



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-02096662v2

HAL Id: hal-02096662

<https://hal.science/hal-02096662v2>

Preprint submitted on 11 Nov 2019 (v2), 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^{1,*}, Eva Löcherbach^{2,**} and Dasha Loukianova^{1,†}

¹*Laboratoire de Mathématiques et Modélisation d'Évry, Université d'Évry Val d'Essonne, UMR CNRS 8071*
e-mail: *xavier.erny@univ-evry.fr; †dasha.loukianova@univ-evry.fr

²*Statistique, Analyse et Modélisation Multidisciplinaire, Université Paris 1 Panthéon-Sorbonne, EA 4543*
e-mail: **eva.locherbach@univ-paris1.fr

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 interactions, 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 networks ((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. More precisely, Z^N 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), (Costa et al. 2018), (Raad 2019) and (Graham 2019)). A number of authors studied the statistical inference for Hawkes processes ((Ogata 1978) and (Reynaud-Bouret and Schbath 2010)). Another field of study, very 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 in (Fournier and Löcherbach 2016), (Delattre, Fournier and Hoffmann 2016) and (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. Namely, in this model, the multivariate Hawkes processes $(Z^{N,i})_{1 \leq i \leq N}$ ($N \in \mathbb{N}^*$) are of the form

$$Z_t^{N,i} = \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} \mathbb{1}_{\{z \leq \lambda_s^N\}} d\pi_i(s, z, u), \quad 1 \leq i \leq N, \quad (2)$$

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 $dt dz d\mu(u)$ and μ is a centered probability measure on \mathbb{R} having a finite second moment σ^2 . The stochastic intensity of $Z^{N,i}$ is given by

$$\lambda_t^{N,i} = \lambda_t^N = f(X_{t-}^N),$$

where

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

Moreover we consider a function h of the form $h(t) = e^{-\alpha t}$ so that the process $(X_t^N)_t$ 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 random jump heights u , chosen according to the measure μ , 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/\sqrt{N} is given to all other neurons in the system. As a consequence, the process X^N has the following dynamic

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

Its infinitesimal generator is given by

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

for sufficiently smooth functions g . As N goes to infinity, the above expression converges to

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

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

$$d\bar{X}_t = -\alpha \bar{X}_t dt + \sigma \sqrt{f(\bar{X}_t)} dB_t. \quad (3)$$

It is classical to show in this framework that the convergence of generators implies the convergence of X^N to \bar{X} in distribution in Skorohod space. In this article we establish explicit bounds for the weak error for this convergence by means of a Trotter-Kato like formula. 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), (Dawson and Vaillancourt 1995) and (Kurtz and Xiong 1999) where all interacting components are subject to common noise. In our case, the common noise, that is, the Brownian motion B of (3), emerges in the limit as a consequence of the central limit theorem.

To obtain a precise control of the speed of convergence of X^N to \bar{X} we use analytical methods showing first the convergence of the generators from which we deduce the convergence of the semigroups via the formula

$$\bar{P}_t g(x) - P_t^N g(x) = \int_0^t P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x) ds. \quad (4)$$

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 (4) for extended generators is stated and proved in Appendix (Proposition 5.4).

It is well-known that under suitable assumptions on f , the solution of (3) 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 the limit of the law of X_t^N is π , for $(N, t) \rightarrow (\infty, \infty)$, under suitable conditions on the joint convergence of (N, t) . We also prove that there exists a parameter α^* such that for all $\alpha > \alpha^*$, this convergence holds whenever $(N, t) \rightarrow (\infty, \infty)$ jointly, without any further condition, and we provide a control of the error (Theorem 1.6).

The paper is organized as follows: in Section 1, we state the assumptions and formulate the main results. Section 2 is devoted to the proof of the convergence of the semigroup of X^N to that of \bar{X} (Theorem 1.4.(i)), and Section 3 to the study of the limit of the law of X_t^N as $N, t \rightarrow \infty$ (Theorem 1.6). In Section 4, we prove the convergence of the systems of point processes $(Z^{N,i})_{1 \leq i \leq N}$ to $(\bar{Z}^i)_{i \geq 1}$ (Theorem 1.7). Finally in Appendix, we prove some results on the extended generators, and some other technical results that we use throughout the paper.

1. Notation, assumptions and main results

1.1. Notation

The following notation are used throughout the paper:

- If X is a random variable, we note $\mathcal{L}(X)$ its distribution.
- If g is a real-valued function which is n times differentiable, we note $\|g\|_{n,\infty} = \sum_{k=0}^n \|g^{(k)}\|_\infty$.
- If $g : \mathbb{R} \rightarrow \mathbb{R}$ is a real-valued measurable function and π a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that g is integrable with respect to π , we write πg for $\int_{\mathbb{R}} g d\pi$.
- 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 for short $C_b(\mathbb{R})$ instead of $C_b^0(\mathbb{R})$. Finally, $C^n(\mathbb{R})$ denotes the set of n times continuously differentiable functions that are not necessarily bounded nor have bounded derivatives.
- 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)), and $D(\mathbb{R}_+, \mathbb{R}_+)$ for this space restricted to non-negative functions.
- $\mathcal{M}^\#$ denotes the space of locally finite measures on $\mathbb{R}_+ \times \mathbb{R}_+$ endowed with the topology of the weak convergence, and $\mathcal{N}^\#$ the subspace that contains only the simple point measures.
- α is a positive constant, L, σ and m_k ($1 \leq k \leq 4$) are fixed parameters defined in Assumptions 1, 2 and 3 below. Finally, we note C any arbitrary constant, so the value of C can change from line to line in an equation. Moreover, if C depends on some non-fixed parameter θ , we write C_θ .

1.2. Assumptions

Let X^N satisfy

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

where ν_0^N is a probability measure on \mathbb{R} . Under natural assumptions on f , the SDE (5) admits a unique non-exploding strong solution (see Proposition 5.6).

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

$$\begin{cases} d\bar{X}_t = -\alpha \bar{X}_t dt + \sigma \sqrt{f(\bar{X}_t)} dB_t, \\ \bar{X}_0 \sim \bar{\nu}_0, \end{cases} \quad (6)$$

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} .

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 .

Under Assumption 1, it is classical that the SDE (6) 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 for every $N \in \mathbb{N}^*$, $\int_{\mathbb{R}} x^4 d\nu_0^N(x) < \infty$.
- 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 2.1) and to prove the convergence of the generators of the processes $(X_t^N)_t$ (see Proposition 2.3).

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 .

Remark 1.1. By definition $m_1 = L$, since $m_1 := \|(\sqrt{f})'\|_{\infty}$ and L is the Lipschitz constant of \sqrt{f} .

Assumption 3 guarantees that the stochastic flow associated to (6) 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 2.4.

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

Assumption 4. X_0^N converges in distribution to \bar{X}_0 .

Obviously, Assumption 4 is a necessary condition for the convergence in distribution of X^N to \bar{X} .

1.3. Main results

Our first main result is the convergence of the process X^N to \bar{X} in distribution in Skorohod space, with an explicit rate of convergence for their semigroups. This rate of convergence will be expressed in terms of the following parameters

$$\beta := \max\left(\frac{1}{2}\sigma^2 L^2 - \alpha, 2\sigma^2 L^2 - 2\alpha, \frac{7}{2}\sigma^2 L^2 - 3\alpha\right) \quad (7)$$

and, for any $T > 0$ and any fixed $\varepsilon > 0$,

$$K_T := (1 + 1/\varepsilon) \int_0^T (1 + s^2) e^{\beta s} \left(1 + e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)(T-s)}\right) ds. \quad (8)$$

Remark 1.3. If $\alpha > 7/6 \sigma^2 L^2$, then $\beta < 0$, and one can choose $\varepsilon > 0$ such that $\sigma^2 L^2 - 2\alpha + \varepsilon < 0$, implying that $\sup_{T>0} K_T < \infty$.

Theorem 1.4. If Assumptions 1 and 2 hold, then the following assertions are true.

(i) Under Assumption 3, for all $T \geq 0$, for each $g \in C_b^3(\mathbb{R})$ and $x \in \mathbb{R}$,

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

In particular, if $\alpha > \frac{7}{6} \sigma^2 L^2$, then

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

(ii) If in addition Assumption 4 holds, then $(X^N)_N$ converges in distribution to \bar{X} in $D(\mathbb{R}_+, \mathbb{R})$.

We refer to Proposition 2.4 for the form of β given in (7). Theorem 1.4 is proved in the end of Subsection 2.2. (ii) is a consequence of Theorem IX.4.21 of (Jacod and Shiryaev 2003), using that X^N is a semimartingale. Alternatively, it can be proved as a consequence of (i), using that X^N is a Markov process.

Below we give some simulations of the trajectories of the process $(X_t^N)_{t \geq 0}$ in Figure 1.

Remark 1.5. As \bar{X} is (almost surely) continuous, Theorem 1.4.(ii), Skorohod's representation theorem (see Theorem 6.7 of (Billingsley 1999)) and the discussion in the subsection "The Skorohod Topology" of Section 12 of (Billingsley 1999) imply the convergence of X^N to \bar{X} in distribution in the topology of the uniform convergence on compact sets.

Under our assumptions, \bar{P} admits an invariant probability measure π , and we can even control the speed of convergence of $P_t^N g(x)$ to $\pi(g)$, as (N, t) goes to infinity, for suitable conditions on the joint convergence of N and t .

Theorem 1.6. Under Assumptions 1 and 2, \bar{X} is recurrent in the sense of Harris, having invariant probability measure $\pi(dx) = p(x)dx$ with density

$$p(x) = C \frac{1}{f(x)} \exp\left(-\frac{2\alpha}{\sigma^2} \int_0^x \frac{y}{f(y)} dy\right).$$

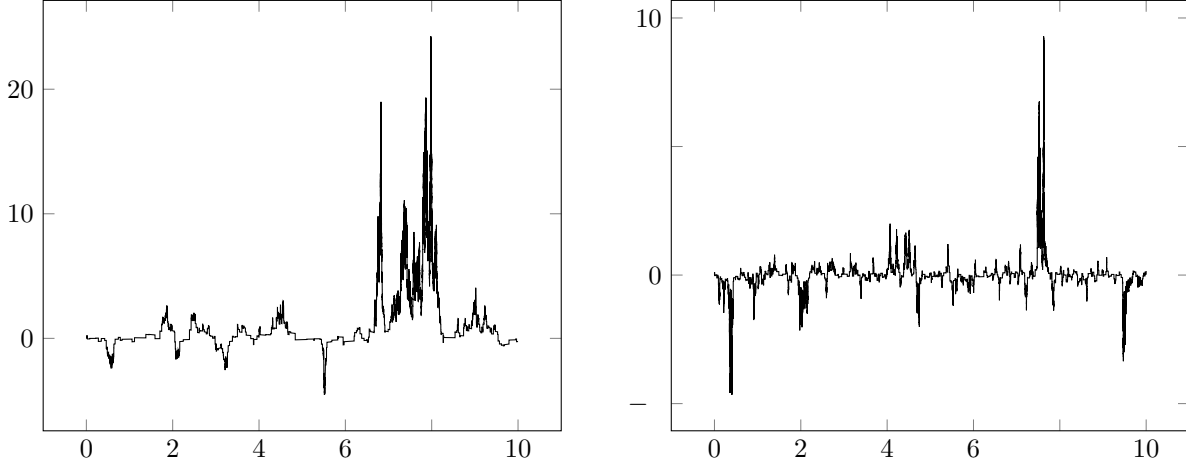


FIGURE 1. Simulation of trajectories of $(X_t^N)_{0 \leq t \leq 10}$ with $X_0^N = 0$, $\alpha = 1$, $\mu = \mathcal{N}(0, 1)$, $f(x) = 1 + x^2$, $N = 10$ (left picture) and $N = 50$ (right picture).

Besides, if Assumption 3 holds, then for all $g \in C_b^3(\mathbb{R})$ and $x \in \mathbb{R}$,

$$|P_t^N g(x) - \pi g| \leq C \|g\|_{3, \infty} (1 + x^2) \left(\frac{K_t}{\sqrt{N}} + e^{-\gamma t} \right),$$

where C and γ are positive constants independent of N and t , and where K_t is defined in (8). In particular, $P_t^N(x, \cdot)$ converges weakly to π as $(N, t) \rightarrow (\infty, \infty)$, provided $K_t = o(\sqrt{N})$.

If we assume, in addition, that $\alpha > \frac{7}{6} \sigma^2 L^2$, then $P_t^N(x, \cdot)$ converges weakly to π as $(N, t) \rightarrow (\infty, \infty)$ without any condition on the joint convergence of (t, N) , and we have, for any $g \in C_b^3(\mathbb{R})$ and $x \in \mathbb{R}$,

$$|P_t^N g(x) - \pi g| \leq C \|g\|_{3, \infty} (1 + x^2) \left(\frac{1}{\sqrt{N}} + e^{-\gamma t} \right).$$

Theorem 1.6 is proved in the end of Section 3.

Finally, using Theorem 1.4.(ii), we show the convergence of the point processes $Z^{N,i}$ defined in (2) 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 4.3), and it allows us to consider the joint distributions $(\bar{X}, \bar{\pi}_1, \dots, \bar{\pi}_k, \dots)$, where \bar{X} is defined as the solution of (6) 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

$$\bar{Z}_t^i = \int_{]0, t] \times \mathbb{R}_+ \times \mathbb{R}} \mathbb{1}_{\{z \leq f(\bar{X}_s)\}} d\pi_i(s, z, u). \quad (9)$$

Theorem 1.7. *Under Assumptions 1, 2 and 4, the sequence $(Z^{N,1}, Z^{N,2}, \dots, Z^{N,k}, \dots)_N$ converges to $(\bar{Z}^1, \bar{Z}^2, \dots, \bar{Z}^k, \dots)$ in distribution in $D(\mathbb{R}_+, \mathbb{R})^{\mathbb{N}^*}$ for the product topology.*

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 1.7 is proved in the end of Section 4.

Remark 1.8. *In Theorem 1.7, we implicitly define $Z^{N,i} := 0$ for each $i \geq N + 1$.*

2. Proof of Theorem 1.4

The goal of this section is to prove Theorem 1.4. To prove the convergence of the semigroups of $(X^N)_N$, we show in a first time the convergence of their generators. We start with useful a priori bounds on the moments of X^N and \bar{X} .

Lemma 2.1. *Under Assumptions 1 and 2, the following holds.*

- (i) *For all $\varepsilon > 0$, $t > 0$ and $x \in \mathbb{R}$, $\mathbb{E}_x [(X_t^N)^2] \leq C(1 + 1/\varepsilon)(1 + x^2)(1 + e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)t})$, for some $C > 0$ independent of N, t, x and ε .*
- (ii) *For all $\varepsilon > 0$, $t > 0$ and $x \in \mathbb{R}$, $\mathbb{E}_x [(\bar{X}_t)^2] \leq C(1 + 1/\varepsilon)(1 + x^2)(1 + e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)t})$, for some $C > 0$ independent of t, x and ε .*
- (iii) *For all $N \in \mathbb{N}^*, T > 0$, $\mathbb{E} [(\sup_{0 \leq t \leq T} |X_t^N|)^2] < +\infty$ and $\mathbb{E} [(\sup_{0 \leq t \leq T} |\bar{X}_t|)^2] < +\infty$.*
- (iv) *For all $T > 0, N \in \mathbb{N}^*$, $\sup_{0 \leq t \leq T} \mathbb{E}_x [(X_t^N)^4] \leq C_T(1 + x^4)$ and $\sup_{0 \leq t \leq T} \mathbb{E}_x [(\bar{X}_t)^4] \leq C_T(1 + x^4)$.*

We postpone the proof of Lemma 2.1 to Appendix.

2.1. 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 5.1). As this definition slightly differs from one reference to another, we define explicitly the extended generator in Definition 5.1 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} , and $\mathcal{D}'(A^N)$ and $\mathcal{D}'(\bar{A})$ their extended domains. 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 2.1.(i).

Lemma 2.2. *$C_b^2(\mathbb{R}) \subseteq \mathcal{D}'(\bar{A})$, 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 2.3. *If Assumptions 1 and 2 hold, then for all $g \in C_b^3(\mathbb{R})$,*

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

Proof. For $g \in C_b^3(\mathbb{R})$, if we note U a random variable having distribution μ , we have, since $\mathbb{E}[U] = 0$,

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

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

2.2. 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 (12) of Proposition 5.4, 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 norm $\|(\bar{P}_s g)'''\|_\infty$. This is done in the next proposition.

Proposition 2.4. *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 $x \mapsto \bar{P}_t g(x)$ is in $C_b^3(\mathbb{R})$ and satisfies*

$$\left\| (\bar{P}_t g)''' \right\|_\infty \leq C \|g\|_{3,\infty} (1 + t^2) e^{\beta t},$$

with $\beta = \max(\frac{1}{2}\sigma^2 L^2 - \alpha, 2\sigma^2 L^2 - 2\alpha, \frac{7}{2}\sigma^2 L^2 - 3\alpha)$.

The proof of Proposition 2.4 requires some detailed calculus to obtain the explicit expression for β , so we postpone it to Appendix.

Proof of Theorem 1.4. Step 1. The main part of the proof of the point (i) will be to show that Proposition 5.4 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 2.3, since

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

$$\begin{aligned}
&\leq \int_0^t \mathbb{E}_x^N [|\bar{A}(\bar{P}_s g)(X_{t-s}^N) - A^N(\bar{P}_s g)(X_{t-s}^N)|] ds \\
&\leq C \frac{1}{\sqrt{N}} \int_0^t \left\| (\bar{P}_s g)''' \right\|_{\infty} \mathbb{E}_x [f(X_{t-s}^N)] ds \\
&\leq C \frac{1}{\sqrt{N}} \|g\|_{3,\infty} \int_0^t \left((1+s^2)e^{\beta s} \left(1 + \mathbb{E}_x \left[(X_{t-s}^N)^2 \right] \right) \right) ds \\
&\leq C \left(1 + \frac{1}{\varepsilon} \right) \frac{1}{\sqrt{N}} \|g\|_{3,\infty} (1+x^2) \int_0^t (1+s^2)e^{\beta s} \left(1 + e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)(t-s)} \right) ds,
\end{aligned}$$

where we have used respectively Proposition 2.4 and Lemma 2.1.(i) to obtain the two last inequalities above, and ε is any positive constant.

Step 2. Now we show that X^N and \bar{X} satisfy the hypotheses of Proposition 5.4. To begin with we know that \bar{X} and X^N satisfy the hypotheses (i), (ii) and (iii), using Lemma 2.1. Then the hypothesis (iv) can be proved for the processes X^N and \bar{X} solving the SDEs (5) and (6) with straightforward calculations using Lemma 2.1. We know that \bar{P} satisfies hypothesis (v) thanks to Proposition 2.4.

Besides one can note that \bar{P} satisfies hypothesis (vi) using the calculations of the proof of Proposition 2.4. Then using Lemma 2.2, we see directly that \bar{A} and A^N satisfy the hypotheses (vii) and (ix). In addition (viii) is straightforward for \bar{A} , and it is a consequence of Lemma 2.5 below for A^N . The only remaining hypothesis (x) is a straightforward consequence of Lemma 2.6 below.

Step 3. We finally give the proof of the point (ii) of the theorem. With the notation of Theorem IX.4.21 of (Jacod and Shiryaev 2003), we have $K^N(x, dy) := Nf(x)\mu(\sqrt{N}dy)$, $b^N(x) = -\alpha x + \int K^N(x, dy)y = -\alpha x$, and $c^N(x) = \int K^N(x, dy)y^2 = \sigma^2 f(x)$. Then, an immediate adaptation of Theorem IX.4.21 of (Jacod and Shiryaev 2003) to our frame implies the result. \square

Lemma 2.5. For all $g \in C_c^2(\mathbb{R})$ such that $\text{Supp } g \subseteq [-M, M]$, we have

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

for some constant $C > 0$.

Proof. We have

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

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

$$\left| (A^N g)'(x) \right| \leq C \|g\|_{1,\infty} (1 + M^2) + \left| Nf'(x)\mathbb{E} \left[g \left(x + \frac{U}{\sqrt{N}} \right) \right] \right| + \left| Nf(x)\mathbb{E} \left[g' \left(x + \frac{U}{\sqrt{N}} \right) \right] \right|. \quad (10)$$

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:

$$\mathbb{E} \left[\left| g' \left(x + \frac{U}{\sqrt{N}} \right) \right| \right] \leq \|g'\|_{\infty} \mathbb{E} \left[\mathbf{1}_{\{|x+U/\sqrt{N}| \leq M\}} \right]$$

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

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

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

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

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

This concludes the proof. \square

Lemma 2.6. *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. Proof of Theorem 1.6

In this section, we prove Theorem 1.6. We begin by proving some properties of the invariant measure of \bar{P}_t . In what follows we use the total variation distance between two probability measures ν_1 and ν_2 defined by

$$\|\nu_1 - \nu_2\|_{TV} = \sup_{g: \|g\|_\infty \leq 1} |\nu_1(g) - \nu_2(g)|.$$

Proposition 3.1. *If Assumptions 1 and 2 hold, then the invariant measure π of $(\bar{P}_t)_t$ exists and is unique. Its density is given, up to multiplication with a constant, by*

$$p(x) = C \frac{1}{f(x)} \exp \left(-\frac{2\alpha}{\sigma^2} \int_0^x \frac{y}{f(y)} dy \right).$$

In addition, if Assumption 3 holds, then for every $0 < q < 1/2$, there exists some $\gamma > 0$ such that, for all $t \geq 0$,

$$\|\bar{P}_t(x, \cdot) - \pi\|_{TV} \leq C (1+x^2)^q e^{-\gamma t}.$$

Proof. In a first time, let us prove the positive Harris recurrence of \bar{X} implying the existence and uniqueness of π . According to Example 3.10 of (Khasminskii 2012) it is sufficient to show that $S(x) := \int_0^x s(y)dy$ goes to $+\infty$ (resp. $-\infty$) as x goes to $+\infty$ (resp. $-\infty$), where

$$s(x) := \exp\left(\frac{2\alpha}{\sigma^2} \int_0^y \frac{v}{f(v)} dv\right).$$

For $x > 0$,

$$s(x) \geq \exp\left(C \int_0^y \frac{2v}{1+v^2} dv\right) = \exp(C \ln(1+y^2)) = (1+y^2)^C \geq 1,$$

implying that $S(x)$ goes to $+\infty$ as x goes to $+\infty$. With the same reasoning, we obtain that $S(x)$ goes to $-\infty$ as x goes to $-\infty$. Finally, the associated invariant density is given, up to a constant, by

$$p(x) = \frac{C}{f(x)s(x)}.$$

For the second part of the proof, take $V(x) = (1+x^2)^q$, for some $q < 1/2$, then

$$V'(x) = 2qx(1+x^2)^{q-1}, V''(x) = 2q(1+x^2)^{q-2}[2x^2(q-1) + (1+x^2)].$$

As $q < \frac{1}{2}$, $V''(x) < 0$ for x^2 sufficiently large, say, for $|x| \geq K$. In this case, for $|x| \geq K$,

$$\bar{A}V(x) \leq -2\alpha qx^2(1+x^2)^{q-1} \leq -2\alpha q \frac{x^2}{1+x^2} V(x) \leq -2\alpha q \frac{K^2}{1+K^2} V(x) = -cV(x).$$

So we obtain all in all for suitable constants c and d that, for any $x \in \mathbb{R}$,

$$\bar{A}V(x) \leq -cV(x) + d. \quad (11)$$

Obviously, the sampled chain $(\bar{X}_{kT})_{k \geq 0}$ is Feller and π -irreducible. The support of π being \mathbb{R} , Theorem 3.4 of (Meyn and Tweedie 1992) implies that every compact set is petite for the sampled chain. Then, as (11) implies the condition (CD3) of Theorem 6.1 of (Meyn and Tweedie 1993), we have the following bound: introducing for any probability measure μ the weighted norm

$$\|\mu\|_V := \sup_{g: |g| \leq 1+V} |\mu(g)|,$$

there exist $C, \gamma > 0$ such that

$$\|\bar{P}_t(x, \cdot) - \pi\|_V \leq C(1+V(x))e^{-\gamma t}.$$

This implies the result, since $\|\cdot\|_{TV} \leq \|\cdot\|_V$. \square

Now the proof of Theorem 1.6 is straightforward.

Proof of Theorem 1.6. The first part of the theorem has been proved in Proposition 3.1. For the second part, for any $g \in C_b^3(\mathbb{R})$,

$$\begin{aligned} |P_t^N g(x) - \pi g| &\leq |P_t^N g(x) - \bar{P}_t g(x)| + |\bar{P}_t g(x) - \pi g| \\ &\leq \frac{K_t}{\sqrt{N}}(1+x^2)\|g\|_{3,\infty} + \|g\|_\infty \|\bar{P}_t(x, \cdot) - \pi\|_{TV} \\ &\leq \|g\|_{3,\infty} C \left(\frac{K_t}{\sqrt{N}}(1+x^2) + e^{-\gamma t}(1+x^2)^q \right), \end{aligned}$$

where we have used Theorem 1.4 and Proposition 3.1. \square

4. Proof of Theorem 1.7

In this section we prove Theorem 1.7, that is the convergence in distribution of $(Z^{N,1}, \dots, Z^{N,k}, \dots)$ to $(\bar{Z}^1, \dots, \bar{Z}^k, \dots)$ in $D(\mathbb{R}_+, \mathbb{R})^{\mathbb{N}^*}$. This convergence is a consequence of the following theorem, that will be proved in Appendix.

Theorem 4.1. *Let $(Y^N)_{N \in \mathbb{N}^*}$ be a sequence of $D(\mathbb{R}_+, \mathbb{R}_+)$ -valued random variables, and \bar{Y} be a $D(\mathbb{R}_+, \mathbb{R}_+)$ -valued random variable. Let $(\pi^k)_{k \in \mathbb{N}^*}$ and $(\bar{\pi}^k)_{k \in \mathbb{N}^*}$ be i.i.d. families of Poisson measures on $\mathbb{R}_+ \times \mathbb{R}_+$ having Lebesgue intensity. Let $Z^{N,k}$ and \bar{Z}^k be the point processes defined as follows*

$$\begin{aligned} Z_t^{N,k} &:= \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{z \leq Y_{s-}^N\}} d\pi^k(s, z), \quad 1 \leq k \leq N, \\ Z_t^{N,k} &:= 0, \quad k \geq N + 1, \\ \bar{Z}_t^k &:= \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{z \leq \bar{Y}_{s-}\}} d\bar{\pi}^k(s, z), \quad k \geq 1. \end{aligned}$$

Assume that \bar{Y} is independent of $(\bar{\pi}^k)_{k \in \mathbb{N}^*}$, and that Y^N converges to \bar{Y} in distribution in Skorohod topology. Then, as N goes to infinity, $(Z^{N,k})_{k \in \mathbb{N}^*}$ converges to $(\bar{Z}^k)_{k \in \mathbb{N}^*}$ in distribution in $D(\mathbb{R}_+, \mathbb{R})^{\mathbb{N}^*}$ endowed with the product topology.

In order to apply Theorem 4.1, we need to prove two properties: the convergence of $(f(X_t^N))_{t \geq 0}$ to $(f(\bar{X}_t))_{t \geq 0}$ in Skorohod topology, and the independence between \bar{X} and the Poisson measures. The first property is a mere consequence of Theorem 1.4.(ii) and the following lemma, that is proved in Appendix.

Lemma 4.2. *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.*

Besides, according to our definition of \bar{Z}^i in (9), it is obvious that \bar{X} is independent of $(\bar{\pi}_j)_{j \in \mathbb{N}^*}$. The goal of Proposition 4.3 is to justify the way we introduced \bar{Z}^i .

Proposition 4.3. *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 $\mathcal{D}(\bar{X}) \otimes \mathcal{D}(\pi_1) \otimes \dots \otimes \mathcal{D}(\pi_k)$ in $D(\mathbb{R}_+, \mathbb{R}) \times (\mathcal{M}^\#)^k$, where $\mathcal{M}^\#$ is the space of locally finite measures on $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}$ endowed with the topology of the weak convergence.*

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.

Finally, we can apply directly Theorem 4.1 to prove Theorem 1.7.

5. Appendix

5.1. Extended generators

In this subsection, we define precisely the notion of generators we use and we prove the results needed to prove formula (4). 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

$(C_b(\mathbb{R}), \|\cdot\|_\infty)$. In this context, the generator A of a semigroup $(P_t)_t$ is defined on the set of functions $\mathcal{D}(A) = \{g \in C_b(\mathbb{R}) : \exists h \in C_b(\mathbb{R}), \|\frac{1}{t}(P_t g - g) - h\|_\infty \rightarrow 0 \text{ as } t \rightarrow 0\}$. Then one denotes the previous function h as Ag . In general, we can only guarantee that $\mathcal{D}(A)$ contains the functions that have a compact support, but to prove Proposition 5.4, 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 $C_c^\infty(\mathbb{R})$.

This 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, verifying the fundamental martingale property, which 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 5.1. *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 measurable functions g such that $\mathbb{E}_x [|g(X_t)|]$ is finite for every $x \in \mathbb{R}$. Then we define $\mathcal{D}'(A)$ to be the set of measurable functions g such that for each $x \in \mathbb{R}$, $\frac{1}{t}(P_t g(x) - g(x))$ converges as t goes to 0 to some limit that we note $Ag(x)$ and such that:*

- for all $t \geq 0$, for all x , $\int_0^t |Ag(X_s)| ds$ is \mathbb{P}_x -almost surely defined and \mathbb{P}_x -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 .

Remark 5.2. *One can note that if we assume the function $t \mapsto P_t Ag(x)$ to be continuous and that $\int_0^t \mathbb{E}_x [|Ag(X_s)|] ds$ is finite for any $x \in \mathbb{R}$, then the martingale property of Definition 5.1 implies that $Ag(x)$ is the derivative of $t \mapsto P_t g(x)$ at $t = 0$.*

In our Definition 5.1, contrarily e.g. to (Meyn and Tweedie 1993), we impose this additional condition of differentiability. We do this because this is the main property we need in our applications. In practice, this differentiability condition is satisfied for C^2 functions g and for "common" Markov processes X with finite $g^{(k)}$ -moments ($0 \leq k \leq 2$).

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. The difference is that here we have to replace the uniform convergence by point-wise convergence, hence we need boundedness assumptions on the Markov process.

Lemma 5.3. *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 $t \geq 0$, $\mathbb{E}_x [\sup_{0 \leq s \leq t} |P_s Ag(X_t)|]$ 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 $|P_t Ag(y) - Ag(y)| \leq CM(y)\varepsilon(t)$ for some constant C that is allowed to 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:

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

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 $\sup_{0 \leq s \leq t} |P_sAg(X_t)| + |Ag(X_t)|$ (using the fact that $P_hg(y) - g(y) = \int_0^h P_sAg(y)ds$ which is a consequence of the last point of Definition 5.1), we know that this expectation vanishes as h goes to 0 by dominated convergence. This means exactly that $\frac{d^+}{dt}(P_tg(x))$ exists and is $P_tAg(x)$.

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

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 $C(\sup_{0 \leq s \leq h} \varepsilon(s))(\sup_{0 \leq r \leq t} \mathbb{E}_x [M(X_r)])$. \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 5.4. *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 $T > 0$, $\sup_{0 \leq t \leq T} \mathbb{E}_x [(Y_t^N)^4] \leq C_T(1 + x^4)$ and $\sup_{0 \leq t \leq T} \mathbb{E}_x^N [(Y_t^N)^4] \leq C_T(1 + x^4)$ for some constant $C_T > 0$ not depending on N .*
- (ii) *for all $T > 0$, $\mathbb{E} [(\sup_{0 \leq t \leq T} |Y_t^N|)^2] < +\infty$.*
- (iii) *for all $0 \leq s, t \leq T$ and $x \in \mathbb{R}$,*

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

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

- (iv) *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 $Q_T > 0$.*

- (v) 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.
- (vi) $C_b^3(\mathbb{R}) \subseteq \mathcal{D}'(A^N) \cap \mathcal{D}'(\bar{A})$. For all $g \in C_b^3(\mathbb{R})$ and $x \in \mathbb{R}$, $|Ag(x)| \leq C\|g\|_{2,\infty}(1+x^2)$ and $|A^N g(x)| \leq C\|g\|_{2,\infty}(1+x^2)$.
- (vii) for all $g \in C_c^3(\mathbb{R})$ such that $\text{Supp } g \subseteq [-M, M]$, $\|(\bar{A}g)'\|_\infty \leq C\|g\|_{3,\infty}(1+M^2)$ and $\|(A^N g)'\|_\infty \leq C\|g\|_{3,\infty}(1+M^2)$.
- (viii) there exists some $C > 0$ such that for all $x, y \in \mathbb{R}$, for all $g \in C_b^3(\mathbb{R})$, $|\bar{A}g(x) - \bar{A}g(y)| \leq C(1+x^2+y^2)|x-y|$ and $|A^N g(x) - A^N g(y)| \leq C(1+x^2+y^2)|x-y|$.
- (ix) 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 $t \in \mathbb{R}_+$:

$$(\bar{P}_t - P_t^N)g(x) = \int_0^t P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x) ds. \quad (12)$$

Remark 5.5. Notice that the conditions of Proposition 5.4 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. To begin with, let us emphasize the fact that hypothesis (i) implies

$$\sup_{0 \leq t \leq T} \mathbb{E}_x [(\bar{Y}_t)^2] \leq C_T(1+x^2) \text{ and } \sup_{0 \leq t \leq T} \mathbb{E}_x^N [(Y_t^N)^2] \leq C_T(1+x^2), \quad (13)$$

since

$$\mathbb{E}_x [(\bar{Y}_t)^2] \leq \mathbb{E}_x [(\bar{Y}_t)^4]^{1/2} \leq C_T \sqrt{1+x^4} \leq C_T(1+x^2).$$

We fix $t \geq 0$, $N \in \mathbb{N}^*$, $g \in C_b^3(\mathbb{R})$, $x \in \mathbb{R}$ in the rest of 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 5.3. The condition of the point (1) of the lemma is a straightforward consequence of (13) and hypothesis (vi), and the conditions of the point (2) are satisfied for $M(x) = \sqrt{1+x^4}$ using (13) and hypothesis (i), (iii) and (viii). As a consequence, and thanks to hypothesis (iv), 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 5.3 (see hypothesis (iv) and (vi)).

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

$$\begin{aligned} |P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x) - P_{t-s_k}^N (\bar{A} - A^N) \bar{P}_{s_k} g(x)| &\leq |(P_{t-s}^N - P_{t-s_k}^N) (\bar{A} - A^N) g_s(x)| \quad (14) \\ &+ |P_{t-s_k}^N (\bar{A} - A^N) (\bar{P}_s - \bar{P}_{s_k}) g(x)|, \quad (15) \end{aligned}$$

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

To show that the term (14) 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 (14) 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 (iii), (iv) and (vii), we have

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

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 (ix)), and then we can apply dominated convergence (using hypothesis (ii) and (vi)).

We just proved that the term (14) vanishes. To finish the proof, we need to show that the term (15) vanishes. We note that the term (15) 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 (ii), (iv), (v) and (ix)). Dominated convergence, using (13) and hypothesis (ii), (iv) and (vi), then implies the result. \square

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

Proposition 5.6. *If Assumptions 1 and 2 hold, the equation (5) admits a unique non-exploding strong solution.*

Proof. It is well known that if f is bounded, there is a unique strong solution of (5) (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 (5) 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 . With a classical reasoning relying on Ito's formula and Grönwall's lemma, we can prove that

$$\mathbb{E} \left[\left(X_{t \wedge \tau_K^N}^{N,K} \right)^2 \right] \leq C_t (1 + x^2), \quad (16)$$

where $C_t > 0$ does not depend on K . As a consequence, we know that almost surely, $\tau^N = +\infty$. 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 X^N satisfies equation (5). 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 (5) with f replaced by f_K , we know that $X^N(\omega)$ verifies the equation (5) on $[0, t]$. This holds for all $t > 0$. As a consequence, we know that X^N satisfies the equation (5). This proves the existence of strong solution. The uniqueness is a consequence of the uniqueness of strong solutions of (5), if we replace f by f_K in (5), 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

5.3. Proof of Lemma 2.1

We begin with the proof of (i). Let $\Phi(x) = x^2$ and A^N be the extended generator of $(X_t^N)_{t \geq 0}$. One can note that, applying Fatou's lemma to the inequality (16), one obtains for all $t \geq 0$, $\mathbb{E} [(X_t^N)^2]$ is finite. As a consequence $\Phi \in \mathcal{D}'(A^N)$ (in the sense of Definition 5.1), and for all $x \in \mathbb{R}$,

$$\begin{aligned} A^N \Phi(x) &= -2\alpha\Phi(x) + \sigma^2 f(x) \leq -2\alpha\Phi(x) + \sigma^2 \left(L|x| + \sqrt{f(0)} \right)^2 \\ &\leq (\sigma^2 L^2 - 2\alpha)\Phi(x) + 2\sigma^2 L|x|\sqrt{f(0)} + \sigma^2 f(0). \end{aligned}$$

Let $\varepsilon > 0$ be fixed, and $\eta_\varepsilon = 2\sigma^2 L \sqrt{f(0)} / \varepsilon$. Using that, for every $x \in \mathbb{R}$, $|x| \leq x^2 / \eta_\varepsilon + \eta_\varepsilon$, we have

$$A^N \Phi(x) \leq c_\varepsilon \Phi(x) + d_\varepsilon, \quad (17)$$

with $c_\varepsilon = \sigma^2 L^2 - 2\alpha + \varepsilon$ and $d_\varepsilon = O(1/\varepsilon)$. Let us assume that $c_\varepsilon \neq 0$, possibly by reducing $\varepsilon > 0$. Considering $Y_t^N := e^{-c_\varepsilon t} \Phi(X_t^N)$, by Ito's formula,

$$\begin{aligned} dY_t^N &= -c_\varepsilon e^{-c_\varepsilon t} \Phi(X_t^N) dt + e^{-c_\varepsilon t} d\Phi(X_t^N) \\ &= -c_\varepsilon e^{-c_\varepsilon t} \Phi(X_t^N) dt + e^{-c_\varepsilon t} A^N \Phi(X_t^N) dt + e^{-c_\varepsilon t} dM_t, \end{aligned}$$

where $(M_t)_{t \geq 0}$ is a \mathbb{P}_x -martingale. Using (17), we obtain

$$dY_t^N \leq d_\varepsilon e^{-c_\varepsilon t} dt + e^{-c_\varepsilon t} dM_t,$$

implying

$$\mathbb{E}_x [Y_t^N] \leq \mathbb{E}_x [Y_0^N] + \frac{d_\varepsilon}{c_\varepsilon} e^{-c_\varepsilon t} (1 - e^{-c_\varepsilon t}).$$

One deduces

$$\mathbb{E}_x \left[(X_t^N)^2 \right] \leq x^2 e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)t} + \frac{C}{\varepsilon} \left(e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)t} + 1 \right), \quad (18)$$

for some constant $C > 0$ independent of t, ε, N .

The proof of (ii) is analogous and therefore omitted.

Now we prove (iii). From

$$X_t^N = X_0^N - \alpha \int_0^t X_s^N ds + \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{s-}^N)\}} d\pi_j(s, z, u),$$

we deduce

$$\begin{aligned} \left(\sup_{0 \leq s \leq t} |X_s^N| \right)^2 &\leq 3(X_0^N)^2 + 3\alpha^2 t \int_0^t (X_s^N)^2 ds \\ &\quad + 3 \sum_{j=1}^N \left(\sup_{0 \leq s \leq t} \left| \int_{]0,s] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_j(r, z, u) \right| \right)^2. \end{aligned} \quad (19)$$

Applying Burkholder-Davis-Gundy inequality to the last term above in (19), we can bound its expectation by

$$\begin{aligned} 3N \mathbb{E} \left[\int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} u^2 \mathbf{1}_{\{z \leq f(X_{s-}^N)\}} d\pi_j(s, z, u) \right] &\leq 3N \sigma^2 \int_0^t \mathbb{E} [f(X_{s-}^N)] ds \\ &\leq 3N \sigma^2 C \int_0^t (1 + \mathbb{E} [(X_s^N)^2]) ds. \end{aligned} \quad (20)$$

Now, bounding (19) by (20), and using point (i) of the lemma we conclude the proof of (iii). Finally, (iv) can be proved in classical way, applying Ito's formula and Grönwall's lemma.

5.4. Proof of Proposition 2.4

To begin with, we use Theorem 1.4.1 of (Kunita 1986) to prove that the flow associated to the SDE (6) 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):

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

In the following, we consider the process $(\bar{X}_t^{(x)})_t$, solution of the SDE (6) and satisfying $\bar{X}_0^{(x)} = x$. 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

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

$$\begin{aligned}
(\bar{P}_t g)''(x) &= \mathbb{E} \left[\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} g'(\bar{X}_t^{(x)}) + \left(\frac{\partial \bar{X}_t^{(x)}}{\partial x} \right)^2 g''(\bar{X}_t^{(x)}) \right], \\
(\bar{P}_t g)'''(x) &= \mathbb{E} \left[\frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3} g'(\bar{X}_t^{(x)}) + 3 \frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} \cdot \frac{\partial \bar{X}_t^{(x)}}{\partial x} g''(\bar{X}_t^{(x)}) + \left(\frac{\partial \bar{X}_t^{(x)}}{\partial x} \right)^3 g'''(\bar{X}_t^{(x)}) \right]. \quad (21)
\end{aligned}$$

We start with the representation

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

This implies

$$\frac{\partial \bar{X}_t^{(x)}}{\partial x} = e^{-\alpha t} + \sigma \int_0^t e^{-\alpha(t-s)} \frac{\partial \bar{X}_s^{(x)}}{\partial x} (\sqrt{f})'(\bar{X}_s^{(x)}) dB_s. \quad (22)$$

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

$$M_t = \int_0^t \sigma (\sqrt{f})'(\bar{X}_s^{(x)}) dB_s, \quad (23)$$

we obtain $U_t = 1 + \int_0^t U_s dM_s$, whence

$$U_t = \exp \left(M_t - \frac{1}{2} \langle M \rangle_t \right).$$

Notice that this implies $U_t > 0$ almost surely, whence $\frac{\partial \bar{X}_t^{(x)}}{\partial x} > 0$ almost surely. We deduce from this that

$$U_t^p = e^{pM_t - \frac{p}{2} \langle M \rangle_t} = \exp \left(pM_t - \frac{1}{2} p^2 \langle M \rangle_t \right) e^{\frac{1}{2} p(p-1) \langle M \rangle_t} = \mathcal{E}(M)_t e^{\frac{1}{2} p(p-1) \langle M \rangle_t}.$$

Since $(\sqrt{f})'$ is bounded, M_t is a martingale, thus $\mathcal{E}(M)$ is an exponential martingale with expectation 1, implying that

$$\mathbb{E} U_t^p \leq e^{\frac{1}{2} p(p-1) \sigma^2 m_1^2 t}, \quad (24)$$

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

$$\mathbb{E} \left[\left(\frac{\partial \bar{X}_t^{(x)}}{\partial x} \right)^2 \right] \leq e^{(\sigma^2 m_1^2 - 2\alpha)t} \quad \text{and} \quad \mathbb{E} \left[\left| \frac{\partial \bar{X}_t^{(x)}}{\partial x} \right|^3 \right] \leq e^{(3\sigma^2 m_1^2 - 3\alpha)t}. \quad (25)$$

Differentiating (22) with respect to x , we obtain

$$\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} = \sigma \int_0^t e^{-\alpha(t-s)} \left[\frac{\partial^2 \bar{X}_s^{(x)}}{\partial x^2} (\sqrt{f})'(\bar{X}_s^{(x)}) + \left(\frac{\partial \bar{X}_s^{(x)}}{\partial x} \right)^2 (\sqrt{f})^{(2)}(\bar{X}_s^{(x)}) \right] dB_s. \quad (26)$$

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

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

which can be rewritten as

$$dV_t = V_t dM_t + Y_t dB_t, V_0 = 0, Y_t = \sigma e^{-\alpha t} U_t^2 \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_t^{(x)}\right),$$

with M_t as in (23). Applying Ito's formula to $Z_t := V_t/U_t$ (recall that $U_t > 0$), we obtain

$$dZ_t = \frac{Y_t}{U_t} dB_t - \frac{Y_t}{U_t} d \langle M, B \rangle_t,$$

such that, by the precise form of Y_t and since $Z_0 = 0$,

$$Z_t = \sigma \int_0^t e^{-\alpha s} U_s \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right) dB_s - \sigma^2 \int_0^t e^{-\alpha s} U_s \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right) \left(\sqrt{f}\right)' \left(\bar{X}_s^{(x)}\right) ds.$$

Using Jensen's inequality, (24) and Burkholder-Davis-Gundy inequality, for all $t \geq 0$,

$$\begin{aligned} \mathbb{E} [Z_t^4] &\leq C \left(\mathbb{E} \left[\left(\int_0^t e^{-\alpha s} U_s \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right) dB_s \right)^4 \right] \right. \\ &\quad \left. + \mathbb{E} \left[\left(\int_0^t e^{-\alpha s} U_s \left(\sqrt{f}\right)' \left(\bar{X}_s^{(x)}\right) \left(\sqrt{f}\right)^{(2)} \left(\bar{X}_s^{(x)}\right) ds \right)^4 \right] \right) \\ &\leq C (t + t^3) \int_0^t e^{(6\sigma^2 m_1^2 - 4\alpha)s} ds \\ &\leq C (t + t^3) \left(1 + t + e^{(6\sigma^2 m_1^2 - 4\alpha)t} \right). \end{aligned} \quad (27)$$

We deduce that

$$\mathbb{E} [V_t^2] \leq \mathbb{E} [Z_t^4]^{1/2} \mathbb{E} [U_t^4]^{1/2} \leq C(t^{1/2} + t^2) e^{3\sigma^2 m_1^2 t},$$

whence

$$\mathbb{E} \left[\left(\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} \right)^2 \right] \leq C(t^{1/2} + t^2) e^{(3\sigma^2 m_1^2 - 2\alpha)t}. \quad (28)$$

Finally, differentiating (26), we get

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

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

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

Once again we can rewrite this as

$$dW_t = W_t dM_t + Y_t' dB_t, W_0 = 0,$$

where

$$Y'_t = \sigma \left(3e^{-\alpha t} U_t V_t (\sqrt{f})^{(2)} (\bar{X}_t^{(x)}) + e^{-2\alpha t} U_t^3 (\sqrt{f})^{(3)} (\bar{X}_t^{(x)}) \right),$$

whence, introducing $Z'_t = \frac{W_t}{U_t}$,

$$Z'_t = \int_0^t \frac{Y'_s}{U_s} dB_s - \int_0^t \frac{Y'_s}{U_s} d \langle M, B \rangle_s.$$

As previously, we obtain,

$$\begin{aligned} \mathbb{E} \left[(Z'_t)^2 \right] &\leq C(1+t) \int_0^t \mathbb{E} \left[\left(\frac{Y'_s}{U_s} \right)^2 \right] ds \\ &\leq C(1+t) \int_0^t (e^{-2\alpha s} \mathbb{E} [V_s^2] + e^{-4\alpha s} \mathbb{E} [U_s^4]) ds \\ &\leq C(1+t) \int_0^t ((s^{1/2} + s^2) e^{(3\sigma^2 m_1^2 - 2\alpha)s} + e^{(6\sigma^2 m_1^2 - 4\alpha)s}) ds \\ &\leq C(1+t^3) \int_0^t (e^{(3\sigma^2 m_1^2 - 2\alpha)s} + e^{(6\sigma^2 m_1^2 - 4\alpha)s}) ds \\ &\leq C(1+t^3) \int_0^t (1 + e^{(6\sigma^2 m_1^2 - 4\alpha)s}) ds \leq C(1+t^4) (1 + e^{(6\sigma^2 m_1^2 - 4\alpha)t}). \end{aligned} \quad (29)$$

As a consequence,

$$\begin{aligned} \mathbb{E} [|W_t|] &\leq \mathbb{E} [(Z'_t)^2]^{1/2} \mathbb{E} [U_t^2]^{1/2} \leq C(1+t^2) (1 + e^{(3\sigma^2 m_1^2 - 2\alpha)t}) e^{\frac{1}{2}\sigma^2 m_1^2 t} \\ &\leq C(1+t^2) (e^{\frac{1}{2}\sigma^2 m_1^2 t} + e^{(\frac{7}{2}\sigma^2 m_1^2 - 2\alpha)t}), \end{aligned}$$

implying

$$\mathbb{E} \left[\left| \frac{\partial^3 \bar{X}_t^{(x)}}{\partial^3 x} \right| \right] \leq C(1+t^2) (e^{(\frac{1}{2}\sigma^2 m_1^2 - \alpha)t} + e^{(\frac{7}{2}\sigma^2 m_1^2 - 3\alpha)t}). \quad (30)$$

Finally, using Cauchy-Schwarz inequality, and inserting (25), (28) and (30) in (21),

$$\left\| (\bar{P}_t g)''' \right\|_\infty \leq C \|g\|_{3,\infty} (1+t^2) (e^{(\frac{1}{2}\sigma^2 m_1^2 - \alpha)t} + e^{2(\sigma^2 m_1^2 - \alpha)t} + e^{(\frac{7}{2}\sigma^2 m_1^2 - 3\alpha)t}),$$

which proves the proposition.

5.5. Proof of Theorem 4.1

One of the key argument to prove Theorem 4.1 is the following result. Let us recall that $\mathcal{M}^\#$ denotes the space of locally finite measures on $\mathbb{R}_+ \times \mathbb{R}_+$, and $\mathcal{N}^\#$ the subspace of locally finite simple point measures.

Theorem 5.7. Let $\Phi : D(\mathbb{R}_+, \mathbb{R}_+) \times \mathcal{N}^\# \rightarrow D(\mathbb{R}_+, \mathbb{R}_+)$ be defined as

$$\Phi(x, \pi)_t := \int_{[0, t] \times \mathbb{R}_+} \mathbf{1}_{\{z \leq x_{s-}\}} d\pi(s, z).$$

Let $(x, \pi) \in D(\mathbb{R}_+, \mathbb{R}_+) \times \mathcal{N}^\#$ satisfy the following conditions:

- for each $t \geq 0$ such that $\pi(\{t\} \times \mathbb{R}_+) = 0$, x is continuous on t ,
- for every $t \geq 0$, $\pi(\{t\} \times \mathbb{R}_+) \leq 1$,
- $\pi(\{(t, x_{t-}) : t \geq 0\}) = 0$.

Then, Φ is continuous at the point (x, π) , where $D(\mathbb{R}_+, \mathbb{R}_+)$ is endowed with Skorohod topology, and $\mathcal{N}^\#$ is endowed with the topology of the weak convergence.

Before proving Theorem 5.7, we point out that Φ is not continuous on every point of $D(\mathbb{R}_+, \mathbb{R}_+) \times \mathcal{N}^\#$ (see Example 5.8).

Example 5.8. Let us consider the point measure $\pi = \delta_{(1,1)}$ 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.

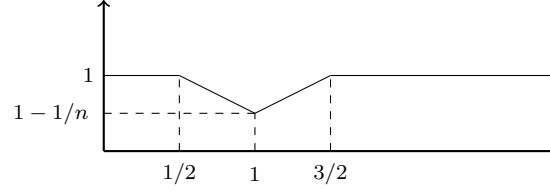


FIGURE 2. Graph of x^n

To prove Theorem 5.7, we need the following lemmas. We omit their proofs.

Lemma 5.9. 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 $(t_N)_N$ be a sequence that converges to some $t > 0$. If x is continuous on t , then $x_N(t_N-) \rightarrow x(t)$.

Lemma 5.10. Let $T > 0$, increasing sequences $0 = t_0 < t_1 < \dots < t_{n-1} < t_n = T$, $0 = t_0^N < t_1^N < \dots < t_{n^N-1}^N < t_{n^N}^N = T$ ($N \in \mathbb{N}^*$), and define the functions $g, g_N \in D([0, T], \mathbb{R})$ by

$$\begin{cases} g(t) = \sum_{j=0}^{n-1} \mathbf{1}_{[t_j, t_{j+1}[}(t)j & t \in [0, T[, \\ g(T) = n - 1, \end{cases}$$

$$\begin{cases} g_N(t) = \sum_{j=0}^{n^N-1} \mathbf{1}_{[t_j^N, t_{j+1}^N[}(t)j & t \in [0, T[, \\ g_N(T) = n^N - 1. \end{cases}$$

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 $g(t)$. Then g_N converges to g in $D([0, T], \mathbb{R})$.

Proof of Theorem 5.7. Let us consider a sequence $(x^N, \pi^N)_N$ that converges to (x, π) . Let $Z := \Phi(x, \pi)$ and $Z^N := \Phi(x^N, \pi^N)$.

We fix $t \geq 0$ such that $\pi(\{t\} \times \mathbb{R}_+) = 0$ and for all $N \in \mathbb{N}^*$, $\pi^N(\{t\} \times \mathbb{R}_+) = 0$. In particular t is a point of continuity of Z and of each Z^N . We consider some $T > 0$ such that $T > \max(t, \|x\|_{\infty, [0, t]}, \sup_N \|x^N\|_{\infty, [0, t]})$ and such that $\pi(\{T\} \times [0, T] \cup [0, T] \times \{T\}) = 0$. Let us consider $n := \pi([0, T]^2)$ and $n^N := \pi^N([0, T]^2)$. In the rest of the proof, we identify each simple point measure with its set of atoms. We write $\pi \cap [0, T]^2 = \{(\tau_j, \zeta_j) : 1 \leq j \leq n\}$ and $\pi^N \cap [0, T]^2 = \{(\tau_j^N, \zeta_j^N) : 1 \leq j \leq n^N\}$, where the pairs are lexicographically ordered.

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

Now we show that for all $1 \leq j \leq n$, τ_j^N and ζ_j^N converge respectively to τ_j and ζ_j . The idea of the proof consists in defining disjoint sets U_j^ε of radius ε that contain each (τ_j, ζ_j) and to use Proposition A2.6.II.(iv) of (Daley and Vere-Jones 2003) again to show that (τ_j^N, ζ_j^N) is necessarily in U_j^ε for all ε . We fix some $\varepsilon > 0$ and we consider $\gamma = \min_{1 \leq j \leq n} (\tau_{j+1} - \tau_j) > 0$. We can choose $0 < \eta < \varepsilon \wedge \gamma/3$ such that for all $1 \leq j_1, j_2 \leq n$, if $j_1 \neq j_2$ then $B((\tau_{j_1}, \zeta_{j_1}), \eta) \cap B((\tau_{j_2}, \zeta_{j_2}), \eta) = \emptyset$ (where we endow \mathbb{R}_+^2 with $\|\cdot\|_\infty$). Then we know that for all $1 \leq j \leq n$, $|\pi^N \cap B((\tau_j, \zeta_j), \eta)|$ converges to $|\pi \cap B((\tau_j, \zeta_j), \eta)| = 1$. This means that for all $1 \leq j \leq n$, there exists a unique $l_j^N \in \llbracket 1, n \rrbracket$ such that $(\tau_{l_j^N}^N, \zeta_{l_j^N}^N) \in B((\tau_j, \zeta_j), \eta)$. We note that for all $1 \leq j \leq n-1$, $\tau_{l_j^N}^N < \tau_j + \gamma/3 < \tau_{j+1} - \gamma/3 < \tau_{l_{j+1}^N}^N$, so this implies that $\tau_{l_1^N}^N < \tau_{l_2^N}^N < \dots < \tau_{l_n^N}^N$, 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), $(\tau_j^N, \zeta_j^N) \in B((\tau_j, \zeta_j), \eta)$,

$$i.e. \quad |\tau_j - \tau_j^N| \vee |\zeta_j - \zeta_j^N| < \eta < \varepsilon.$$

Thus, τ_j^N and ζ_j^N converge respectively to τ_j and ζ_j .

Notice that

$$Z_t^n = \sum_{j=1}^n \mathbf{1}_{\{\zeta_j^n \leq x_{\tau_j^n}^n\}} \mathbf{1}_{\{\tau_j^n \leq t\}}.$$

Now we argue that $\mathbf{1}_{\{\zeta_j^n \leq x_{\tau_j^n}^n\}}$ converges to $\mathbf{1}_{\{\zeta_j \leq x_{\tau_j}\}}$. Indeed: there are two cases, either $\zeta_j < x_{\tau_j}$, or $\zeta_j > x_{\tau_j}$, in the first case we consider $\varepsilon > 0$ such that $\zeta_j + \varepsilon < x_{\tau_j}$. Then using Lemma 5.9, for n big enough, we have $\zeta_j^n < \zeta_j + \varepsilon/3 < x_{\tau_j} - \varepsilon/3 < x_{\tau_j^n}^n$, implying the convergence of $\mathbf{1}_{\{\zeta_j^n \leq x_{\tau_j^n}^n\}}$. The second case is handled in the same way. For the same reason, $\mathbf{1}_{\{\tau_j^n \leq t\}}$ converges to $\mathbf{1}_{\{\tau_j \leq t\}}$ (since we chose t such that $\pi(\{t\} \times \mathbb{R}_+) = 0$).

To resume, we have shown that for all $t \geq 0$ satisfying that, $\pi(\{t\} \times \mathbb{R}_+) = 0$ and for all $n \in \mathbb{N}^*$, $\pi^n(\{t\} \times \mathbb{R}_+) = 0$, Z_t^n converges to Z_t . Observing that these points are dense in \mathbb{R}_+ , we can apply Lemma 5.10 to obtain that Z^n converges to Z in $D([0, t], \mathbb{R})$ for all t with the above properties. We observe that such t are points of continuity of Z , and that we can choose an increasing sequence $(t_n)_n$ of such points that goes to infinity. As a consequence, Proposition 16.2 of (Billingsley 1999) implies that Z^n converges to Z in $D(\mathbb{R}_+, \mathbb{R}_+)$. \square

Before proving Theorem 4.1, we state two lemmas, whose proofs are omitted.

Lemma 5.11. *Let $(E_n)_{n \in \mathbb{N}}$ be a sequence of topological spaces, and, for each n , X_n be a tight random variable on E_n . Then $(X_n)_{n \in \mathbb{N}}$ is tight on $\prod_{n \in \mathbb{N}} E_n$ for the product topology.*

Lemma 5.12. *Let $(E_n, d_n)_{n \in \mathbb{N}}$ be a sequence of separable metric spaces. Then $\prod_{n \in \mathbb{N}} E_n$ is a separable metric space.*

We can end this subsection with the

Proof of Theorem 4.1. Firstly, let us prove that $(Y^N, \pi^1, \dots, \pi^k, \dots)$ converges in distribution to $(\bar{Y}, \bar{\pi}^1, \dots, \bar{\pi}^k, \dots)$ in $D(\mathbb{R}_+, \mathbb{R}) \times (\mathcal{M}^\#)^{\mathbb{N}^*}$.

By hypothesis, the sequence $(Y^N)_N$ converges in $D(\mathbb{R}_+, \mathbb{R})$, so it is tight, and each π^k is tight on $\mathcal{M}^\#$ thanks to Ulam's theorem (see Theorem 1.3 of (Billingsley 1999)), since $\mathcal{M}^\#$ is a Polish space (see Theorem A2.6.III.(i) of (Daley and Vere-Jones 2003)). Then, by Lemma 5.11, $(\pi^1, \dots, \pi^k, \dots)$ is tight on $(\mathcal{M}^\#)^{\mathbb{N}^*}$. This implies that the sequence $(Y^N, \pi^1, \dots, \pi^k, \dots)_N$ is tight on $D(\mathbb{R}_+, \mathbb{R}) \times (\mathcal{M}^\#)^{\mathbb{N}^*}$.

To show that this sequence converges, it is sufficient to prove that each converging subsequence converges to the same limit (see Corollary of Theorem 5.1 of (Billingsley 1999)). Let P be a limit distribution for the sequence of tuples $(Y^N, \pi^1, \dots, \pi^k, \dots)$. The marginals of P are respectively the distributions of \bar{Y} and those of $\bar{\pi}^k$ ($k \in \mathbb{N}^*$). Since the variables $\bar{Y}, \bar{\pi}^1, \dots, \bar{\pi}^k, \dots$ are assumed to be independent, the limit distribution P is uniquely determined by its marginals. As a consequence $(Y^N, \pi^1, \dots, \pi^k, \dots)$ converges in distribution to $(\bar{Y}, \bar{\pi}^1, \dots, \bar{\pi}^k, \dots)$ in $D(\mathbb{R}_+, \mathbb{R}) \times (\mathcal{M}^\#)^{\mathbb{N}^*}$.

As $D(\mathbb{R}_+, \mathbb{R}) \times (\mathcal{M}^\#)^{\mathbb{N}^*}$ is a separable metric space (see Theorem 16.3 of (Billingsley 1999), Theorem A2.6.III.(i) of (Daley and Vere-Jones 2003) and Lemma 5.12), Skorohod's representation theorem (Theorem 6.7 of (Billingsley 1999)) implies that there exist random variables \tilde{Y}, \tilde{Y}^N ($N \in \mathbb{N}^*$), $\tilde{\pi}^k$ ($k \in \mathbb{N}^*$), $\tilde{\pi}^{N,k}$ ($N, k \in \mathbb{N}^*$) defined on some probability space Ω' such that:

- $(\tilde{Y}, \tilde{\pi}^1, \dots, \tilde{\pi}^k, \dots)$ has the same distribution as $(\bar{Y}, \bar{\pi}^1, \dots, \bar{\pi}^k, \dots)$,
- $(\tilde{Y}^N, \tilde{\pi}^{N,1}, \dots, \tilde{\pi}^{N,k}, \dots)$ has the same distribution as $(Y^N, \pi^1, \dots, \pi^k, \dots)$,
- \tilde{Y}^N converges almost surely to \tilde{Y} in Skorohod topology,
- $\tilde{\pi}^{N,k}$ converges almost surely to $\tilde{\pi}^k$ in (\mathcal{N}, d) , for each $k \in \mathbb{N}^*$.

One can note that the hypothesis Theorem 5.7 are satisfied because of the independence between \tilde{Y} and $\tilde{\pi}^k$, and because $\tilde{\pi}^k$ ($k \geq 1$) are Poisson measures with Lebesgue intensity. As a consequence, for every $k \in \mathbb{N}^*$, the point process $\tilde{Z}^{N,k} := \Phi(\tilde{Y}^N, \tilde{\pi}^{N,k})$ converges almost surely to $\tilde{Z}^k := \Phi(\tilde{Y}, \tilde{\pi}^k)$ in $D(\mathbb{R}_+, \mathbb{R})$. This implies the almost sure convergence of $(\tilde{Z}^{N,k})_{k \in \mathbb{N}^*}$ to $(\tilde{Z}^k)_{k \in \mathbb{N}^*}$ in $D(\mathbb{R}_+, \mathbb{R})^{\mathbb{N}^*}$ as N goes to infinity. Consequently, the theorem is proved. \square

Remark 5.13. *Under the hypothesis of Theorem 4.1, we can prove that, for all $k \in \mathbb{N}^*$, the system $(Z^{N,i})_{1 \leq i \leq k}$ converges to $(\bar{Z}^i)_{1 \leq i \leq k}$ in the space $D(\mathbb{R}_+, \mathbb{R}^k)$. However, the convergence in $D(\mathbb{R}_+, \mathbb{R}^{\mathbb{N}^*})$ is false.*

5.6. Proof of Lemma 4.2

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

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

5.7. Proof of Proposition 4.3

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 converging 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

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

Step 1. We show that π is a $(P', (\mathcal{F}'_t)_{t \geq 0})$ -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 $\lambda_1, \dots, \lambda_k \geq 0$,

$$\mathbb{E} \left(\exp \left[- \sum_{i=1}^k \lambda_i \pi([s, t] \times U_i) \right] \middle| \mathcal{F}'_s \right) = \exp \left[(t-s) \sum_{i=1}^k (e^{-\lambda_i} - 1) \nu(U_i) \right]. \quad (31)$$

To prove (31), 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 $V_1, \dots, V_n \in \mathcal{B}(\mathbb{R}_+ \times \mathbb{R})$,

$$\begin{aligned} & \mathbb{E} \left(\exp \left[- \sum_{i=1}^k \lambda_i \pi([s, t] \times U_i) \right] \varphi_1(Z_{s_1}) \times \dots \times \varphi_n(Z_{s_n}) \right) = \\ & \exp \left[(t-s) \sum_{i=1}^k (e^{-\lambda_i} - 1) \nu(U_i) \right] \mathbb{E} (\varphi_1(Z_{s_1}) \times \dots \times \varphi_n(Z_{s_n})), \end{aligned} \quad (32)$$

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

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

$$\varphi(Y_t) - \varphi(Y_0) + \alpha \int_0^t \varphi'(Y_v) Y_v dv - \frac{\sigma^2}{2} \int_0^t \varphi''(Y_v) f(Y_v) dv \quad (33)$$

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

$$\mathbb{E} \left[\left(\varphi(Y_t) - \varphi(Y_s) + \alpha \int_s^t \varphi'(Y_v) Y_v dv - \frac{\sigma^2}{2} \int_s^t \varphi''(Y_v) f(Y_v) dv \right) \prod_{i=1}^n \psi_i(Z_{s_i}) \right] = 0. \quad (34)$$

To prove (34), 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 $N \geq 1$,

$$\mathbb{E} \left[\left(W_t^N - W_s^N \right) \prod_{k=1}^n \psi_k(Z_{s_k}^N) \right] = 0. \quad (35)$$

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,

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

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 $F_{s,t}^M(x, m) := F_{s,t}(x, m) \cdot \xi_M(\sup_{0 \leq r \leq t} |x_r|)$, 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 (34), that is, $\mathbb{E}[F_{s,t}(Y, \pi)] = 0$. Recalling (35), we start from

$$\begin{aligned} |\mathbb{E}[F_{s,t}(Y, \pi)]| &= \left| \mathbb{E}[F_{s,t}(Y, \pi)] - \mathbb{E}\left[\left(W_t^N - W_s^N\right) \prod_{k=1}^N \psi_k(Z_{s_k}^N)\right] \right| \\ &\leq \left| \mathbb{E}\left[F_{s,t}(Y, \pi) \left(1 - \xi_M\left(\sup_{0 \leq r \leq t} |Y_r|\right)\right)\right] \right| \end{aligned} \quad (36)$$

$$+ \left| \mathbb{E}\left[F_{s,t}(Y, \pi) \xi_M\left(\sup_{0 \leq r \leq t} |Y_r|\right)\right] - \mathbb{E}\left[F_{s,t}(X^N, \pi_1) \xi_M\left(\sup_{0 \leq r \leq t} |X_r^N|\right)\right] \right| \quad (37)$$

$$+ \left| \mathbb{E}\left[F_{s,t}(X^N, \pi_1) \left(1 - \xi_M\left(\sup_{0 \leq r \leq t} |X_r^N|\right)\right)\right] \right|. \quad (38)$$

Using the fact that $1 - \xi_M(x) \leq \mathbb{1}_{\{|x| > M\}}$, Cauchy-Schwarz's inequality, Markov's inequality and Lemma 2.1, we can bound (36) and (38) by C/\sqrt{M} for some $C > 0$ that is independent of N .

Now, fix some $\varepsilon > 0$ and consider a constant $M_\varepsilon > 0$ such that (36) and (38) are smaller than ε . In a next step, we choose an integer N_ε big enough such that (37) 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 (33) 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 (33) 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 (33) with $\varphi_K \in C_b^3(\mathbb{R})$ is a martingale.

Hence we have shown that the expression in (33) 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 (33) 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 $(\mathcal{F}_t^0)_{t \geq 0}$. 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 (33), for the test-function $\varphi = h$ that is C^3 . Thus we know that $(B'_t)_t$ is a $(P', (\mathcal{F}'_{t \geq 0}))$ -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 .

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.
- DAWSON, D. and VAILLANCOURT, J. (1995). Stochastic McKean-Vlasov equations. *Nonlinear Differential Equations and Applications NoDEA* **2** 199-229.
- 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.
- 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.
- GRAHAM, C. (2019). Regenerative properties of the linear hawkes process with unbounded memory. *arXiv:1905.11053 [math, stat]*.
- 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.
- JACOD, J. and SHIRYAEV, A. N. (2003). *Limit Theorems for Stochastic Processes*, Second ed. Springer-Verlag Berlin Heidelberg New York.
- KHASHMINSKII, R. (2012). *Stochastic stability of differential equations*, Second ed. Springer.
- 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.
- KURTZ, T. G. and XIONG, J. (1999). Particle representations for a class of nonlinear SPDEs. *Stochastic Processes and their Applications* **83** 103–126.
- 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. (1992). Stability of Markovian Processes I: Criteria for Discrete-Time Chains. *Advances in Applied Probability* **24** 542–574.
- 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.
- RAAD, M. B. (2019). Renewal Time Points for Hawkes Processes. *arXiv:1906.02036 [math]*.
- 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.
- 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).