_n\|_{3,\infty} \frac{1}{\sqrt{N}}, \end{aligned}$$

which goes to 0 when N goes to infinity (using Lemma 3.1).

Since $C_b^3(\mathbb{R})$ is dense in $C_b(\mathbb{R})$, standard arguments allow to conclude that $\forall n \in \mathbb{N}^*, \forall g_1, \dots, g_n \in C_b(\mathbb{R}), \forall 0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq T$,

$$\mathbb{E} [g_1(X_{t_1}^N) \cdots g_n(X_{t_n}^N)] \xrightarrow[N \rightarrow +\infty]{} \mathbb{E} [g_1(\bar{X}_{t_1}) \cdots g_n(\bar{X}_{t_n})].$$

6.7. Proof of Proposition 5.2

We just prove the proposition for $k = 1$ to simplify the proof, but the general case is almost the same.

Recall that $D(\mathbb{R}_+, \mathbb{R})$ is separable and complete (see Theorem 16.3 of Billingsley (1999)), and $\mathcal{M}^\#$ is also separable and complete (Theorem A2.6.III.(i) of Daley and Vere-Jones (2003)). Hence the product of the metric spaces $(D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M}^\#)$ is also separable and complete. Since the sequence $(X_N)_N$ is tight on $D(\mathbb{R}_+, \mathbb{R})$ and π^1 is tight on $\mathcal{M}^\#$, (see Theorem 1.3 of Billingsley (1999)), the couple (X_N, π^1) is tight on $(D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M}^\#)$.

Thus it suffices to show that any weakly convergent subsequence of $\mathcal{D}(X^N, \pi_1)$ converges to $\mathcal{D}(\bar{X}) \otimes \mathcal{D}(\pi^1)$ (see Corollary of Theorem 5.1 of Billingsley (1999)). To simplify the notations we assume that $\mathcal{D}(X^N, \pi^1)$ is already a weakly-converging subsequence, converging to some limit P .

Let $(Y, \pi) \in (D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M}^\#)$ such that $(Y, \pi) \sim P$. It is easy to see that

$$Y \sim \bar{X} \text{ and } \pi \sim \pi^1,$$

but we do not know yet if both are independent.

In the sequel we suppose that (Y, π) is defined on a filtered probability space $(\Omega', \mathcal{A}', (\mathcal{F}_t)_{t \geq 0}, P')$, where

$$\mathcal{F}_t' = \bigcap_{T > t} \mathcal{F}_T^0, \mathcal{F}_t^0 = \sigma(Y_s, \pi([0, s] \times A), A \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}), s \leq t).$$

Step 1. We show that π is a $(P', (\mathcal{F}_t^0)_{t \geq 0})$ -Poisson random measure on $[0, +\infty[\times \mathbb{R}_+ \times \mathbb{R}$, with non-random compensator measure $dt \times \nu$ where $\nu = dz \times \mu(du)$.

For that sake, it is sufficient to show that for all $s < t$, disjoint sets $U_1, \dots, U_k \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R})$, and $\lambda_1, \dots, \lambda_k \geq 0$,

$$\mathbb{E} \left(\exp \left[- \sum_{i=1}^k \lambda_i \pi([s, t] \times U_i) \right] \middle| \mathcal{F}_s^0 \right) = \exp \left[(t-s) \sum_{i=1}^k (e^{-\lambda_i} - 1) \nu(U_i) \right]. \quad (37)$$

To prove (37), it suffices to show that for all $s_1 < \dots < s_n < s$, all bounded $\varphi_1, \dots, \varphi_n$, disjoint sets $U_1, \dots, U_k \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R})$, and sets $V_1, \dots, V_n \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R})$,

$$\begin{aligned} & \mathbb{E} \left(\exp \left[- \sum_{i=1}^k \lambda_i \pi([s, t] \times U_i) \right] \varphi_1(Z_{s_1}) \times \dots \times \varphi_n(Z_{s_n}) \right) = \\ & \exp \left[(t-s) \sum_{i=1}^k (e^{-\lambda_i} - 1) \nu(U_i) \right] \mathbb{E} (\varphi_1(Z_{s_1}) \times \dots \times \varphi_n(Z_{s_n})), \end{aligned} \quad (38)$$

where $Z_{s_i} = (Y_{s_i}, \pi([0, s_i] \times V_i))$

The previous equality holds if we replace Y by X^N and π by π_1 , because $\pi^1([s, t] \times U_i)$ and $Z_{s_1}^N, \dots, Z_{s_n}^N$ are independent, where $Z_{s_i}^N = (X_{s_i}^N, \pi_1([0, s_i] \times V_i))$.

This implies that π is a $(P', (\mathcal{F}_t^0)_t)$ -Poisson random measure. By right continuity of $s \mapsto \exp \left[(t-s) \sum_{i=1}^m (e^{-\lambda_i} - 1) \nu(U_i) \right]$, this implies that π is also a Poisson random measure with respect to $(P', (\mathcal{F}_t^0)_t)$.

Step 2. Fix a test function $\varphi \in C_b^3$. Now we show that

$$\varphi(Y_t) - \varphi(Y_0) + \alpha \int_0^t \varphi'(Y_v) Y_v dv - \frac{\sigma^2}{2} \int_0^t \varphi''(Y_v) f(Y_v) dv \quad (39)$$

is a $(\mathcal{F}_t^0)_t$ -martingale. Fix $s_1 < s_2 < \dots < s_n \leq s < t$ together with continuous and bounded test functions ψ_i and disjoint sets $U_1, \dots, U_n \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R})$. Denote $Z_{s_i} = (Y_{s_i}, \pi([0, s_i] \times U_i))$. It suffices to show that

$$\mathbb{E} \left[\left(\varphi(Y_t) - \varphi(Y_s) + \alpha \int_s^t \varphi'(Y_v) Y_v dv - \frac{\sigma^2}{2} \int_s^t \varphi''(Y_v) f(Y_v) dv \right) \prod_{i=1}^n \psi_i(Z_{s_i}) \right] = 0. \quad (40)$$

To prove (40), we shall use that

$$W_t^N = \varphi(X_t^N) + \alpha \int_0^t \varphi'(X_v^N) X_v^N dv - N \int_0^t dv \int_{\mathbb{R}} d\mu(u) \left[\varphi \left(X_v^N + \frac{u}{\sqrt{N}} \right) - \varphi(X_v^N) \right] f(X_v^N)$$

is a $(\mathcal{F}_t^0)_t$ -martingale. As a consequence, for all $N \geq 1$,

$$\mathbb{E} \left[(W_t^N - W_s^N) \prod_{k=1}^n \psi_k(Z_{s_k}^N) \right] = 0.$$

Using the integral form of the remainder in Taylor's formula applied in the jump term of W_t^N , we can write $W_t^N - W_s^N$ as

$$\varphi(X_t^N) - \varphi(X_s^N) + \alpha \int_s^t \varphi'(X_v^N) X_v^N dv - \frac{\sigma^2}{2} \int_s^t \varphi''(X_v^N) f(X_v^N) dv + \frac{1}{\sqrt{N}} \Phi,$$

where Φ is a random variable whose expectation is bounded uniformly in N . Thus,

$$\mathbb{E} \left[(W_t^N - W_s^N) \prod_{k=1}^n \psi_k(Z_{s_k}^N) \right] = \mathbb{E} [F_{s,t}(X^N, \pi_1)] + \frac{1}{\sqrt{N}} \mathbb{E} \left[\Phi \prod_{k=1}^n \psi_k(Z_{s_k}^N) \right],$$

where

$$F_{s,t}(x, m) = \left(\varphi(x_t) - \varphi(x_s) + \alpha \int_s^t \varphi'(x_v) x_v dv - \frac{\sigma^2}{2} \int_s^t \varphi''(x_v) f(x_v) dv \right) \prod_{k=1}^n \psi_k(x_{s_k}, m([0, s_k] \times U_k))$$

is a continuous function on $D(\mathbb{R}_+, \mathbb{R}) \times \mathcal{M}^\#$. If $F_{s,t}$ was bounded we could make N go to infinity in the previous expression (since (X^N, π_1) converge in distribution to (Y, π)). So we have to truncate and consider $F_{s,t}^M(x, m) := F_{s,t}(x, m) \cdot \xi_M \left(\sup_{0 \leq r \leq t} |x_r| \right)$, where $\xi_M : \mathbb{R} \rightarrow [0, 1]$ is C^∞ and verifies

$$\mathbb{1}_{\{|x| \leq M\}} \leq \xi_M(x) \leq \mathbb{1}_{\{|x| \leq M+1\}}.$$

Recall that we want to show (40), that is, $\mathbb{E} [F_{s,t}(Y, \pi)] = 0$. We start from

$$\begin{aligned} |\mathbb{E} [F_{s,t}(Y, \pi)]| &= \left| \mathbb{E} [F_{s,t}(Y, \pi)] - \mathbb{E} \left[(W_t^N - W_s^N) \prod_{k=1}^N \psi_k(Z_{s_k}^N) \right] \right| \\ &\leq \left| \mathbb{E} \left[F_{s,t}(Y, \pi) \left(1 - \xi_M \left(\sup_{0 \leq r \leq t} |Y_r| \right) \right) \right] \right| \end{aligned} \quad (41)$$

$$+ \left| \mathbb{E} \left[F_{s,t}(Y, \pi) \xi_M \left(\sup_{0 \leq r \leq t} |Y_r| \right) \right] - \mathbb{E} \left[F_{s,t}(X^N, \pi_1) \xi_M \left(\sup_{0 \leq r \leq t} |X_r^N| \right) \right] \right| \quad (42)$$

$$+ \left| \mathbb{E} \left[F_{s,t}(X^N, \pi_1) \left(1 - \xi_M \left(\sup_{0 \leq r \leq t} |X_r^N| \right) \right) \right] \right|. \quad (43)$$

Using the fact that $1 - \xi_M(x) \leq \mathbb{1}_{\{|x| > M\}}$, Cauchy-Schwarz's inequality, Markov's inequality and Lemma 3.1, we can bound (41) and (43) by Γ/\sqrt{M} for some $\Gamma > 0$ that is independent of N .

Now, fix some $\varepsilon > 0$ and consider a constant $M_\varepsilon > 0$ such that (41) and (43) are smaller than ε . In a next step, we choose an integer N_ε big enough such that (42) is smaller than ε . As a consequence, $|\mathbb{E} [F_{s,t}(Y, \pi)]| \leq 3\varepsilon$ for all $\varepsilon > 0$, whence $\mathbb{E} [F_{s,t}(Y, \pi)] = 0$ which means that for all $\varphi \in C_b^3(\mathbb{R})$, the expression (39) is a $(\mathcal{F}_t^0)_t$ -martingale.

In the following we need to prove that for all $\varphi \in C^3$ (not necessarily bounded), the expression (39) is a $(\mathcal{F}_t^0)_t$ -local martingale. So we introduce the stopping times $\tau_K = \inf\{t > 0 : |Y_t| > K\}$, and for $\varphi \in C^3(\mathbb{R})$, we define $\varphi_K \in C_c^3(\mathbb{R})$ by $\varphi_K(x) = \varphi(x)\xi_K(x)$. Now if $F_{s,t}^{\varphi}$ denotes the function $F_{s,t}$ we used previously, by definition of F, τ_K and φ_K , we know that $\mathbb{E} [F_{s \wedge \tau_K, t \wedge \tau_K}^{\varphi}(Y, \pi)] = \mathbb{E} [F_{s \wedge \tau_K, t \wedge \tau_K}^{\varphi_K}(Y, \pi)]$ which equals 0, since the expression (39) with $\varphi_K \in C_b^3(\mathbb{R})$ is a martingale.

Hence we have shown that the expression in (39) is a $(\mathcal{F}_t^0)_t$ -martingale if $\varphi \in C_b^3(\mathbb{R})$, and that it is a $(\mathcal{F}_t^0)_t$ -local martingale if $\varphi \in C^3(\mathbb{R})$. By right-continuity of $s \mapsto Y_s$, this implies that the expression in (39) is martingale (resp. local martingale) with respect to $(\mathcal{F}_t^0)_t$ for $\varphi \in C_b^3(\mathbb{R})$ (resp. $\varphi \in C^3(\mathbb{R})$).

Step 3 . Now we show that Y and π are independent. By Theorem II.2.42 of Jacod and Shiryaev (2003), step 2 implies that Y is a $(P', (\mathcal{F}_t^0)_{t \geq 0})$ -semi-martingale with characteristics

$B_t = -\alpha \int_0^t Y_s ds$, $\nu(ds, dx) = 0$, $C_t = \int_0^t \sigma^2 f(Y_s) ds$. Moreover, Theorem III.2.26 of [Jacod and Shiryaev \(2003\)](#) implies that there exists a Brownian motion B' defined on $(\Omega', \mathcal{A}', (\mathcal{G}_t)_{t \geq 0}, P')$, such that Y is solution of

$$Y_t = Y_0 - \alpha \int_0^t Y_s ds + \sigma \int_0^t \sqrt{f(Y_s)} dB'_s.$$

So B' is defined on the same space, but for the moment we do not know that this Brownian motion is indeed a Brownian with respect to the filtration we are interested in, that is, with respect to $(\mathcal{F}'_t)_{t \geq 0}$. To understand this last point we use the Lamperti transform. To do so, we need to introduce

$$h(x) := \int_0^x \frac{1}{\sqrt{f(t)}} dt.$$

Using Ito's formula, one gets that $\tilde{Y}_t := h(Y_t)$ solves

$$d\tilde{Y}_t = -\alpha h'(Y_t) Y_t dt + \sigma h'(Y_t) \sqrt{f(Y_t)} dB'_t + \frac{\sigma^2}{2} h''(Y_t) f(Y_t) dt.$$

In other words,

$$\sigma B'_t = h(Y_t) - h(Y_0) + \alpha \int_0^t h'(Y_s) Y_s ds - \frac{\sigma^2}{2} \int_0^t h''(Y_s) f(Y_s) ds$$

is exactly of the form as in (39), for the test-function $\varphi = h$ that is C^3 . Thus we know that $(B'_t)_t$ is a $(P', (\mathcal{F}'_t)_{t \geq 0})$ -local martingale.

By Theorem II.6.3 of [Ikeda and Watanabe \(1989\)](#) we can then conclude that B' and the Poisson random measure π - which are defined with respect to the same filtration, living on the same space - are independent, and thus also Y and π .

6.8. Lemmas on Skorohod space

Lemma 6.7. *Let $(x_N)_N$ be a sequence of $D(\mathbb{R}_+, \mathbb{R})$ that converges to some $x \in D(\mathbb{R}_+, \mathbb{R})$, and a sequence $(t_N)_N$ that converges to $t > 0$. If x is continuous on t , then $x_N(t_N-) \rightarrow x(t)$.*

Lemma 6.8. *Let $T > 0$, $k \in \mathbb{N}^*$, increasing sequences $0 = t_{i,0} < t_{i,1} < \dots < t_{i,n_i-1} < t_{i,n_i} = T$ ($1 \leq i \leq k$), $0 = t_{i,0}^N < t_{i,1}^N < \dots < t_{i,n_i^N-1}^N < t_{i,n_i^N}^N = T$ ($1 \leq i \leq k$), and we define the functions $g, g_N \in D([0, T], \mathbb{R}^k)$ by*

$$\begin{cases} g(t) = \left(\sum_{j=0}^{n_i-1} \mathbb{1}_{[t_{i,j}, t_{i,j+1}]}(t) j \right)_{1 \leq i \leq k} & t \in [0, T[, \\ g(T) = (n_i - 1)_{1 \leq i \leq k}, \end{cases}$$

$$\begin{cases} g_N(t) = \left(\sum_{j=0}^{n_i^N-1} \mathbb{1}_{[t_{i,j}^N, t_{i,j+1}^N]}(t) j \right)_{1 \leq i \leq k} & t \in [0, T[, \\ g_N(T) = (n_i^N - 1)_{1 \leq i \leq k}. \end{cases}$$

We assume that there exists a dense subset $A \subseteq [0, T]$ that contains T such that, for all $t \in A$, $g_N(t)$ converges to $g(t)$, and we assume that for all $i_1 \neq i_2$ for all $j_1 \in \llbracket 1, n_{i_1-1} \rrbracket$ and $j_2 \in \llbracket 1, n_{i_2-1} \rrbracket$, $t_{i_1, j_1} \neq t_{i_2, j_2}$. Then g_N converges to g in $D([0, T], \mathbb{R}^k)$.

Bibliography

- AÏT-SAHALIA, Y., CACHO-DIAZ, J. and LAEVEN, R. J. A. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics* **117** 585-606.
- BACRY, E. and MUZY, J. F. (2016). Second order statistics characterization of Hawkes processes and non-parametric estimation. *Trans. in Inf. Theory* **2**.
- BAUWENS, L. and HAUTSCH, N. (2009). *Modelling financial high frequency data using point processes*. Springer Berlin Heidelberg.
- 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. *Ann. Probab.* **44** 3740-3803.
- COSTA, M., GRAHAM, C., MARSALLE, L. and TRAN, V. C. (2018). Renewal in Hawkes processes with self-excitation and inhibition. *arXiv e-prints* arXiv:1801.04645.
- 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.
- DAVIS, M. H. A. (1993). *Markov Models and Optimization*, First ed. Springer Science+Business Media Dordrecht.
- DELATTRE, S., FOURNIER, N. and HOFFMANN, M. (2016). Hawkes processes on large networks. *Ann. Appl. Probab.* **26** 216-261.
- DITLEVSEN, S. and LÖCHERBACH, E. (2017). Multi-class Oscillating Systems of Interacting Neurons. *Stochastic Processes and their Applications* **127** 1840-1869.
- DUARTE, A., LÖCHERBACH, E. and OST, G. (2018). Stability, Convergence to Equilibrium and Simulation of Non-linear Hawkes Processes with Memory Kernels given by the Sum of Erlang Kernels. *ArXiv e-prints* arXiv:1610.03300v3.
- 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.
- GRÜN, S., DIEDSMANN, M. and AERTSEN, A. M. (2010). *Analysis of Parallel Spike Trains*. Rotter, Springer series in computational neurosciences.
- HAWKES, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58** 83-90.
- HAWKES, A. G. and OAKES, D. (1974). A Cluster Process Representation of a Self-Exciting Process. *Journal of Applied Probability* **11** 493-503.
- HELMSTETTER, A. and SORNETTE, D. (2002). Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *Journal of Geophysical Research* **107**.
- HEWLETT, P. (2006). Clustering of order arrivals, price impact and trade path optimisation. In Workshop on Financial Modeling with Jump processes. Ecole Polytechnique.
- IKEDA, N. and WATANABE, S. (1989). *Stochastic Differential Equations and Diffusion Processes*, Second ed. North-Holland Publishing Company.
- GRÖNWALL'S INEQUALITY (2019). Grönwall's inequality — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Grönwall's_inequality#Integral_form_with_locally_finite_measures.

- 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.
- KUNITA, H. (1990). *Stochastic flows and stochastic differential equations*. Cambridge University Press.
- LU, X. and ABERGEL, F. (2018). High dimensional Hawkes processes for limit order books Modelling, empirical analysis and numerical calibration. *Quantitative Finance* 1-16.
- MEYN, S. P. and TWEEDIE, R. L. (1993). Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes. *Applied Probability Trust* **25** 518-548.
- OGATA, Y. (1978). The asymptotic behavior of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics* **30** 243-261.
- OGATA, Y. (1999). Seismicity Analysis through Point-process Modeling: A Review. *pure and applied geophysics* **155** 471-507.
- OKATAN, M., A WILSON, M. and N BROWN, E. (2005). Analyzing Functional Connectivity Using a Network Likelihood Model of Ensemble Neural Spiking Activity. *Neural computation* **17** 1927-61.
- PILLOW, J. W., WILSON, M. A. and BROWN, E. N. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* **454** 995-999.
- REYNAUD-BOURET, P. and SCHBATH, S. (2010). Adaptive estimation for Hawkes processes; application to genome analysis. *Ann. Statist.* **38** 2781-2822.
- REYNAUD-BOURET, P., RIVOIRARD, V., GRAMMONT, F. and TULEAU-MALOT, C. (2014). Goodness-of-Fit Tests and Nonparametric Adaptive Estimation for Spike Train Analysis. *Journal of Mathematical Neuroscience* **4** 3 - 330325.
- VILLANI, C. (2008). Optimal transport, old and new. Lecture notes for the 2005 Saint-Flour summer school.
- Y. KAGAN, Y. (2009). Statistical Distributions of Earthquake Numbers: Consequence of Branching Process. *Geophys. J. Int.* **180**.
- ZHOU, K., ZHA, H. and SONG, L. (2013). Learning triggering kernels for multi-dimensional Hawkes processes. Proceedings of the 30th International Conf. on Machine Learning (ICML).