



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. *Bernoulli*, 2022, 28 (1). hal-02096662v4

HAL Id: hal-02096662

<https://hal.science/hal-02096662v4>

Submitted on 23 Nov 2020

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

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 et FR FP2M 2036 CNRS*

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 the Skorokhod 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 neuroscience 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, 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).

In (Delattre, Fournier and Hoffmann, 2016), the authors 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 $ds dz d\mu(u)$ 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

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

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 given as solution of the SDE

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

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 the Skorokhod 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 the Skorokhod 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

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

Here $\bar{P}_t g(x) = \mathbb{E}_x [g(\bar{X}_t)]$ and $P_t^N g(x) = \mathbb{E}_x [g(X_t^N)]$ 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 $t \mapsto P_t^N g(x)$ in 0. The proof of formula (4) for our extended generators is given in the Appendix (Proposition 5.6).

It is well-known that under suitable assumptions on f , the solution of (3) admits a unique invariant measure λ whose density is explicitly known. Thus, a natural question is to consider the limit of the law $\mathcal{L}(X_t^N)$ of X_t^N when t and N go simultaneously to infinity. We prove that the limit of $\mathcal{L}(X_t^N)$ is λ , 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 converges 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 the Appendix, we collect some results about extended generators and we give the proof of (4) together with 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 λ a measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that g is integrable with respect to λ , we write $\lambda(g)$ for $\int_{\mathbb{R}} g d\lambda$.
- 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 Skorokhod space of càdlàg functions from \mathbb{R}_+ to \mathbb{R} , endowed with the Skorokhod metric (see Chapter 3 Section 16 of (Billingsley, 1999)), and $D(\mathbb{R}_+, \mathbb{R}_+)$ for this space restricted to non-negative functions.
- 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 5.8).

The aim of this paper is to provide explicit bounds for the convergence of X^N in the 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 $\bar{\nu}_0$ is 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 1 is used in many computations of the paper in one of the following forms:

- $\forall x \in \mathbb{R}, f(x) \leq (\sqrt{f(0)} + L|x|)^2$,

or, if we do not need the accurate dependency on the parameter,

- $\forall x \in \mathbb{R}, f(x) \leq C(1 + x^2)$.

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 finite 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 belongs to $C^4(\mathbb{R})$ and that 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 the Skorokhod space, with an explicit rate of convergence for their semigroups. This rate of convergence will be expressed in terms of the following parameters

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

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

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

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

Recall that $\bar{P}_t g(x) = \mathbb{E}_x [g(\bar{X}_t)]$ and $P_t^N g(x) = \mathbb{E}_x [g(X_t^N)]$ denote the Markovian semigroups of \bar{X} and X^N .

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 $\alpha > \frac{7}{6}\sigma^2L^2$, 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. Theorem 1.4.(ii) states the convergence of X^N to \bar{X} in the Skorokhod topology. Since \bar{X} is almost surely continuous, this implies the, a priori stronger, convergence in distribution in the topology of the uniform convergence on compact sets. Indeed, according to Skorokhod's representation theorem (see Theorem 6.7 of (Billingsley, 1999)), we can assume that X^N converges almost surely to \bar{X} in the Skorokhod space, and this classically entails the uniform convergence on every compact set (see the discussion at the bottom of page 124 in Section 12 of (Billingsley, 1999)).

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

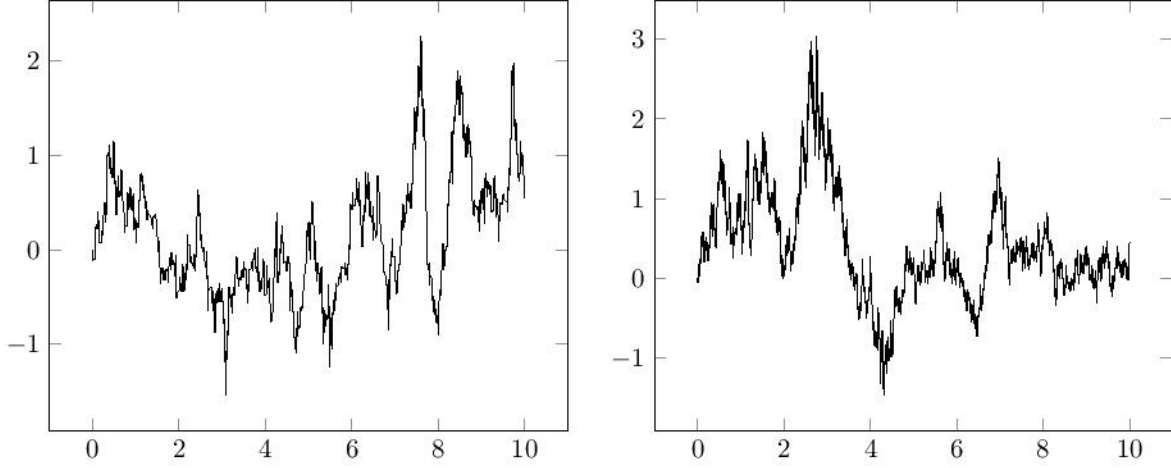


FIG 1. Simulation of trajectories of $(X_t^N)_{0 \leq t \leq 10}$ with $X_0^N = 0$, $\alpha = 1$, $\mu = \mathcal{N}(0,1)$, $f(x) = 1 + x^2$, $N = 100$ (left picture) and $N = 500$ (right picture).

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

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

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

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

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 λ 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 λ 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 $x \in \mathbb{R}$,

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

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 $(B_t)_{t \geq 0}$. The limit point processes \bar{Z}^i are then defined by

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

Theorem 1.7. *Under Assumptions 1, 2 and 4, for every $k \in \mathbb{N}^*$, the sequence $(Z^{N,1}, \dots, Z^{N,k})_N$ converges to $(\bar{Z}^1, \dots, \bar{Z}^k)$ in distribution in $D(\mathbb{R}_+, \mathbb{R}^k)$. Consequently, the sequence $(Z^{N,j})_{j \geq 1}$ converges to $(\bar{Z}^j)_{j \geq 1}$ 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} , for any $k > 1$, $\bar{Z}^1, \dots, \bar{Z}^k$ 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 $\sup_{0 \leq t \leq T} \mathbb{E}_x [(\bar{X}_t)^4] \leq C_T(1 + x^4)$.*
- (v) *For all $0 \leq s, t \leq T$ and $x \in \mathbb{R}$,*

$$\mathbb{E}_x [(\bar{X}_t - \bar{X}_s)^2] \leq C_T(1 + x^2)|t - s| \text{ and } \mathbb{E}_x [(X_t^N - X_s^N)^2] \leq C_T(1 + x^2)|t - s|.$$

We postpone the proof of Lemma 2.1 to the Appendix. The inequalities of points (i) and (ii) of the lemma hold for any fixed $\varepsilon > 0$. This parameter ε appears for the following reason. We prove the above points using the Lyapunov function $x \mapsto x^2$. When applying the generators to this function, there are terms of order x that appear and that we bound by $x^2\varepsilon + \varepsilon^{-1}$ to be able to compare it to x^2 .

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 below 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} the one 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 Itô'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

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

The following result is the first step towards the proof of our main result.

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 the following representation

$$(\bar{P}_t - P_t^N) g(x) = \int_0^t P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x) ds, \quad (10)$$

which is proven in Proposition 5.6 in the Appendix.

Obviously, to be able to apply Proposition 2.3 to 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 $\|\bar{P}_s g\|_{3,\infty}$. 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 $x \mapsto \bar{P}_t g(x)$ belongs to $C_b^3(\mathbb{R})$ and satisfies

$$\left\| (\bar{P}_t g)''' \right\|_{\infty} \leq C \|g\|_{3,\infty} (1 + t^2) e^{\beta t}, \quad (11)$$

with $\beta = \max(\frac{1}{2}\sigma^2 L^2 - \alpha, 2\sigma^2 L^2 - 2\alpha, \frac{7}{2}\sigma^2 L^2 - 3\alpha)$. Moreover, for all $T > 0$,

$$\sup_{0 \leq t \leq T} \|\bar{P}_t g\|_{3,\infty} \leq Q_T \|g\|_{3,\infty} \quad (12)$$

for some $Q_T > 0$, and for all $i \in \{0, 1, 2\}$ 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.

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

Proof of Theorem 1.4. Step 1. The proof of point (i) is a straightforward consequence of Proposition 2.3, since, applying formula (10),

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

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. We now give the proof of 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

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

$$\|\nu_1 - \nu_2\|_{TV} = \frac{1}{2} \sup_{g: \|g\|_{\infty} \leq 1} |\nu_1(g) - \nu_2(g)|.$$

Proposition 3.1. *If Assumptions 1 and 2 hold, then the invariant measure λ 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 $t \geq 0$,

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

Proof. In a first time, let us prove the positive Harris recurrence of \bar{X} implying the existence and uniqueness of λ . 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

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

For $x > 0$, and using that f is subquadratic,

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

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

$$\bar{A}V(x) \leq -cV(x) + d. \tag{13}$$

Obviously, for any fixed $T > 0$, the sampled chain $(\bar{X}_{kT})_{k \geq 0}$ is Feller and λ -irreducible. The support of λ being \mathbb{R} , Theorem 3.4 of (Meyn and Tweedie, 1993) implies that every compact set is petite for the sampled chain. Then, as (13) 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

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

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 $g \in C_b^3(\mathbb{R})$,

$$|P_t^N g(x) - \lambda(g)| \leq |P_t^N g(x) - \bar{P}_t g(x)| + |\bar{P}_t g(x) - \lambda(g)|$$

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

where we have used Theorem 1.4 and Proposition 3.1. Since $(1+x^2)^q \leq 1+x^2$, q being smaller than $1/2$, this implies the result. \square

4. Proof of Theorem 1.7

We prove the result using Theorem IX.4.15 of (Jacod and Shiryaev, 2003).

Let $k \in \mathbb{N}^*$, let us note $Y^N := (X^N, Z^{N,1}, \dots, Z^{N,k})$ and $\bar{Y} := (\bar{X}, \bar{Z}^1, \dots, \bar{Z}^k)$. Using the notation of Theorem IX.4.15 of (Jacod and Shiryaev, 2003) with the semimartingales Y^N ($N \in \mathbb{N}^*$) and \bar{Y} and denoting e^j ($0 \leq j \leq k$) the j -th unit vector, we have:

- $b^{N,0}(x) = b^0(x) = -\alpha x$ and $b^{N,i}(x) = b^i(x) = 0$ for $1 \leq i \leq k$,
- $\tilde{c}^{N,0,0}(x) = c^{0,0}(x) = \sigma^2 f(x^0)$ and $c^{N,i,j}(x) = c^{i,j}(x) = 0$ for $(i,j) \neq (0,0)$,
- $g * K^N(x) = f(x^0) \sum_{j=1}^k \int_{\mathbb{R}} g\left(\frac{u}{\sqrt{N}}e^0 + e^j\right) d\mu(u) + (N-k) \int_{\mathbb{R}} g\left(\frac{u}{\sqrt{N}}e^0\right) d\mu(u)$,
- $g * K(x) = f(x^0) \sum_{j=1}^k g(e^j)$.

The only condition of Theorem IX.4.15 that is not straightforward is the convergence of $g * K^N$ to $g * K$ for $g \in C_1(\mathbb{R}^{k+1})$. The complete definition of $C_1(\mathbb{R}^{k+1})$ is given in VII.2.7 of (Jacod and Shiryaev, 2003), but here, we just use the fact that $C_1(\mathbb{R}^{k+1})$ is a subspace of $C_b(\mathbb{R}^{k+1})$ containing functions which are zero around zero. This convergence follows from the fact that any $g \in C_1(\mathbb{R}^{k+1})$ can be written as $g(x) = h(x)\mathbb{1}_{\{|x|>\varepsilon\}}$ where $h \in C_b(\mathbb{R}^{k+1})$ and $\varepsilon > 0$. This allows to show that, for this kind of function g ,

$$\begin{aligned} \left| (N-k)f(x^0) \int_{\mathbb{R}} g\left(\frac{u}{\sqrt{N}}e^0\right) d\mu(u) \right| &\leq (N-k)f(x^0)\|h\|_{\infty} \int_{\mathbb{R}} \mathbb{1}_{\{|u|>\varepsilon\sqrt{N}\}} d\mu(u) \\ &\leq f(x^0)C \frac{N-K}{N^2} \leq Cf(x^0)N^{-1}, \end{aligned}$$

where the second inequality follows from the fact that we assume that μ is a probability measure having a finite fourth moment.

Theorem IX.4.15 of (Jacod and Shiryaev, 2003) implies that for all $k \geq 1$, $(Z^{N,1}, \dots, Z^{N,k})$ converges to $(\bar{Z}^1, \dots, \bar{Z}^k)$ in distribution in $D(\mathbb{R}_+, \mathbb{R}^k)$.

This implies the weaker convergence in $D(\mathbb{R}_+, \mathbb{R})^k$ for any $k \in \mathbb{N}^*$. Then, the convergence in $D(\mathbb{R}_+, \mathbb{R})^{\mathbb{N}^*}$ is classical (see e.g. Theorem 3.29 of (Kallenberg, 1997)).

5. Appendix

5.1. Extended generators

There are different definitions of infinitesimal generators in the literature. The aim of this subsection is to define precisely the notion of generator we use in this paper. Moreover we establish some properties of these generators and prove formula (10). 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 5.6, 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)). These extended generators are defined by the point-wise convergence on \mathbb{R} instead of the uniform convergence. Moreover, they verify 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.

Let $(X_t)_t$ be a Markov process taking values in \mathbb{R} . We set

$$\mathcal{D}(P) = \{g : \mathbb{R} \rightarrow \mathbb{R}, \text{ measurable, s.t. } \forall x \in \mathbb{R}, \forall t \geq 0, \mathbb{E}_x |g(X_t)| < \infty\}.$$

For $g \in \mathcal{D}(P)$, $x \in \mathbb{R}$, $t \geq 0$, we define $P_t g(x) = \mathbb{E}_x [g(X_t)]$.

Definition 5.1. We define $\mathcal{D}'(A)$ to be the set of $g \in \mathcal{D}(P)$ for which there exists a measurable function $Ag : \mathbb{R} \rightarrow \mathbb{R}$, such that $Ag \in \mathcal{D}(P)$, $t \mapsto P_t Ag(x)$ is continuous in 0, and $\forall x \in \mathbb{R}, \forall t \geq 0$,

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

Remark 5.2. Using Fubini's theorem and (ii) we can rewrite (i) in the following form:

$$P_t g(x) - g(x) = \int_0^t P_s Ag(x) ds. \quad (14)$$

Then (14) implies immediately that if $g \in \mathcal{D}'(A)$, then

$$\lim_{t \rightarrow 0} \frac{1}{t} (P_t g(x) - g(x)) = Ag(x). \quad (15)$$

Note also that it follows from the Markov property and the definition of Ag that the process $g(X_t) - g(X_0) - \int_0^t Ag(X_s) ds$ is a \mathbb{P}_x -martingale w.r.t. to the filtration generated by $(X_t)_t$.

The following result is classical and stated without proof. It is a straightforward consequence of (14) and (15).

Proposition 5.3. Suppose that A is the extended generator of the semigroup $(P_t)_t$, $g \in \mathcal{D}'(A)$, and the map $s \rightarrow P_s Ag(x)$ is continuous on \mathbb{R}_+ for some $x \in \mathbb{R}$. Then

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

Moreover, if for all $t \geq 0$, $P_t g \in \mathcal{D}'(A)$, then $\frac{d}{dt} P_t g(x) = A P_t g(x) = P_t Ag(x)$.

In what follows, we give some sufficient conditions to verify the continuity and the derivability of the map $s \mapsto P_s h(x)$. These conditions are not intended to be optimal, they are stated such that it is easy to check them both for X^N and \bar{X} .

Proposition 5.4. *Let $(X_t)_t$ be a Markov process with semigroup $(P_t)_t$ and extended generator A .*

1. *Let $h \in \mathcal{D}(P)$, $x \in \mathbb{R}$. Suppose that*

(i) *the map $t \rightarrow X_t$ is continuous in \mathbb{L}^2 , i.e. $\lim_{|t-s| \rightarrow 0} \mathbb{E}_x |X_s - X_t|^2 = 0$;*

(ii) *for all $T > 0$, $\sup_{0 \leq t \leq T} \mathbb{E}_x (|X_t|^4) < +\infty$;*

(iii) *there exists $C > 0$, such that $\forall x, y \in \mathbb{R}$, $|h(x) - h(y)| \leq C(1 + x^2 + y^2)|x - y|$.*

Then the map $s \mapsto P_s h(x)$ is continuous on \mathbb{R}_+ .

2. *Suppose moreover that (i), (ii) and (iii)' are satisfied with*

(iii)' *$g \in \mathcal{D}'(A)$ such that for some $C > 0$, and for all $x, y \in \mathbb{R}$, we have that $|Ag(x) - Ag(y)| \leq C(1 + x^2 + y^2)|x - y|$.*

Then the map $s \rightarrow P_s g(x)$ is differentiable on \mathbb{R}_+ , and $\frac{d}{dt} P_t g(x) = P_t Ag(x)$.

Proof. The proof of point 1. follows from the following chain of inequalities

$$\begin{aligned} |P_t h(x) - P_s h(x)| &\leq \mathbb{E}_x |h(X_t) - h(X_s)| \leq C \mathbb{E}_x [(1 + X_t^2 + X_s^2) |X_t - X_s|] \leq \\ &C [\mathbb{E}_x (1 + X_t^4 + X_s^4)]^{1/2} [\mathbb{E}_x |X_t - X_s|^2]^{1/2} \leq C \sup_{u \leq s \vee t} [\mathbb{E}_x X_u^4]^{1/2} \|X_s - X_t\|_2 \xrightarrow{|t-s| \rightarrow 0} 0. \end{aligned}$$

The second assertion of the proof follows from point 1. and Proposition 5.3, observing that $h := Ag$ satisfies point (iii). □

5.2. Proof of (10)

In this section, we first collect some useful results about the extended generators A^N of X^N and \bar{A} of \bar{X} . Then we give the proof of (10). We start with the following result.

Proposition 5.5. 1. *For all $g \in C_b^3(\mathbb{R})$, for all $x, y \in \mathbb{R}$,*

$$|\bar{A}g(x) - \bar{A}g(y)| \leq C \|g\|_{3,\infty} (1 + x^2 + y^2) |x - y| \quad \text{and} \quad |\bar{A}g(x)| \leq C \|g\|_{2,\infty} (1 + x^2).$$

In particular, for any $g \in C_b^3$, the map $t \rightarrow \bar{P}_t g(x)$ is differentiable on \mathbb{R}_+ , and $\frac{d}{dt} \bar{P}_t g(x) = \bar{P}_t \bar{A}g(x) = \bar{A} \bar{P}_t g(x)$.

2. *For all $g \in C_b^2(\mathbb{R})$, for all $x, y \in \mathbb{R}$,*

$$|A^N g(x) - A^N g(y)| \leq C \|g\|_{2,\infty} (1 + x^2 + y^2) |x - y| \quad \text{and} \quad |A^N g(x)| \leq C \|g\|_{1,\infty} (1 + x^2).$$

In particular, for any $g \in C_b^2$, the map $t \rightarrow P_t^N g(x)$ is differentiable on \mathbb{R}_+ , and $\frac{d}{dt} P_t^N g(x) = P_t^N A^N g(x)$.

Proof. The result follows from Proposition 5.4 together with Lemma 2.1 and Lemma 2.2. Finally, to show that $\bar{P}_t \bar{A}g(x) = \bar{A} \bar{P}_t g(x)$, we use Proposition 5.3 and Proposition 2.4. □

We are now able to give the proof of the main result of this section. This result is a Trotter-Kato like formula that allows to obtain a control of the difference between the semigroups \bar{P} and P^N , provided we dispose already of a control of the distance between their generators. It is an adaptation of Lemma 1.6.2 from (Ethier and Kurtz, 2005) to the notion of extended generators.

Proposition 5.6. *Grant Assumptions 1, 2 and 3. Let \bar{A} and A^N be the extended generators of respectively \bar{P} and P^N .*

Then the following equality holds for each $g \in C_b^3(\mathbb{R})$, $x \in \mathbb{R}$ and $t \in \mathbb{R}_+$.

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

Proof. 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. Introduce for $0 \leq s \leq t$ the function $u(s) = P_{t-s}^N \bar{P}_s g(x)$.

One can note that $u = \Phi \circ \Psi$ with $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}$; $\Phi(v_1, v_2) = P_{v_1}^N \bar{P}_{v_2} g(x)$ and $\Psi : \mathbb{R} \rightarrow \mathbb{R}^2$; $\Psi(s) = (t-s, s)$. Let us show that Φ is differentiable w.r.t. to both variables v_1 and v_2 . Indeed, for v_1 it is a consequence of the fact that $h = \bar{P}_{v_2} g \in C_b^3(\mathbb{R})$ by Proposition 2.4 and Proposition 5.5, from which we know that if $h \in C_b^2$, then $v_1 \mapsto P_{v_1}^N h(x)$ is differentiable and

$$\frac{\partial}{\partial v_1} \Phi(v_1, v_2) = \frac{d}{dv_1} P_{v_1}^N h(x) = P_{v_1}^N A^N h(x).$$

To show the differentiability of Φ with respect to v_2 , let us write $\Phi(v_1, v_2) = \mathbb{E}_x [\bar{P}_{v_2} g(X_{v_1}^N)]$. From Proposition 5.5, we know that since $g \in C_b^3$, $v_2 \mapsto \bar{P}_{v_2} g(X_{v_1}^N)$ is a.s. differentiable with derivative

$$\frac{d}{dv_2} \bar{P}_{v_2} g(X_{v_1}^N) = \bar{A} \bar{P}_{v_2} g(X_{v_1}^N) = \bar{P}_{v_2} \bar{A} g(X_{v_1}^N) = \mathbb{E}_{X_{v_1}^N} (\bar{A} g)(\bar{X}_{v_2}).$$

Moreover, $|\bar{A} g(x)| \leq C \|g\|_{2,\infty} (1+x^2)$ by Proposition 5.5. Now, using Lemma 2.1.(ii) we see that

$$\sup_{v_2 \leq T} \left| \frac{d}{dv_2} \bar{P}_{v_2} g(X_{v_1}^N) \right| \leq \mathbb{E}_{X_{v_1}^N} \left[\sup_{v_2 \leq T} |(\bar{A} g)(\bar{X}_{v_2})| \right] \leq C_T (1 + (X_{v_1}^N)^2).$$

By Lemma 2.1.(iii), we see that the last bound is integrable, hence by dominated convergence, $v_2 \mapsto \Phi(v_1, v_2)$ is differentiable with derivative

$$\frac{\partial}{\partial v_2} \Phi(v_1, v_2) = P_{v_1}^N \bar{A} \bar{P}_{v_2} g(x) = P_{v_1}^N \bar{P}_{v_2} \bar{A} g(x).$$

As a consequence, u is differentiable on \mathbb{R}_+ , and we have

$$\begin{aligned} u'(s) &= -\frac{\partial}{\partial v_1} \Phi(t-s, s) + \frac{\partial}{\partial v_2} \Phi(t-s, s) \\ &= -P_{t-s}^N A^N \bar{P}_s g(x) + P_{t-s}^N \bar{P}_s \bar{A} g(x) \\ &= P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x). \end{aligned}$$

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

$$|P_{t-s}^N (\bar{A} - A^N) \bar{P}_s g(x) - P_{t-s_k}^N (\bar{A} - A^N) \bar{P}_{s_k} g(x)| \leq |(P_{t-s}^N - P_{t-s_k}^N) (\bar{A} - A^N) g_s(x)| \quad (17)$$

$$+ |P_{t-s_k}^N (\bar{A} - A^N) (\bar{P}_s - \bar{P}_{s_k}) g(x)|, \quad (18)$$

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

To show that the term (17) vanishes when k goes to infinity, denote $h_s(x) = (\bar{A} - A^N)g_s(x)$. Using Proposition 5.5 and the fact that $g_s \in C_b^3(\mathbb{R})$, we have

$$|h_s(x) - h_s(y)| \leq C(1 + x^2 + y^2)|x - y|.$$

Proposition 5.4 applied to h_s and to P^N implies that $u \rightarrow P_u^N h_s(x)$ is continuous. As a consequence the term (17) vanishes as $k \rightarrow \infty$.

To finish the proof, we need to show that the term (18) vanishes. Denote $g_k = (\bar{P}_s - \bar{P}_{s_k})g$. We have to show that

$$\mathbb{E}_x [(\bar{A} - A^N) g_k(X_{t-s_k}^N)] \rightarrow 0, \text{ when } k \rightarrow \infty.$$

In what follows we will in fact show that

$$\mathbb{E}_x [\bar{A}g_k(X_{t-s_k}^N)] \rightarrow 0 \text{ and } \mathbb{E}_x [A^N g_k(X_{t-s_k}^N)] \rightarrow 0, \text{ when } k \rightarrow \infty. \quad (19)$$

To begin with, using Proposition 2.4, the functions g_k belong to $C_b^3(\mathbb{R})$, and for any $i \in \{0, 1, 2\}$, for all $y \in \mathbb{R}$, $g_k^{(i)}(y)$ vanishes as k goes to infinity. Using again Proposition 2.4, we see that for all $i \in \{0, 1, 2, 3\}$, $\|g_k^{(i)}\|_\infty$ is uniformly bounded in k . It follows that each sequence $(g_k^{(i)})_k$, $i \in \{0, 1, 2\}$, is uniformly equicontinuous and thus converges to zero uniformly on each compact interval.

We next show that this implies that also the sequences $(A^N g_k)_k$ and $(\bar{A}g_k)_k$ converge to zero uniformly on each compact interval. For $(\bar{A}g_k)_k$, this is immediate, since \bar{A} is a local operator having continuous coefficients. For $(A^N g_k)_k$, it follows from the fact that $A^N g_k(x) \rightarrow 0$ as $k \rightarrow \infty$ for each fixed x and the fact that by Lemma 5.7 given below, this sequence is uniformly (in k , for fixed N) equicontinuous on each compact.

We are now able to conclude. The sequence $(X_{t-s_k}^N)_k$ is almost surely bounded by $\sup_{0 \leq r \leq t} |X_r^N|$ which is finite almost surely by Lemma 2.1.(iii). Hence, almost surely as $k \rightarrow \infty$, $\bar{A}g_k(X_{t-s_k}^N) \rightarrow 0$ and $A^N g_k(X_{t-s_k}^N) \rightarrow 0$.

We now apply dominated convergence to prove (19). Using that by Proposition 5.5, for all $g \in C_b^3(\mathbb{R})$ and $x \in \mathbb{R}$,

$$|\bar{A}g(x)| \leq C\|g\|_{2,\infty}(1 + x^2),$$

we can bound the expression in the first expectation by

$$C\|g_k\|_{2,\infty} \left(1 + \left(\sup_{0 \leq r \leq t} |X_r^N| \right)^2 \right) \leq 2C \left(\sup_{0 \leq r \leq t} \|\bar{P}_r g\|_{2,\infty} \right) \left(1 + \left(\sup_{0 \leq r \leq t} |X_r^N| \right)^2 \right),$$

whose expectation is finite thanks to Lemma 2.1(iii). The same arguments work for A^N . This implies that (18) vanishes as $k \rightarrow \infty$, and this concludes the proof. \square

We now prove the missing lemma

Lemma 5.7. *For all $g \in C_b^2(\mathbb{R})$ and any $M > 0$,*

$$\sup_{x \in [-M, M]} |(A^N g)'(x)| \leq C_N \|g\|_{2,\infty} (1 + M^2),$$

for some constant $C_N > 0$ that can depend on N , but not on M .

Proof. We have

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

Since

$$\mathbb{E} \left[\left| g \left(x + \frac{U}{\sqrt{N}} \right) - g(x) \right| \right] \leq \frac{\|g'\|_\infty}{\sqrt{N}} \mathbb{E} [|U|], \quad \mathbb{E} \left[\left| g' \left(x + \frac{U}{\sqrt{N}} \right) - g'(x) \right| \right] \leq \frac{\|g''\|_\infty}{\sqrt{N}} \mathbb{E} [|U|],$$

we obtain

$$\sup_{x \in [-M, M]} |(A^N g)'(x)| \leq |\alpha| \|g\|_{2, \infty} (1 + M) + \sqrt{N} (|f'(x)| \vee |f(x)|) \|g\|_{2, \infty} \mathbb{E} [|U|].$$

Assumption 3 implies that $|f'(x)| \leq m_1 C \sqrt{1 + x^2}$ for all x . Together with the sub-quadraticity of f , this concludes the proof. \square

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

Proposition 5.8. *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 Löcherbach, 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 Itô's formula and Grönwall's lemma, we can prove that

$$\sup_{0 \leq s \leq t} \mathbb{E} \left[\left(X_{s \wedge \tau_K^N}^{N,K} \right)^2 \right] \leq C_t (1 + x^2), \quad (20)$$

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 $X_s^{N,K}(\omega)$ satisfies equation (5) with f replaced by f_K , we know that $X^N(\omega)$ verifies equation (5) on $[0, t]$. This holds for all $t > 0$. As a consequence, we know that X^N satisfies equation (5). This proves the existence of a 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

5.4. Proof of Lemma 2.1

Proof. 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 (20), one obtains for all $t \geq 0$, $\sup_{0 \leq s \leq t} \mathbb{E}[(X_s^N)^2]$ is finite. As a consequence $\Phi \in \mathcal{D}'(A^N)$ (in the sense of Definition 5.1). And, recalling that μ is centered and that $\sigma^2 := \int_{\mathbb{R}} u^2 d\mu(u)$, we have for all $x \in \mathbb{R}$,

$$\begin{aligned} A^N \Phi(x) &= -\alpha x \Phi'(x) + N f(x) \int_{\mathbb{R}} \left[\Phi\left(x + \frac{u}{\sqrt{N}}\right) - \Phi(x) \right] d\mu(u) \\ &= -2\alpha x^2 + N f(x) \int_{\mathbb{R}} \left[2x \frac{u}{\sqrt{N}} + \frac{u^2}{N} \right] d\mu(u) \\ &= -2\alpha \Phi(x) + \sigma^2 f(x) \leq -2\alpha \Phi(x) + \sigma^2 \left(L|x| + \sqrt{f(0)} \right)^2 \\ &\leq (\sigma^2 L^2 - 2\alpha) \Phi(x) + 2\sigma^2 L|x| \sqrt{f(0)} + \sigma^2 f(0). \end{aligned}$$

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

$$A^N \Phi(x) \leq c_\varepsilon \Phi(x) + d_\varepsilon, \quad (21)$$

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 Itô'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, denoting by $\tilde{\pi}_j(dt, dx, du) := \pi_j(dt, dx, du) - dt dx d\mu(u)$ the compensated measure of π_j ($1 \leq j \leq N$), $(M_t)_{t \geq 0}$ is the \mathbb{P}_x -local martingale defined as

$$M_t = \sum_{j=1}^N \int_{[0, t] \times \mathbb{R}_+ \times \mathbb{R}} \left[\Phi\left(X_{s-}^N + \frac{u}{\sqrt{N}}\right) - \Phi(x) \right] \mathbf{1}_{\{z \leq f(X_{s-}^N)\}} d\tilde{\pi}_j(s, z, u).$$

One can note that, since $\sup_{0 \leq s \leq t} \mathbb{E}[(X_s^N)^2]$ is finite for any $t \geq 0$, $(M_t)_{t \geq 0}$ is a locally square integrable \mathbb{P}_x -local martingale, and as a consequence, it is a \mathbb{P}_x -martingale.

Using (21), we obtain

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

implying

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

One deduces

$$\mathbb{E}_x \left[(X_t^N)^2 \right] \leq x^2 e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)t} + \frac{C}{\varepsilon} \left(e^{(\sigma^2 L^2 - 2\alpha + \varepsilon)t} + 1 \right), \quad (22)$$

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

$$\begin{aligned} \left(\sup_{0 \leq s \leq t} |X_t^N| \right)^2 &\leq 3 (X_0^N)^2 + 3\alpha^2 t \int_0^t (X_s^N)^2 ds \\ &\quad + 3 \sum_{j=1}^N \left(\sup_{0 \leq s \leq t} \left| \int_{]0,s] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_j(r, z, u) \right| \right)^2. \end{aligned} \quad (23)$$

Applying Burkholder-Davis-Gundy inequality to the last term above in (23), we can bound its expectation by

$$\begin{aligned} 3N \mathbb{E} \left[\int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} u^2 \mathbf{1}_{\{z \leq f(X_{s-}^N)\}} d\pi_j(s, z, u) \right] &\leq 3N\sigma^2 \int_0^t \mathbb{E} [f(X_{s-}^N)] ds \\ &\leq 3N\sigma^2 C \int_0^t (1 + \mathbb{E} [(X_s^N)^2]) ds. \end{aligned} \quad (24)$$

Now, bounding the expectation of (23) by (24), and using point (i) of the lemma we conclude the proof of (iii).

The assertion (iv) can be proved in classical way, applying Itô's formula and Grönwall's lemma. Let us explain how to prove this property for the process X^N . The proof for \bar{X} is similar. According to Itô's formula, for every $t \geq 0$,

$$\begin{aligned} (X_t^N)^4 &= (X_0^N)^4 - 4\alpha \int_0^t (X_s^N)^4 ds \\ &\quad + \sum_{j=1}^N \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} \left[(X_{s-}^N + \frac{u}{\sqrt{N}})^4 - (X_{s-}^N)^4 \right] \mathbf{1}_{\{z \leq f(X_{s-}^N)\}} d\pi_j(s, z, u) \\ &\leq (X_0^N)^4 + \sum_{j=1}^N \int_{]0,t] \times \mathbb{R}_+ \times \mathbb{R}} \left[(X_{s-}^N + \frac{u}{\sqrt{N}})^4 - (X_{s-}^N)^4 \right] \mathbf{1}_{\{z \leq f(X_{s-}^N)\}} d\pi_j(s, z, u). \end{aligned}$$

Let us recall that u is centered and has a finite fourth moment, and that f is subquadratic. Introducing the stopping times $\tau_K^N := \inf\{t > 0 : |X_t^N| > K\}$ for $K > 0$, and $u_K^N(t) := \mathbb{E} \left[(X_{t \wedge \tau_K^N}^N)^4 \right]$, it follows from the above that for all $t \geq 0$,

$$u_K^N(t) \leq C + Ct + C \int_0^t u_K^N(s) ds,$$

where C is a constant independent of t, N and K . This implies that for all $t \geq 0$,

$$\sup_{N \in \mathbb{N}^*} \sup_{K > 0} \sup_{0 \leq s \leq t} u_K^N(s) < \infty.$$

Consequently, the stopping times τ_K^N tend to infinity as K goes to infinity, and Fatou's lemma allows to conclude.

We finally prove (v). Indeed, by Itô's isometry and Jensen's inequality, for all $0 \leq s \leq t \leq T$, using the sub-quadraticity of f and (i),

$$\begin{aligned} \mathbb{E}_x [(X_t^N - X_s^N)^2] &= \mathbb{E}_x \left[\left(-\alpha \int_s^t X_r^N dr + \frac{1}{\sqrt{N}} \sum_{j=1}^N \int_{]s,t] \times \mathbb{R}_+ \times \mathbb{R}} u \mathbf{1}_{\{z \leq f(X_{r-}^N)\}} d\pi_j(r, z, u) \right)^2 \right] \\ &\leq 2\alpha^2(t-s) \int_s^t \mathbb{E}_x [(X_r^N)^2] dr + 2\sigma^2 \int_s^t \mathbb{E}_x [f(X_r^N)] dr \\ &\leq 2\alpha^2 C_t (1+x^2)(t-s) + 2\sigma^2 C_t (1+x^2)(t-s) \\ &\leq C_T (t-s)(1+x^2). \end{aligned}$$

This proves that X^N satisfies hypothesis (v). A similar computation holds true for \bar{X} . \square

5.5. Proof of Proposition 2.4

Proof. 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 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

$$\begin{aligned} \bar{P}_t g(x) &= \mathbb{E} \left[g \left(\bar{X}_t^{(x)} \right) \right], \quad (\bar{P}_t g)'(x) = \mathbb{E} \left[\frac{\partial \bar{X}_t^{(x)}}{\partial x} g' \left(\bar{X}_t^{(x)} \right) \right], \\ (\bar{P}_t g)''(x) &= \mathbb{E} \left[\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} g' \left(\bar{X}_t^{(x)} \right) + \left(\frac{\partial \bar{X}_t^{(x)}}{\partial x} \right)^2 g'' \left(\bar{X}_t^{(x)} \right) \right], \\ (\bar{P}_t g)'''(x) &= \mathbb{E} \left[\frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3} g' \left(\bar{X}_t^{(x)} \right) + 3 \frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} \cdot \frac{\partial \bar{X}_t^{(x)}}{\partial x} g'' \left(\bar{X}_t^{(x)} \right) + \left(\frac{\partial \bar{X}_t^{(x)}}{\partial x} \right)^3 g''' \left(\bar{X}_t^{(x)} \right) \right]. \quad (25) \end{aligned}$$

We start with the representation

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

This implies

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

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

$$M_t = \int_0^t \sigma (\sqrt{f})' (\bar{X}_s^{(x)}) dB_s, \quad (27)$$

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

$$U_t = \exp \left(M_t - \frac{1}{2} \langle M \rangle_t \right). \quad (28)$$

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

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

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

$$\mathbb{E} U_t^p \leq e^{\frac{1}{2} p(p-1) \sigma^2 m_1^2 t}, \quad (29)$$

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

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

Differentiating (26) with respect to x , we obtain

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

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

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

which can be rewritten as

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

with M_t as in (27). Applying Itô'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 $Z_0 = 0$,

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

Using Jensen's inequality, (29) and Burkholder-Davis-Gundy inequality, for all $t \geq 0$,

$$\begin{aligned} \mathbb{E}[Z_t^4] &\leq C \left(\mathbb{E} \left[\left(\int_0^t e^{-\alpha s} U_s (\sqrt{f})^{(2)} (\bar{X}_s^{(x)}) dB_s \right)^4 \right] \right. \\ &\quad \left. + \mathbb{E} \left[\left(\int_0^t e^{-\alpha s} U_s (\sqrt{f})' (\bar{X}_s^{(x)}) (\sqrt{f})^{(2)} (\bar{X}_s^{(x)}) ds \right)^4 \right] \right) \\ &\leq C \left(\mathbb{E} \left[\left(\int_0^t e^{-2\alpha s} U_s^2 (\sqrt{f})^{(2)} (\bar{X}_s^{(x)})^2 ds \right)^2 \right] \right. \\ &\quad \left. + \mathbb{E} \left[\left(\int_0^t e^{-\alpha s} U_s (\sqrt{f})' (\bar{X}_s^{(x)}) (\sqrt{f})^{(2)} (\bar{X}_s^{(x)}) ds \right)^4 \right] \right) \\ &\leq C (t + t^3) \int_0^t e^{-4\alpha s} \mathbb{E}[U_s^4] ds \leq C (t + t^3) \int_0^t e^{(6\sigma^2 m_1^2 - 4\alpha)s} ds \\ &\leq C (t + t^3) (1 + t + e^{(6\sigma^2 m_1^2 - 4\alpha)t}) \leq C (t + t^4) e^{(6\sigma^2 m_1^2 - 4\alpha)t}. \end{aligned} \quad (32)$$

We deduce that

$$\mathbb{E}[V_t^2] \leq \mathbb{E}[Z_t^4]^{1/2} \mathbb{E}[U_t^4]^{1/2} \leq C(t^{1/2} + t^2) e^{3\sigma^2 m_1^2 - 2\alpha t} e^{3\sigma^2 m_1^2 t} \leq C(t^{1/2} + t^2) e^{6\sigma^2 m_1^2 - 2\alpha t},$$

whence

$$\mathbb{E} \left[\left(\frac{\partial^2 \bar{X}_t^{(x)}}{\partial x^2} \right)^2 \right] \leq C(t^{1/2} + t^2) e^{(6\sigma^2 m_1^2 - 4\alpha)t}. \quad (33)$$

Finally, differentiating (31), we get

$$\begin{aligned} \frac{\partial^3 \bar{X}_t^{(x)}}{\partial x^3} &= \sigma \int_0^t e^{-\alpha(t-s)} \left[\frac{\partial^3 \bar{X}_s^{(x)}}{\partial x^3} (\sqrt{f})' (\bar{X}_s^{(x)}) + 3 \frac{\partial^2 \bar{X}_s^{(x)}}{\partial x^2} \frac{\partial \bar{X}_s^{(x)}}{\partial x} (\sqrt{f})^{(2)} (\bar{X}_s^{(x)}) \right. \\ &\quad \left. + \left(\frac{\partial \bar{X}_s^{(x)}}{\partial x} \right)^3 (\sqrt{f})^{(3)} (\bar{X}_s^{(x)}) \right] dB_s. \end{aligned} \quad (34)$$

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

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

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,

$$\begin{aligned} \mathbb{E} \left[(Z'_t)^2 \right] &\leq C(1+t) \int_0^t \mathbb{E} \left[\left(\frac{Y'_s}{U_s} \right)^2 \right] ds \\ &\leq C(1+t) \int_0^t (e^{-2\alpha s} \mathbb{E} [V_s^2] + e^{-4\alpha s} \mathbb{E} [U_s^4]) ds \\ &\leq C(1+t) \int_0^t \left((s^{1/2} + s^2) e^{(6\sigma^2 m_1^2 - 4\alpha)s} + e^{(6\sigma^2 m_1^2 - 4\alpha)s} \right) ds \\ &\leq C(1+t^3) \int_0^t e^{(6\sigma^2 m_1^2 - 4\alpha)s} ds \\ &\leq C(1+t^3)(1+t + e^{(6\sigma^2 m_1^2 - 4\alpha)t}) \leq C(1+t^4) \left(1 + e^{(6\sigma^2 m_1^2 - 4\alpha)t} \right). \end{aligned} \quad (35)$$

As a consequence,

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

implying

$$\mathbb{E} \left[\left| \frac{\partial^3 \bar{X}_t^{(x)}}{\partial^3 x} \right| \right] \leq C(1+t^2) \left(e^{(\frac{1}{2}\sigma^2 m_1^2 - \alpha)t} + e^{(\frac{7}{2}\sigma^2 m_1^2 - 3\alpha)t} \right). \quad (36)$$

Finally, using Cauchy-Schwarz inequality, and inserting (30), (33) and (36) in (25),

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

which proves the first assertion of the proposition. The proof of the second assertion, equation (12), follows similarly. Finally to prove the third assertion, we first study the regularity of the first derivative. Notice that $t \mapsto \frac{\partial \bar{X}_t^{(x)}}{\partial x}$ is almost surely continuous by equation (26). Now take any sequence $t_n \rightarrow t$. By (30), the family of random variables $\left\{ \frac{\partial \bar{X}_{t_n}^{(x)}}{\partial x} g'(\bar{X}_{t_n}^{(x)}), n \geq 1 \right\}$ is uniformly integrable. As a consequence, the second formula in (25) implies that $(\bar{P}_{t_n} g)'(x) \rightarrow (\bar{P}_t g)'(x)$ as $n \rightarrow \infty$, whence the desired continuity. The argument is similar for the second derivative, using (31) and (33). That concludes the proof. \square

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.
- CLINET, S. and YOSHIDA, N. (2017). Statistical inference for ergodic point processes and application to Limit Order Book. *Stochastic Processes and their Applications* **127** 1800–1839.
- 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 BerlinHeidelberg NewYork.
- KALLENBERG, O. (1997). *Foundations of Modern Probability. Probability and Its Applications*. Springer-Verlag, 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. (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).