



**HAL**  
open science

# Fluctuations for mean-field interacting age-dependent Hawkes processes

Julien Chevallier

► **To cite this version:**

Julien Chevallier. Fluctuations for mean-field interacting age-dependent Hawkes processes. 2016. hal-01393373v1

**HAL Id: hal-01393373**

**<https://hal.science/hal-01393373v1>**

Preprint submitted on 7 Nov 2016 (v1), last revised 7 May 2017 (v2)

**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.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# FLUCTUATIONS FOR MEAN-FIELD INTERACTING AGE-DEPENDENT HAWKES PROCESSES

Julien Chevallier\*

Université de Cergy-Pontoise, AGM UMR-CNRS 8088,  
95302 Cergy-Pontoise

## Abstract

The propagation of chaos and associated law of large numbers for mean-field interacting age-dependent Hawkes processes (when the number of processes  $n$  goes to  $+\infty$ ) being granted by the study performed in [9], the aim of the present paper is to prove the resulting functional central limit theorem. It involves the study of a measure-valued process describing the fluctuations (at scale  $n^{-1/2}$ ) of the empirical measure of the ages around its limit value. This fluctuation process is proved to converge towards a limit process characterized by a limit system of stochastic differential equations driven by a Gaussian noise instead of Poisson (which occurs for the law of large numbers limit).

*Keywords:* Hawkes process, central limit theorem, interacting particle systems, stochastic partial differential equation, neural network.

*Mathematical Subject Classification:* 60G55, 60F05, 60G57, 60H15, 92B20

## I Introduction

In the recent years, the self-exciting point process known as the Hawkes process [20] has been used in very diverse areas. First introduced to model earthquake replicas [24] or [31] (ETAS model), it has been used in criminology to model burglary [30], in genomic data analysis to model occurrences of genes [19, 36], in social networks analysis to model viewing or popularity [4, 12], as well as in finance [2, 3]. We refer to [25] or [43] for more extensive reviews on applications of Hawkes processes.

Part of our analysis finds its motivation in the use of Hawkes processes for the modelling in neuroscience. They are used to describe spike trains associated with several neurons (see e.g. [11]). In that case, it is common to consider a multivariate framework : multivariate Hawkes processes consist of multivariate point processes  $(N^1, \dots, N^n)$  whose intensities are respectively given for  $i = 1, \dots, n$  by

$$\lambda_t^i = \Phi \left( \sum_{j=1}^n \int_0^{t-} h_{j \rightarrow i}(t-z) N^j(dz) \right), \quad (1)$$

where  $\Phi : \mathbb{R} \rightarrow \mathbb{R}_+$  is called the intensity function and  $h_{j \rightarrow i}$  is the *interaction function* describing the influence of each point of  $N^j$  on the appearance of a new point onto  $N^i$ , via its intensity  $\lambda^i$ . Notice that we implicitly assume here that there is no

---

\*Corresponding author: e-mail: [julien.chevallier@u-cergy.fr](mailto:julien.chevallier@u-cergy.fr)

influence of the possible points of  $N^j$  that are before time 0.

In the present paper, as in [9], we study a generalization of multivariate Hawkes process by adding an age dependence.

**Definition I.1.** *For any point process  $N$ , we call predictable age process associated with  $N$ , the process defined by*

$$S_{t-} := t - \sup\{T \in N, T < t\} = t - T_{N_{t-}}, \quad \text{for all } t > 0, \quad (2)$$

and extended by continuity in  $t = 0$ . In particular, its value in  $t = 0$  is entirely determined by  $N \cap \mathbb{R}_-$  and is well-defined as soon as there is a point therein.

In comparison with the standard multivariate Hawkes processes (1), we add an age dependence, as it is done in [9], by assuming that the intensity function  $\Phi$  in (1) (which is then denoted by  $\Psi$  to avoid confusion) may also depend on the predictable age process  $(S_{t-}^i)_{t \geq 0}$  associated with the point process  $N^i$ , like for instance

$$\lambda_t^i = \Psi \left( S_{t-}^i, \frac{1}{n} \sum_{j=1}^n \int_0^{t-} h(t-z) N^j(dz) \right). \quad (3)$$

We refer to [9] where the neurobiological motivation for such a form of intensity is given. Under suitable assumptions, it is shown in [9] that a multivariate point process satisfying (3) exists and we call it an age dependent Hawkes process (ADHP). Furthermore, ADHPs are well approximated, when the dimension  $n$  goes to infinity, by i.i.d. *limit* point processes of the McKean-Vlasov type whose stochastic intensity depends on the time  $t$  and on the age [9, Theorem IV.1.]. More precisely, the intensity of the limit process associated with the framework (3), denoted by  $\bar{N}$ , is given by the following implicit formula  $\bar{\lambda}_t = \Psi(\bar{S}_{t-}, \int_0^t h(t-z) \mathbb{E}[\bar{\lambda}_z] dz)$  where  $(\bar{S}_{t-})_{t \geq 0}$  is the predictable age process associated with  $\bar{N}$ .

As usual with McKean-Vlasov dynamics, the asymptotic evolution (when  $n$  goes to infinity) of the distribution of the population at hand can be described as the solution of a nonlinear partial differential equation (PDE). In our case, it is shown that, starting from a density, the distribution of the limit predictable age process  $(\bar{S}_{t-})_{t \geq 0}$ , denoted by  $u_t$ , admits a density for all time  $t \geq 0$  which is furthermore the unique solution of the non-linear system

$$\begin{cases} \frac{\partial u(t, s)}{\partial t} + \frac{\partial u(t, s)}{\partial s} + \Psi(s, X(t)) u(t, s) = 0, \\ u(t, 0) = \int_{s \in \mathbb{R}} \Psi(s, X(t)) u(t, s) ds, \end{cases} \quad (4)$$

with initial condition that  $u(0, \cdot) = u_0$  (the initial density of the age at time 0), where for all  $t \geq 0$ ,  $X(t) = \int_0^t h(t-z) u(z, 0) dz$  [9, Proposition III.8.]. Such a form of PDE system is known either as *age-structure system* or *refractory density equation* or even *von Foerster-McKendrick system*. Here, the age is represented by the variable  $s$ . We refer to [14] for a linear version of (4) and its theoretical connection with the integrate and fire model, and to [17, 32, 33] for analytical studies of (4).

The relation between mean-field age dependent Hawkes processes and the PDE system (4) is completed by the following law of large numbers (consequence of the functional law of large numbers [9, Corollary IV.4.]):

$$\bar{\mu}_{S_{t-}}^n := \frac{1}{n} \sum_{i=1}^n \delta_{S_{t-}^{n,i}} \xrightarrow{n \rightarrow +\infty} u_t. \quad (5)$$

Moreover, the rate of this convergence is at least  $n^{-1/2}$ . In light of this bound obtained on the rate of convergence, the fluctuation process defined, for all  $t \geq 0$ , by  $\eta_t^n = \sqrt{n}(\bar{\mu}_{S_{t-}}^n - u_t)$  is expected to describe, on the right scale, the second order term appearing in the expansion of the mean-field approximation, the first order term being given by the law of large numbers.

The study of the random fluctuations allows to go beyond the first order mean field limit and its main drawback: propagation of chaos. It means independence of the neurons' activities which is unrealistic from the biological viewpoint [40, 15]. Hence, the derivation of the second order term is of great importance regarding neural networks modelling since it gives an approximation of the fluctuations coming from the finiteness of the number of neurons  $n$  (finite size effects) [7, 8, 27]. A partial but promising answer to this problematic is given by highlighting a stochastic partial differential equation system which could be interpreted as an intermediate modelling scale between the microscopic scale given by ADHP and the macroscopic one given by (4).

Following the approach developed in [16], we prove in the present article that the fluctuations satisfy a functional central limit theorem (CLT) in a suitable distributional space: the limit of the normalized fluctuations is described by means of a stochastic differential equation in infinite dimension driven by a Gaussian noise in comparison with the Poisson noise appearing in [9]. To do so, we regard the fluctuation process  $\eta^n$  as taking values in a Hilbert space, namely the dual of some Sobolev space of test functions. The index of regularity of the dual space, in one-to-one correspondence with the regularity of the test functions in the Sobolev space, is prescribed by the tightness property we are able to provide to the sequence  $(\eta^n)_{n \geq 1}$  and by the form of the generator of the limiting McKean-Vlasov dynamics identified in [9]. Let us precise that this generator is the one associated with the renewal dynamics of the system (4) as highlighted by Proposition II.4 given hereafter.

Although the choice of this index of regularity is rather constrained, the choice of the domain supporting the Sobolev space is somewhat larger. Indeed, two options are available, depending on the way we consider the process  $\eta^n$ , either over a finite time horizon, namely  $(\eta_t^n)_{0 \leq t \leq \theta}$  for some  $\theta \geq 0$ , or in infinite horizon, namely  $(\eta_t^n)_{t \geq 0}$ .

In the first case, we may use the fact that there exists a compact  $K_\theta$  (which is growing with  $\theta$ ) such that  $\eta_t^n$  is supported in  $K_\theta$  for all  $t$  in  $[0, \theta]$ . Hence, one could regard, for all  $\theta \geq 0$ , the fluctuation process  $(\eta_t^n)_{0 \leq t \leq \theta}$  as a process with values in the dual of a standard Sobolev space of functions with support in  $K_\theta$ . The main drawback of such an approach is that the space of trajectories within which the CLT takes place depends on the time horizon  $\theta$ . To bypass this issue, one may be willing to work directly on the entire positive time line  $\mathbb{R}_+$ , but then, it is not possible anymore to find a compact subset  $K$  supporting the measures  $\eta_t^n$ , for all  $t \geq 0$ , since  $\cup_{\theta \geq 0} K_\theta = \mathbb{R}_+$ . A convenient strategy to sidestep this fact is to use a Sobolev space supported by the entire  $\mathbb{R}_+$ . Yet, standard Sobolev spaces supported by  $\mathbb{R}_+$  fail to accommodate with our purpose, since, as made clear by the proof below, constant functions are required to belong to the space of test functions. Therefore, instead of a standard Sobolev space, we may use a weighted Sobolev space, provided that the weight satisfies suitable integrability properties.

In order to state our CLT on the whole time interval, the second approach is preferred. Furthermore, the weights of the Sobolev spaces are chosen to be polynomial (see Section IV.1 below). This choice is quite convenient because Sobolev spaces with polynomial weights are well-documented in the literature. In particular,

results on the connection between spaces weighted by different powers, Sobolev embedding theorems and Maurin's theorem, are well-known. It is worth noting that, provided that constant functions can be chosen as test functions, the precise value of the power in the polynomial weight of the Sobolev space does not really matter in our analysis: more generally, a different choice of family of weights would have been possible and, somehow, it would have led to a result equivalent to ours. In this regard, we stress, at the end of the paper, the fact that our result in infinite horizon is in fact equivalent to what we would have obtained by implementing the first of the two approaches mentioned above instead of the second one: roughly speaking, one can recover our result by sticking together the CLTs obtained on each finite interval of the form  $[0, \theta]$ , for  $\theta \geq 0$ ; conversely, one can prove, from our statement, that, on any finite interval  $[0, \theta]$ , the CLT holds true in the dual space of a standard Sobolev space supported by  $K_\theta$ .

The Hilbertian approach used in this article has been already implemented in the diffusion processes framework [16, 23, 26, 28]. Let us mention here what are the main differences between these earlier results and ours:

- Under general non-degeneracy conditions, the marginal laws of a diffusion process are not compactly supported. The unboundedness of the support imposes the choice of weighted Sobolev spaces even in finite time horizon. In this framework, Sobolev spaces with polynomial weights are especially adapted to carry solutions with moments that are finite up to some order only. In that case, the choice of the power in the weight is explicitly prescribed by the maximal order up to which the solution has a finite moment. As already mentioned, this differs from our case: in the present article, the particles (namely, the ages of the neurons) are compactly supported over any finite time interval and thus, have finite moments of any order. Once again, this is the reason why the choice of the power, and more generally of the weight, in the Sobolev space is much larger.
- Unlike point processes, diffusion processes are time continuous. Also, their generator is both local and of second order, whereas the generator for the point process identified in the mean-field limit in [9] is both of the first order and nonlocal. As a first consequence, the indices of regularity of the various Sobolev spaces used in this paper differ from those used in the diffusive framework. Also, the space of trajectories cannot be the same: although the limit process in our CLT has continuous trajectories, we must work with a space of càdlàg functions in order to accommodate with the jumps of the fluctuation process. Surprisingly, jumps do not just affect the choice of the functional space used to state the CLT (namely space of càdlàg versus space of continuous functions) but it also dictates the metric used to estimate the error in the Sznitman coupling between the age-dependent Hawkes process and its mean-field limit (which is also a point process). Indeed, the standard trick used for diffusion processes that consists in getting stronger estimates for the Sznitman coupling by considering  $L^p$ -norms, for  $p > 2$ , is not adapted to point processes. Therefore, we develop a specific approach by providing higher order estimates of the error in the Sznitman coupling in the total variation sense. Up to our knowledge, this argument is completely new.

Let us mention that the fluctuations of jump processes have been the object of previous publications [37, 41]. However, the CLTs are established in the fluid limit, namely small jumps at high frequency so that the jumps vanish at the limit. The

techniques developed in those articles are useless here since the framework of the present article does not fall into the fluid limit framework: in our case, the limit processes are also jump processes.

The present paper is organized as follows. The model is described in Section II. Then, the main estimates required in this work are given in Section III. These can be seen as the extension, to higher orders, of the estimates used in [9] to get the bound  $n^{-1/2}$  on the rate of the convergence (5). These key estimates are used to prove tightness for the distribution  $\eta^n$  in a Hilbert space that is the dual of some weighted Sobolev space. Under regularity assumptions on the intensity function  $\Psi$  and the interaction function  $h$ , we finally prove in Section V.2 the convergence of the fluctuation process which states our CLT. Furthermore, its limit is characterized by a system of stochastic differential equations, driven by a Gaussian process with explicit covariance, and involving an auxiliary process with values in  $\mathbb{R}$  (Theorem V.12). Finally, the CLT is applied to give some justification to a stochastic partial differential equation which can be seen as a better approximation than the PDE system (4) in the mean-field limit.

### General notations

- Statistical distributions are referred to as laws of random variables to avoid confusion with distributions in the analytical sense that are linear forms acting on some test function space.
- The space of bounded functions of class  $\mathcal{C}^k$ , with bounded derivatives of each order less than  $k$  is denoted by  $\mathcal{C}_b^k$ .
- The space of càdlàg (right continuous with left limits) functions is denoted by  $\mathcal{D}$ .
- For  $\mu$  a measure on  $E$  and  $\varphi$  a function on  $E$ , we denote  $\langle \mu, \varphi \rangle := \int_E \varphi(x) \mu(dx)$  when it makes sense.
- If a quantity  $Q$  depends on the time variable  $t$ , then we most often use the notation  $Q_t$  when it is a random process in comparison with  $Q(t)$  when it is a deterministic function.
- We say that the quantity  $Q_n(\sigma)$ , which depends on an integer  $n$  and a parameter  $\sigma \in \mathbb{R}^d$ , is bounded up to a locally bounded function (which does not depend on  $n$ ) by  $f(n)$ , denoted by  $Q_n(\sigma) \lesssim_\sigma f(n)$ , if there exists a locally bounded function  $g : \mathbb{R}^d \rightarrow \mathbb{R}_+$  such that, for all  $n$ ,  $|Q_n(\sigma)| \leq g(\sigma)f(n)$ .
- Throughout this paper,  $C$  denotes a constant that may change from line to line.

## II Definitions and propagation of chaos

In all the sequel, we focus on locally finite point processes,  $N$ , on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  that are random countable sets of points of  $\mathbb{R}$  such that for any bounded measurable set  $A \subset \mathbb{R}$ , the number of points in  $N \cap A$  is finite almost surely (a.s.). The associated points define an ordered sequence  $(T_n)_{n \in \mathbb{Z}}$ . For a measurable set  $A$ ,  $N(A)$  denotes the number of points of  $N$  in  $A$ . We are interested in the behaviour of  $N$  on  $(0, +\infty)$  and we denote  $t \in \mathbb{R}_+ \mapsto N_t := N((0, t])$  the associated counting process. Furthermore, the point measure associated with  $N$  is denoted by  $N(dt)$ . In

particular, for any non-negative measurable function  $f$ ,  $\int_{\mathbb{R}} f(t)N(dt) = \sum_{i \in \mathbb{Z}} f(T_i)$ . For any point process  $N$ , we call *age process* associated with  $N$  the process  $(S_t)_{t \geq 0}$  given by

$$S_t = t - \sup\{T \in N, T \leq t\}, \quad \text{for all } t \geq 0. \quad (6)$$

In comparison with the age process, we call *predictable age process* associated with  $N$  the predictable process  $(S_{t-})_{t \geq 0}$  given by

$$S_{t-} = t - \sup\{T \in N, T < t\}, \quad \text{for all } t > 0, \quad (7)$$

and extended by continuity in  $t = 0$ .

We work on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and suppose that the canonical filtration associated with  $N$ , namely  $(\mathcal{F}_t^N)_{t \geq 0}$  defined by  $\mathcal{F}_t^N := \sigma(N \cap (-\infty, t])$ , is such that for all  $t \geq 0$ ,  $\mathcal{F}_t^N \subset \mathcal{F}_t$ . Let us denote  $\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}$ . We call  $\mathbb{F}$ -(predictable) intensity of  $N$  any non-negative  $\mathbb{F}$ -predictable process  $(\lambda_t)_{t \geq 0}$  such that  $(N_t - \int_0^t \lambda_s ds)_{t \geq 0}$  is an  $\mathbb{F}$ -local martingale. Informally,  $\lambda_t dt$  represents the probability that the process  $N$  has a new point in  $[t, t + dt]$  given  $\mathcal{F}_{t-}$ . Under some assumptions that are supposed here, this intensity process exists, is essentially unique and characterizes the point process (see [6] for more insights). In particular, since  $N$  admits an intensity, for any  $t \geq 0$ , the probability that  $t$  belongs to  $N$  is null. Moreover, notice the following properties satisfied by the age processes:

- the two age processes are equal for all  $t \geq 0$  except the positive times  $T$  in  $N$  (almost surely a set of null measure in  $\mathbb{R}_+$ ),
- for any fixed  $t \geq 0$ ,  $S_{t-} = S_t$  almost surely (since  $N$  admits an intensity),
- and the value  $S_{0-} = S_0$  is entirely determined by  $N \cap \mathbb{R}_-$  and is well-defined as soon as there is a point therein.

The exact behaviour of  $N \cap \mathbb{R}_-$  is not of great interest in the present article. We only assume that there is a point in it almost surely such that  $S_{0-} = S_0$  is well-defined. Furthermore, we assume that the random variable  $S_0$  admits  $u_0$  as a probability density.

## II.1 Parameters and list of assumptions

The definition of an *age dependent Hawkes process* (ADHP) is given bellow, but let us first introduce the parameters of the model:

- a positive integer  $n$  which is the number of particles (e.g. neurons) in the network (for  $i = 1, \dots, n$ ,  $N^i$  represents the occurrences of the events, e.g. spikes, associated with the particle  $i$ );
- a probability density  $u_0$ ;
- an interaction function  $h : \mathbb{R}_+ \rightarrow \mathbb{R}$ ;
- an intensity function  $\Psi : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}_+$ .

For sake of simplicity, all the assumptions made on the parameters are gathered here:

$$(\mathcal{A}_{\infty}^{u_0}): \left| \begin{array}{l} \text{The probability density } u_0 \text{ is uniformly bounded with compact support so that there exists a constant } C > 0 \text{ such that } S_0 \leq C \text{ almost surely (a.s.). The smallest possible constant } C \text{ is denoted by } M_{S_0}. \end{array} \right.$$



$(\mathcal{A}_\infty^h)$ :	The interaction function $h$ is locally bounded. Denote by, for all $t \geq 0$ , $h_\infty(t) := \max_{s \in [0, t]} h(s) < +\infty$ .
$(\mathcal{A}_{\text{Höl}}^h)$ :	There exist two positive constants denoted by $\text{Höl}(h)$ and $\beta(h)$ such that for all $t, s \geq 0$ , $ h(t) - h(s)  \leq \text{Höl}(h) t - s ^{\beta(h)}$ .
$(\mathcal{A}_{y, \mathcal{C}^2}^\Psi)$ :	For all $s \geq 0$ , the function $\Psi_s : y \mapsto \Psi(s, y)$ is of class $\mathcal{C}^2$ . Furthermore, $\ \frac{\partial \Psi}{\partial y}\ _\infty := \sup_{s, y}  \frac{\partial \Psi}{\partial y}(s, y)  < +\infty$ and $\ \frac{\partial^2 \Psi}{\partial y^2}\ _\infty < +\infty$ . The constant $\ \frac{\partial \Psi}{\partial y}\ _\infty$ is denoted by $\text{Lip}(\Psi)$ .
$(\mathcal{A}_\infty^\Psi)$ :	The function $\Psi$ is uniformly bounded, that is $\ \Psi\ _\infty < +\infty$ .
$(\mathcal{A}_{s, \mathcal{C}_b^2}^\Psi)$ :	For all $y$ in $\mathbb{R}$ , the functions $s \mapsto \Psi(s, y)$ and $s \mapsto \frac{\partial \Psi}{\partial y}(s, y)$ respectively belong to $\mathcal{C}_b^2$ and $\mathcal{C}_b^1$ . Furthermore, the functions $y \mapsto \ \Psi(\cdot, y)\ _{\mathcal{C}_b^2}$ and $y \mapsto \ \frac{\partial \Psi}{\partial y}(\cdot, y)\ _{\mathcal{C}_b^1}$ are locally bounded <sup>1</sup> .
$(\mathcal{A}_{s, \mathcal{C}_b^4}^\Psi)$ :	For all $y$ in $\mathbb{R}$ , the function $s \mapsto \Psi(s, y)$ belongs to $\mathcal{C}_b^4$ and $y \mapsto \ \Psi(\cdot, y)\ _{\mathcal{C}_b^4}$ is locally bounded.

**Remark II.1.** *Note that:*

- Assumption  $(\mathcal{A}_{\text{Höl}}^h)$  implies Assumption  $(\mathcal{A}_\infty^h)$ ,
- the assumptions regarding the intensity function  $\Psi$  are rather technical, nevertheless Assumptions  $(\mathcal{A}_{y, \mathcal{C}^2}^\Psi)$ ,  $(\mathcal{A}_\infty^\Psi)$  and  $(\mathcal{A}_{s, \mathcal{C}_b^2}^\Psi)$  are satisfied as soon as  $\Psi$  belongs to  $\mathcal{C}_b^2$ . Furthermore, Assumption  $(\mathcal{A}_{s, \mathcal{C}_b^4}^\Psi)$  is satisfied if  $\Psi$  is in  $\mathcal{C}_b^4$ .

Let  $(\mathcal{A}_{\text{LLN}})$  be satisfied if  $(\mathcal{A}_\infty^{u_0})$ ,  $(\mathcal{A}_\infty^h)$ ,  $(\mathcal{A}_{y, \mathcal{C}^2}^\Psi)$  and  $(\mathcal{A}_\infty^\Psi)$  are satisfied. These four assumptions also appear in [9], where they are used to prove propagation of chaos as stressed below. Furthermore, let  $(\mathcal{A}_{\text{TGN}})$  be satisfied if  $(\mathcal{A}_{\text{LLN}})$  and  $(\mathcal{A}_{s, \mathcal{C}_b^2}^\Psi)$  are satisfied. It is used in the present article to prove tightness of the fluctuations. Finally, let  $(\mathcal{A}_{\text{CLT}})$  be satisfied if  $(\mathcal{A}_{\text{TGN}})$ ,  $(\mathcal{A}_{\text{Höl}}^h)$  and  $(\mathcal{A}_{s, \mathcal{C}_b^4}^\Psi)$  are satisfied. It is used in the present article to prove convergence of the fluctuations.

Notice that Assumption  $(\mathcal{A}_\infty^{u_0})$  implies that the age processes associated with  $N$  are such that, almost surely,

$$\text{for all } t \geq 0, S_t \leq M_{S_0} + t \text{ and } S_{t-} \leq M_{S_0} + t. \quad (8)$$

## II.2 Already known results

Below is given the definition of an ADHP by providing its representation as a system of stochastic differential equations (SDE) driven by Poisson noise.

**Representation II.2.** *Let  $(\Pi^i(dt, dx))_{i \geq 1}$  be some i.i.d.  $\mathbb{F}$ -Poisson measures with intensity 1 on  $\mathbb{R}_+^2$ . Let  $(S_0^i)_{i \geq 1}$  be some i.i.d. random variables distributed according to  $u_0$ .*

*Let  $(N_t^i)_{t \geq 0}^{i=1, \dots, n}$  be a family of counting processes such that, for  $i = 1, \dots, n$ , and all  $t \geq 0$ ,*

$$N_t^i = \int_0^t \int_0^{+\infty} \mathbb{1} \left\{ x \leq \Psi \left( S_{t'-}^i, \frac{1}{n} \sum_{j=1}^n \left( \int_0^{t'-} h(t' - z) N^j(dz) \right) \right) \right\} \Pi^i(dt', dx), \quad (9)$$

*where  $(S_{t-}^i)_{t \geq 0}$  is the predictable age process associated with  $N^i$ . Then,  $(N^i)_{i=1, \dots, n}$  is an age dependent Hawkes process (ADHP) with parameters  $(n, h, \Psi, u_0)$ .*

<sup>1</sup>The definitions of the norms  $\|\cdot\|_{\mathcal{C}_b^k}$ , for all  $k \geq 0$ , can be found in Section IV.1



**Remark II.3.** Note that an ADHP is in fact a (deterministic) measurable function of the Poisson measures  $(\Pi^i(dt, dx))_{i \geq 1}$ . More classically, an ADHP can be characterized by its stochastic intensity (3). Going back and forth between the definition via the intensities (3) and Representation II.2 is standard (see [9, Section II.4] for more insights). Furthermore, [9, Proposition II.4] gives that, under Assumption  $(\mathcal{A}_{LLN})$ , there exists an ADHP  $(N^i)_{i=1, \dots, n}$  with parameters  $(n, h, \Psi, u_0)$  such that  $t \mapsto \mathbb{E}[N_t^1]$  is locally bounded.

Notice that, since the initial conditions  $(S_0^i)_{i=1, \dots, n}$  are i.i.d. and the Poisson measures  $(\Pi^i(dt, dx))_{i \geq 1}$  are i.i.d., the processes  $N^i$ ,  $i = 1, \dots, n$ , defined by (9) are exchangeable.

Here, we give a brief overview of the results obtained in [9] in order to set the context of the present article. We expect ADHPs to be well approximated, when  $n$  goes to infinity, by i.i.d. solutions of the following *limit equation*,

$$\forall t > 0, \bar{N}_t = \int_0^t \int_0^{+\infty} \mathbb{1}_{\left\{x \leq \Psi\left(\bar{S}_{t'-}, \int_0^{t'-} h(t' - z) \mathbb{E}[\bar{N}(dz)]\right)\right\}} \Pi(dt', dx), \quad (10)$$

where  $\Pi(dt', dx)$  is an  $\mathbb{F}$ -Poisson measure on  $\mathbb{R}_+^2$  with intensity 1 and  $(\bar{S}_{t-})_{t \geq 0}$  is the predictable age process associated with  $\bar{N}$  where  $\bar{S}_0$  is distributed according to  $u_0$ .

Under Assumption  $(\mathcal{A}_{LLN})$ , [9, Proposition III.6] states existence and uniqueness of the limit process  $\bar{N}$ . In particular, there exists a continuous function  $\bar{\lambda} : \mathbb{R}_+ \rightarrow \mathbb{R}$  (which depends on the parameters  $h, \Psi$  and  $u_0$ ) such that if  $(\bar{N}_t)_{t \geq 0}$  is a solution of (10) then  $\mathbb{E}[\bar{N}(dt)] = \bar{\lambda}(t)dt$ . Let us define the deterministic function  $\bar{\gamma}$  by, for all  $t \geq 0$ ,

$$\bar{\gamma}(t) := \int_0^t h(t - z) \bar{\lambda}(z) dz. \quad (11)$$

Notice that  $\bar{\gamma}(t')$  is the integral term  $\int_0^{t'-} h(t' - z) \mathbb{E}[\bar{N}(dz)]$  appearing in (10).

Furthermore, the limit predictable age process  $(\bar{S}_{t-})_{t \geq 0}$  is closely related to the PDE system (4).

**Proposition II.4** ([9, Proposition III.8]). *Under Assumption  $(\mathcal{A}_{LLN})$ , the unique solution  $u$  to the system (4) with initial condition that  $u_0$  is such that  $u(t, \cdot)$  is the density of the age  $\bar{S}_{t-}$  (or  $\bar{S}_t$  since they are equal a.s.).*

Once the limit equation is well-posed, following the ideas of Sznitman in [39], it is easy to construct a suitable coupling between ADHPs and i.i.d. solutions of the limit equation (10). More precisely, consider

- a sequence  $(S_0^i)_{i \geq 1}$  of i.i.d. random variables distributed according to  $u_0$ ;
- a sequence  $(\Pi^i(dt', dx))_{i \geq 1}$  of i.i.d.  $\mathbb{F}$ -Poisson measures with intensity 1 on  $\mathbb{R}_+^2$ .

Under Assumption  $(\mathcal{A}_{LLN})$ , we have existence of both ADHPs and the limit process  $\bar{N}$ . Hence, one can build simultaneously:

- a sequence (indexed by  $n \geq 1$ )  $(N^{n,i})_{i=1, \dots, n}$  of ADHPs with parameters  $(n, h, \Psi, u_0)$  according to Representation II.2 namely

$$N_t^{n,i} = \int_0^t \int_0^{+\infty} \mathbb{1}_{\left\{x \leq \Psi\left(S_{t'-}^{n,i}, \gamma_{t'}^n\right)\right\}} \Pi^i(dt', dx) \quad (12)$$

where  $S_0^{n,i} = S_0^i$  and  $\gamma_{t'}^n := n^{-1} \sum_{j=1}^n \int_0^{t'-} h(t' - z) N^{n,j}(dz)$ ,

- and a sequence  $(\bar{N}_t^i)_{t \geq 0}^{i \geq 1}$  of i.i.d. solutions of the limit equation namely

$$\bar{N}_t^i = \int_0^t \int_0^{+\infty} \mathbb{1}_{\{x \leq \Psi(\bar{S}_{t'-}, \bar{\gamma}(t'))\}} \Pi^i(dt', dx), \quad (13)$$

where  $\bar{S}_0^i = S_0^i$  and  $\bar{\gamma}$  is defined by (11).

Moreover, denote by  $\lambda_t^{n,i} := \Psi(S_{t-}^{n,i}, \gamma_t^n)$  and  $\bar{\lambda}_t^i := \Psi(\bar{S}_{t-}^i, \bar{\gamma}(t))$  the respective intensities of  $N^{n,i}$  and  $\bar{N}^i$ .

**Remark II.5.** Notice that the coupling above is based on the sharing of common initial conditions  $(S_0^i)_{i \geq 1}$  and a common underlying randomness, that are the  $\mathbb{F}$ -Poisson measures  $(\Pi^i(dt', dx))_{i \geq 1}$ . Note also that the sequence of ADHPs is indexed by the size of the network  $n$  whereas the solutions of the limit equation which represent the behaviour under the mean field approximation are not.

Then, standard computations mainly based on Grönwall lemma lead to the following estimates [9, Corollary IV.3]: for all  $i = 1, \dots, n$  and  $\theta > 0$ ,

$$\mathbb{E} \left[ \sup_{t \in [0, \theta]} |S_{t-}^{n,i} - \bar{S}_{t-}^i| \right] \lesssim_{\theta} \mathbb{P} \left( (S_{t-}^{n,i})_{t \in [0, \theta]} \neq (\bar{S}_{t-}^i)_{t \in [0, \theta]} \right) \lesssim_{\theta} n^{-1/2}. \quad (14)$$

Finally, these estimates ensure the propagation of chaos property<sup>2</sup> [9, Corollary IV.4] and, in particular, the convergence (as  $n \rightarrow +\infty$ ) of the empirical measure  $\bar{\mu}_{S_t}^n := \frac{1}{n} \sum_{i=1}^n \delta_{S_t^{n,i}}$  towards the law of  $\bar{S}_t^1$  for all  $t \geq 0$ .

### II.3 What next ? The purpose of the present paper

As a straight follow-up to the convergence of the empirical measure  $\bar{\mu}_{S_t}^n$ , we are interested in the dynamics of the fluctuations of this empirical measure around its limit. For any  $t \geq 0$ ,  $\bar{S}_t^1$  and  $\bar{S}_{t-}^1$  have the same probability law since they are equal almost surely. Furthermore, this law, denoted by  $u_t$  admits the density  $u(t, \cdot)$  with respect to the Lebesgue measure, where  $u$  is the unique solution of (4) according to Proposition II.4, thus

$$\langle u_t, \varphi \rangle = \int_0^{+\infty} \varphi(s) u(t, s) ds.$$

The analysis of the coupling (Equation (14)) gives a rate of convergence at least in  $n^{-1/2}$  so we want to find the limit law of the fluctuation process defined, for all  $t \geq 0$ , by

$$\eta_t^n := \sqrt{n} (\bar{\mu}_{S_t}^n - u_t). \quad (15)$$

Notice that  $\eta_t^n$  is a distribution in the functional analysis sense on the state space of the ages, i.e.  $\mathbb{R}_+$ , and is devoted to be considered as a linear form acting on test functions  $\varphi$  by means of  $\langle \eta_t^n, \varphi \rangle$ .

## III Estimates in total variation norm

The bound  $(n^{-1/2})$  on the rate of convergence, given by (14), is not sufficient in order to prove convergence or even tightness of the fluctuation process  $\eta^n$ . Some refined estimates are necessary. For instance, when dealing with diffusions, one looks for higher order moment estimates on the difference between the particles driven by

<sup>2</sup>For any fixed integer  $k$ , the processes  $(S_t^{n,1})_{t \geq 0}, \dots, (S_t^{n,k})_{t \geq 0}$  are asymptotically independent.

the real dynamics and the limit particles (see [16, 23, 26, 28] for instance). Here, we deal with pure jump processes and, up to our knowledge, there is no reason why one could obtain better rates for higher order moments. A simple way to catch this fact is by looking at the coupling between the counting processes. Indeed, the difference between two counting processes, say  $\delta_t^{n,i} = |N_t^{n,i} - \bar{N}_t^i|$ , takes value in  $\mathbb{N}$  so that for all  $p \geq 1$ ,  $(\delta_t^{n,i})^p \geq \delta_t^{n,i}$ , and the moment of order  $p$  is greater than the moment of order one.

In order to accommodate this fact, the key idea is to estimate the coupling (12)-(13) in the total variation distance. Hence, the estimates needed in the next section (and proved in the present section) are the analogues of higher order moments but with respect to the total variation norm, i.e. the probabilities

$$\begin{aligned} \chi_n^{(k)}(\theta) &:= \mathbb{P} \left( (S_{t-}^{n,k'})_{t \in [0, \theta]} \neq (\bar{S}_{t-}^{k'})_{t \in [0, \theta]} \text{ for every } k' = 1, \dots, k \right) \\ &= \mathbb{P} \left( (S_t^{n,k'})_{t \in [0, \theta]} \neq (\bar{S}_t^{k'})_{t \in [0, \theta]} \text{ for every } k' = 1, \dots, k \right), \end{aligned} \quad (16)$$

for all positive integer  $k$  and real number  $\theta \geq 0$ .

The heuristics underlying the result stated below, in Proposition III.1, relies on the asymptotic independence between the  $k$  age processes  $(S_{t-}^{n,k'})_{t \in [0, \theta]}$ ,  $k' = 1, \dots, k$ . Indeed, if they were independent then we would have (remind (14)),

$$\chi_n^{(k)}(\theta) = \prod_{k'=1}^k \mathbb{P}((S_{t-}^{n,k'})_{t \in [0, \theta]} \neq (\bar{S}_{t-}^{k'})_{t \in [0, \theta]}) = (\chi_n^{(1)}(\theta))^k \lesssim_{\theta} n^{-k/2},$$

which is exactly the rate of convergence we find below.

**Proposition III.1.** *Under Assumption ( $\mathcal{A}_{LLN}$ ),*

$$\chi_n^{(k)}(\theta) \lesssim_{(\theta, k)} n^{-k/2} \quad \text{and} \quad \xi_n^{(k)}(t) := \mathbb{E} \left[ |\gamma_t^n - \bar{\gamma}(t)|^k \right] \lesssim_{(t, k)} n^{-k/2}.$$

**Remark III.2.** *In addition to the explanation given in the beginning of this section, let us mention that the analogous to the higher moment estimates obtained for diffusions is obtained here for the difference between  $\gamma_t^n$  and  $\bar{\gamma}(t)$ . Indeed, as  $k$  grows, the convergence of  $\xi_n^{(k)}(t)$  quickens. However, this gain in the rate of convergence does not apply when looking at the difference between the ages  $S_t^{n,1}$  and  $\bar{S}_t^1$  or the difference between the intensities  $\lambda_t^{n,1}$  and  $\bar{\lambda}_t^1$  (except if  $\Psi$  does not depend on the age  $s$ ).*

*Proof.* The core of this proof lies on a trick using the exchangeability of the processes in order to obtain Grönwall-type inequalities involving  $\chi_n^{(k)}$  and  $\xi_n^{(k)}$ .

Denote by  $A \Delta B$  the symmetric difference of the sets  $A$  and  $B$ . Then, for any  $i \leq n$ , let us define  $\Delta^{n,i} := N^{n,i} \Delta \bar{N}^i$  that is the set of points that are not common to  $N^{n,i}$  and  $\bar{N}^i$ . From (12)-(13), one has

$$\Delta_t^{n,i} = \int_0^t \int_0^{+\infty} \mathbb{1}_{\left\{ x \in [[\lambda_{t'}^{n,i}, \bar{\lambda}_{t'}^i]] \right\}} \Pi^i(dt', dx),$$

where  $[[\lambda_{t'}^{n,i}, \bar{\lambda}_{t'}^i]]$  is the non empty interval which is either  $[\lambda_{t'}^{n,i}, \bar{\lambda}_{t'}^i]$  or  $[\bar{\lambda}_{t'}^i, \lambda_{t'}^{n,i}]$ . Then, the intensity of the point process  $\Delta^{n,i}$  is given by  $\lambda_t^{\Delta, n, i} := |\lambda_t^{n,i} - \bar{\lambda}_t^i|$ .

Note that, for all  $n \geq 1$  and  $i = 1, \dots, n$ ,  $S_{0-}^{n,i} = \bar{S}_{0-}^i$  so that the equality between the processes  $(S_{t-}^{n,i})_{t \in [0, \theta]}$  and  $(\bar{S}_{t-}^i)_{t \in [0, \theta]}$  is equivalent to  $\Delta_{\theta-}^{n,i} = 0$ . In particular, one has

$$\chi_n^{(k)}(\theta) \leq \mathbb{E} \left[ \prod_{i=1}^k \Delta_{\theta-}^{n,i} \right], \quad (17)$$

since counting processes take value in  $\mathbb{N}$ . For any positive integers  $k$  and  $p$ , let us denote, for all  $n \geq k$ ,

$$\varepsilon_n^{(k,p)}(\theta) := \mathbb{E} \left[ \prod_{i=1}^k \left( \Delta_{\theta-}^{n,i} \right)^p \right].$$

Let us show, by induction on  $k$ , that

$$\varepsilon_n^{(k,p)}(\theta) \lesssim_{(\theta,k,p)} n^{-k/2} \quad (18)$$

which will end the proof thanks to (17). First, note that the case  $k = 1$  and  $p = 1$  is already treated. Indeed, [9, Theorem IV.1] gives

$$\varepsilon_n^{(1,1)}(\theta) = \int_0^\theta \mathbb{E} \left[ |\lambda_t^{n,1} - \bar{\lambda}_t^1| \right] dt \lesssim_\theta n^{-1/2}. \quad (19)$$

Then, note that for any two positive integers  $p$  and  $q$ ,

$$\varepsilon_n^{(k,p)}(\theta) \leq \varepsilon_n^{(k,q)}(\theta) \text{ as soon as } p \leq q. \quad (20)$$

This is due to the fact that counting processes take value in  $\mathbb{N}$ . The rest of the proof is divided in two steps: initialization and inductive step.

**Step one.** For  $k = 1$  and  $p$  a positive integer, it holds that

$$\left( \Delta_{\theta-}^{n,1} \right)^p = \sum_{p'=0}^{p-1} \binom{p}{p'} \int_0^{\theta-} \left( \Delta_{t-}^{n,1} \right)^{p'} \Delta^{n,1}(dt). \quad (21)$$

Indeed, each time the process  $(\Delta_t^{n,1})_{t \geq 0}$  jumps (from  $\Delta_{t-}^{n,1}$  to  $\Delta_{t-}^{n,1} + 1$ ) then  $(\Delta_{t-}^{n,1})^p$  jumps from  $(\Delta_{t-}^{n,1})^p$  to  $(\Delta_{t-}^{n,1} + 1)^p$  so the infinitesimal variation is

$$\left( \Delta_{t-}^{n,1} + 1 \right)^p - \left( \Delta_{t-}^{n,1} \right)^p = \sum_{p'=0}^{p-1} \binom{p}{p'} \left( \Delta_{t-}^{n,1} \right)^{p'}.$$

The right-hand side of (21) involves integrals of predictable processes, that are the  $(\Delta_{t-}^{n,1})^{p'}$ , with respect to a point measure under which it is convenient to take expectation.

More precisely, since  $(\Delta_{t-}^{n,1})^{p'} \leq (\Delta_{t-}^{n,1})^p$  as soon as  $0 < p' \leq p - 1$ , it holds that

$$\begin{aligned} \varepsilon_n^{(1,p)}(\theta) = \mathbb{E} \left[ \left( \Delta_{\theta-}^{n,1} \right)^p \right] &\leq \mathbb{E} \left[ \int_0^\theta \Delta^{n,1}(dt) \right] + 2^p \mathbb{E} \left[ \int_0^\theta \left( \Delta_{t-}^{n,1} \right)^p \Delta^{n,1}(dt) \right] \\ &\leq \varepsilon_n^{(1,1)}(\theta) + 2^p \int_0^\theta \mathbb{E} \left[ \left( \Delta_{t-}^{n,1} \right)^p \lambda_t^{\Delta, n, 1} \right] dt. \end{aligned} \quad (22)$$

Yet the intensity  $\lambda_t^{\Delta, n, 1}$  is bounded by  $\|\Psi\|_\infty$  and  $\varepsilon_n^{(1,1)}(\theta) \lesssim_\theta n^{-1/2}$ , see (19), so

$$\varepsilon_n^{(1,p)}(\theta) \lesssim_{(\theta,p)} n^{-1/2} + \int_0^\theta \varepsilon_n^{(1,p)}(t) dt,$$

and Lemma B.1 gives  $\varepsilon_n^{(1,p)}(\theta) \lesssim_{(\theta,p)} n^{-1/2}$ .

**Step two.** For all integers  $k \geq 2$  and  $p \geq 1$ , one can generalize the argument used to prove (21) in order to end up with

$$\prod_{i=1}^k (\Delta_{\theta^-}^{n,i})^p = \sum_{j=1}^k \sum_{p'=0}^{p-1} \binom{p}{p'} \int_0^{\theta^-} \prod_{i \neq j, i=1}^k (\Delta_{t^-}^{n,i})^p (\Delta_{t^-}^{n,j})^{p'} \Delta^{n,j}(dt), \quad \text{almost surely.}$$

Hence, thanks to the exchangeability of the processes  $(\Delta^{n,i})_{i=1,\dots,n}$  and the predictability of the integrated processes, we have

$$\begin{aligned} \varepsilon_n^{(k,p)}(\theta) &= \sum_{j=1}^k \sum_{p'=0}^{p-1} \binom{p}{p'} \mathbb{E} \left[ \int_0^\theta \prod_{i \neq j, i=1}^k (\Delta_{t^-}^{n,i})^p (\Delta_{t^-}^{n,j})^{p'} \Delta^{n,j}(dt) \right] \\ &= k \sum_{p'=0}^{p-1} \binom{p}{p'} \int_0^\theta \mathbb{E} \left[ (\Delta_{t^-}^{n,1})^{p'} \prod_{i=2}^k (\Delta_{t^-}^{n,i})^p \lambda_t^{\Delta,n,1} \right] dt \\ &\leq k \int_0^\theta \mathbb{E} \left[ \prod_{i=2}^k (\Delta_{t^-}^{n,i})^p \lambda_t^{\Delta,n,1} \right] + 2^p \mathbb{E} \left[ (\Delta_{t^-}^{n,1})^p \prod_{i=2}^k (\Delta_{t^-}^{n,i})^p \lambda_t^{\Delta,n,1} \right] dt, \end{aligned} \quad (23)$$

where we used that  $(\Delta_{t^-}^{n,1})^{p'} \leq (\Delta_{t^-}^{n,1})^p$  as soon as  $0 < p' \leq p-1$ .

On the one hand, using that  $\lambda_t^{\Delta,n,1} \leq \|\Psi\|_\infty$ , the second expectation in (23) is bounded by  $\|\Psi\|_\infty \varepsilon_n^{(k,p)}(t)$ .

On the other hand, we use  $(\mathcal{A}_{y,C^2}^\Psi)$  which gives the following bound on the intensity,

$$\lambda_t^{\Delta,n,1} \leq \text{Lip}(\Psi) |\gamma_t^n - \bar{\gamma}(t)| + \|\Psi\|_\infty \mathbb{1}_{S_{t^-}^{n,1} \neq \bar{S}_{t^-}^1} \leq \text{Lip}(\Psi) |\gamma_t^n - \bar{\gamma}(t)| + \|\Psi\|_\infty (\Delta_{t^-}^{n,1})^p.$$

Hence the first expectation in (23) is bounded by

$$\text{Lip}(\Psi) D(t) + \|\Psi\|_\infty \varepsilon_n^{(k,p)}(t), \quad (24)$$

with  $D(t) := \mathbb{E}[\prod_{i=2}^k (\Delta_{t^-}^{n,i})^p |\gamma_t^n - \bar{\gamma}(t)|]$ . The second term of (24) is convenient to use a Grönwall-type lemma. To deal with the first term, we use a trick involving the exchangeability of the particles. Indeed, using the exchangeability we can replace each of the  $k-1$  terms  $(\Delta_{t^-}^{n,i})^p$  in the expression of  $D(t)$  by the following sum

$$\frac{1}{\lfloor \frac{n}{k} \rfloor} \sum_{j_i=(i-1)\lfloor \frac{n}{k} \rfloor + 1}^{i\lfloor \frac{n}{k} \rfloor} (\Delta_{t^-}^{n,j_i})^p$$

without modifying the value of the expectation since the sums are taken on disjointed indices. Hence, using for the second line a generalization of Hölder's inequality with  $k$  exponents equal to  $1/k$ , we have

$$\begin{aligned} D(t) &\leq \mathbb{E} \left[ \prod_{i=2}^k \left( \frac{1}{\lfloor \frac{n}{k} \rfloor} \sum_{j_i=(i-1)\lfloor \frac{n}{k} \rfloor + 1}^{i\lfloor \frac{n}{k} \rfloor} (\Delta_{t^-}^{n,j_i})^p \right) |\gamma_t^n - \bar{\gamma}(t)| \right] \\ &\leq \left( \prod_{i=2}^k \mathbb{E} \left[ \left( \frac{1}{\lfloor \frac{n}{k} \rfloor} \sum_{j=1}^{\lfloor \frac{n}{k} \rfloor} (\Delta_{t^-}^{n,j})^p \right)^k \right]^{1/k} \right) \xi_n^{(k)}(t)^{1/k} \leq E_{n,k,p}(t)^{\frac{k-1}{k}} \xi_n^{(k)}(t)^{1/k}, \end{aligned} \quad (25)$$

with  $E_{n,k,p}(t) := \mathbb{E}[(1/\lfloor \frac{n}{k} \rfloor) \sum_{j=1}^{\lfloor \frac{n}{k} \rfloor} (\Delta_{t-}^{n,j,p})^k]$ . Yet, computations given in Section A.1 give the two following statements: there exists a constant  $C(k)$  which does not depend on  $n$  or  $p$  such that

$$E_{n,k,p}(t) \leq C(k) \left( \sum_{k'=1}^{k-1} n^{k'-k} \varepsilon_n^{(k',pk)}(t) + \varepsilon_n^{(k,p)}(t) \right), \quad (26)$$

and  $\xi_n^{(k)}(t)$  satisfy the following bound,

$$\xi_n^{(k)}(t) \lesssim_{(t,k)} n^{-k/2} + \sum_{k'=1}^{k-1} n^{k'-k} \varepsilon_n^{(k',k)}(t) + \varepsilon_n^{(k,1)}(t). \quad (27)$$

Then, using the induction hypothesis (18), that is for all  $1 \leq k' \leq k-1$  and for all positive integer  $p$ ,  $\varepsilon_n^{(k',p)}(t) \lesssim_{(t,k,p)} n^{-k'/2}$ , one has,

$$\begin{cases} E_{n,k,p}(t) \lesssim_{(t,k,p)} \sum_{k'=1}^{k-1} n^{k'-k} n^{-k'/2} + \varepsilon_n^{(k,p)}(t) \lesssim_{(t,k,p)} n^{-(k+1)/2} + \varepsilon_n^{(k,p)}(t) \\ \xi_n^{(k)}(t) \lesssim_{(t,k,p)} n^{-k/2} + \sum_{k'=1}^{k-1} n^{k'-k} n^{-k'/2} + \varepsilon_n^{(k,1)}(t) \lesssim_{(t,k,p)} n^{-k/2} + \varepsilon_n^{(k,1)}(t). \end{cases} \quad (28)$$

Gathering (23), (24), (25) and (28) gives (remind that  $\varepsilon_n^{(k,1)}(t) \leq \varepsilon_n^{(k,p)}(t)$ )

$$\varepsilon_n^{(k,p)}(\theta) \lesssim_{(\theta,k,p)} n^{-k/2} + \int_0^\theta \varepsilon_n^{(k,p)}(t) dt,$$

and so the Grönwall-type Lemma B.1 gives  $\varepsilon_n^{(k,p)}(\theta) \lesssim_{(\theta,k,p)} n^{-k/2}$  which ends the proof thanks to (17).  $\square$

## IV Tightness

The aim of this section is to prove tightness of the sequence of the laws of  $(\eta^n)_{n \geq 1}$  regarded as stochastic processes (in time) with values in a suitable space of distributions. Thus, we consider  $(\eta_t^n)_{t \geq 0}$  as a random process with values in the dual space of some well-chosen space of test functions. In Section IV.1, we give the definition of these spaces of test functions. Following the Hilbertian approach developed in [16], we work with weighted Sobolev Hilbert spaces. Finally, the tightness result is stated in Theorem IV.11.

The following study takes benefit of the Hilbert structure of the Sobolev spaces considered. Let us state here the Aldous tightness criterion for Hilbert space valued stochastic processes (cf. [22, p. 34-35]) used in the present paper. Let  $H$  be a separable Hilbert space. A sequence of processes  $(X^n)_{n \geq 1}$  in  $\mathcal{D}(\mathbb{R}_+, H)$  defined on the respective filtered probability spaces  $(\Omega^n, \mathcal{F}^n, (\mathcal{F}_t^n)_{t \geq 0}, \mathbb{P}^n)$  is tight if both conditions below hold true:

- |                    |   |
|--------------------|---|
| (A <sub>1</sub> ): | for every $t \geq 0$ and $\varepsilon > 0$ , there exists a compact set $K \subset H$ such that $\sup_{n \geq 1} \mathbb{P}^n (X_t^n \notin K) \leq \varepsilon,$   |
| (A <sub>2</sub> ): | for every $\varepsilon_1, \varepsilon_2 > 0$ and $\theta \geq 0$ , there exists $\delta_0 > 0$ and an integer $n_0$ such that for all $(\mathcal{F}_t^n)_{t \geq 0}$ -stopping time $\tau_n \leq \theta$ , $\sup_{n \geq n_0} \sup_{\delta \leq \delta_0} \mathbb{P}^n (\ X_{\tau_n + \delta}^n - X_{\tau_n}^n\ _H \geq \varepsilon_1) \leq \varepsilon_2.$ |

Note that  $(\mathbf{A}_1)$  is implied by the condition  $(\mathbf{A}_{1'})$  stated below which is much easier to ensue.

$$(\mathbf{A}_{1'}): \left\{ \begin{array}{l} \text{There exists a Hilbert space } H_0 \text{ such that } H_0 \hookrightarrow_K H \text{ and, for all } \\ t \geq 0, \\ \sup_{n \geq 1} \mathbb{E}^n [\|X_t^n\|_{H_0}^2] < +\infty, \\ \text{where the notation } \hookrightarrow_K \text{ means that the embedding is compact and} \\ \mathbb{E}^n \text{ denotes the expectation associated with the probability } \mathbb{P}^n. \end{array} \right.$$

The fact that  $(\mathbf{A}_{1'})$  implies  $(\mathbf{A}_1)$  is easily checked: by compactness of the embedding, closed balls in  $H_0$  are compact in  $H$  so, Markov's inequality gives  $(\mathbf{A}_1)$ .

#### IV.1 Preliminaries on weighted Sobolev spaces

Here are listed some definitions and technical results about the weighted Sobolev spaces used in the present article. To avoid confusion, let us stress the fact that the test functions we use are supported in the state space of the ages, namely  $\mathbb{R}_+$ . For any integer  $k$  and any real  $\alpha$  in  $\mathbb{R}_+$ , we denote by  $\mathcal{W}_0^{k,\alpha} := \mathcal{W}_0^{k,\alpha}(\mathbb{R}_+)$  the completion of the set of compactly supported (in  $\mathbb{R}_+$ ) functions of class  $\mathcal{C}^\infty$  for the following norm

$$\|f\|_{k,\alpha} := \left( \sum_{k'=0}^k \int_{\mathbb{R}_+} \frac{|f^{(k')}(x)|^2}{1+|x|^{2\alpha}} dx \right)^{1/2},$$

where  $f^{(j)}$  denotes the  $j^{\text{th}}$  derivative of  $f$ . Then,  $\mathcal{W}_0^{k,\alpha}$  equipped with the norm  $\|\cdot\|_{k,\alpha}$  is a separable Hilbert space and we denote  $(\mathcal{W}_0^{-k,\alpha}, \|\cdot\|_{-k,\alpha})$  its dual space. Notice that

$$\begin{cases} \text{if } k' \geq k, \text{ then } \|\cdot\|_{k,\alpha} \leq \|\cdot\|_{k',\alpha} \text{ and } \|\cdot\|_{-k',\alpha} \leq \|\cdot\|_{-k,\alpha}, \\ \text{if } \alpha' \geq \alpha, \text{ then } \mathcal{W}_0^{k,\alpha} \hookrightarrow \mathcal{W}_0^{k,\alpha'} \text{ and } \mathcal{W}_0^{-k,\alpha'} \hookrightarrow \mathcal{W}_0^{-k,\alpha}, \end{cases} \quad (29)$$

where the notation  $\hookrightarrow$  means that the embedding is continuous.

Let  $\mathcal{C}^{k,\alpha}$  be the space of functions  $f$  on  $\mathbb{R}_+$  with continuous derivatives up to order  $k$  such that, for all  $k' \leq k$ ,  $\sup_{x \in \mathbb{R}_+} |f^{(k')}(x)|/(1+|x|^\alpha) < +\infty$ . We equip this space with the norm

$$\|f\|_{\mathcal{C}^{k,\alpha}} := \sum_{k'=0}^k \sup_{x \in \mathbb{R}_+} \frac{|f^{(k')}(x)|}{1+|x|^\alpha}.$$

Recall that  $\mathcal{C}_b^k$  is the space of bounded functions of class  $\mathcal{C}^k$  with bounded derivatives of every order less than  $k$ . Notice that  $\mathcal{C}_b^k = \mathcal{C}^{k,0}$  as normed spaces. Denote by  $\mathcal{C}_b^{-k}$  its dual space. For any  $\alpha > 1/2$  and any integer  $k$  (so that  $\int_{\mathbb{R}_+} 1/(1+|x|^{2\alpha}) dx < +\infty$ ), we have  $\mathcal{C}_b^k \hookrightarrow \mathcal{W}_0^{k,\alpha}$ , i.e. there exists a constant  $C$  such that

$$\|\cdot\|_{k,\alpha} \leq C \|\cdot\|_{\mathcal{C}_b^k}. \quad (30)$$

We recall the following Sobolev embeddings (see [16, Section 2.1]):

- (i) Sobolev embedding theorem:  $\mathcal{W}_0^{m+k,\alpha} \hookrightarrow \mathcal{C}^{k,\alpha}$  for  $m \geq 1$ ,  $k \geq 0$  and  $\alpha$  in  $\mathbb{R}_+$ , i.e. there exists a constant  $C$  such that

$$\|f\|_{\mathcal{C}^{k,\alpha}} \leq C \|f\|_{m+k,\alpha}. \quad (31)$$

- (ii) Maurin's theorem:  $\mathcal{W}_0^{m+k,\alpha} \hookrightarrow_{H.S.} \mathcal{W}_0^{k,\alpha+\beta}$  for  $m \geq 1$ ,  $k \geq 0$ ,  $\alpha$  in  $\mathbb{R}_+$  and  $\beta > 1/2$ , where *H.S.* means that the embedding is of Hilbert-Schmidt type<sup>3</sup>.

<sup>3</sup>Here, it means that  $\sum_{j \geq 1} \|\varphi_j\|_{k,\alpha+\beta}^2 < +\infty$  if  $(\varphi_j)_{j \geq 1}$  is an orthonormal basis of  $\mathcal{W}_0^{m+k,\alpha}$ .



In particular, the embedding is compact and there exists a constant  $C$  such that

$$\|f\|_{k,\alpha+\beta} \leq C\|f\|_{k+m,\alpha}. \quad (32)$$

Hence, the following dual embeddings hold true:

$$\begin{cases} \mathcal{W}_0^{-k,\alpha} \hookrightarrow \mathcal{C}_b^{-k}, \text{ for } k \geq 0 \text{ and } \alpha > 1/2, \text{ (dual embedding of (30))} \\ \mathcal{W}_0^{-k,\alpha+\beta} \hookrightarrow_{H.S.} \mathcal{W}_0^{-(m+k),\alpha}, \text{ for } m \geq 1, k \geq 0, \alpha \text{ in } \mathbb{R}_+ \text{ and } \beta > 1/2. \end{cases} \quad (33)$$

In some of the proofs given in the next section, we consider an orthonormal basis  $(\varphi_j)_{j \geq 1}$  of  $\mathcal{W}_0^{k,\alpha}$  composed of  $\mathcal{C}^\infty$  functions with compact support. The existence of such a basis follows from the fact that the functions of class  $\mathcal{C}^\infty$  with compact support are dense in  $\mathcal{W}_0^{k,\alpha}$ . Furthermore, if  $(\varphi_j)_{j \geq 1}$  is an orthonormal basis of  $\mathcal{W}_0^{k,\alpha}$  and  $w$  belongs to  $\mathcal{W}_0^{-k,\alpha}$ , then  $\|w\|_{-k,\alpha}^2 = \sum_{j \geq 1} \langle w, \varphi_j \rangle^2$  thanks to Parseval's identity. Let us precise that we stick with the notation  $(\varphi_j)_{j \geq 1}$  even if the space  $\mathcal{W}_0^{k,\alpha}$  (in particular the regularity  $k$ ) may differ from page to page.

The three lemmas below are useful throughout the analysis.

**Lemma IV.1.** *For every test function  $\varphi$  in  $\mathcal{W}_0^{2,\alpha}$ ,  $\|\varphi'\|_{1,\alpha} \leq \|\varphi\|_{2,\alpha}$ . If  $f$  belongs to  $\mathcal{C}_b^k$  for some  $k \geq 1$  then, for any fixed  $\alpha$  in  $\mathbb{R}_+$ , there exists a constant  $C$  such that for every test function  $\varphi$  in  $\mathcal{W}_0^{k,\alpha}$ ,  $\|f\varphi\|_{k,\alpha} \leq C\|f\|_{\mathcal{C}_b^k}\|\varphi\|_{k,\alpha}$ .*

*Proof.* The first assertion follows from the definition of  $\|\cdot\|_{2,\alpha}$ , and the second one follows from Leibniz's rule and the definition of  $\|\cdot\|_{k,\alpha}$ .  $\square$

Let us denote  $R$  (for *reset*) the linear mapping defined by  $R\varphi := \varphi(0) - \varphi(\cdot)$  where  $\varphi$  is some test function. This mapping naturally appears in our problem since the age process jumps to the value 0 at each point of the underlying point process, as it appears below in Proposition IV.5.

**Lemma IV.2.** *For any integer  $k \geq 1$  and  $\alpha > 1/2$ , the linear mapping  $R$  is continuous from  $\mathcal{W}_0^{k,\alpha}$  to itself.*

*Proof.* The function  $R\varphi$  only differs from  $\varphi$  by a constant so the derivatives of  $R\varphi$  are equal to the derivatives of  $\varphi$ . Hence, using the convexity of the square function, we have

$$\begin{aligned} \|R\varphi\|_{k,\alpha}^2 &\leq \int_{\mathbb{R}_+} \frac{2|\varphi(0)|^2}{1+|x|^{2\alpha}} dx + \int_{\mathbb{R}_+} \frac{2|\varphi(x)|^2}{1+|x|^{2\alpha}} dx + \sum_{k'=1}^k \int_{\mathbb{R}_+} \frac{|\varphi^{(k')}(x)|^2}{1+|x|^{2\alpha}} dx \\ &\leq 2 \int_{\mathbb{R}_+} \frac{1}{1+|x|^{2\alpha}} dx |\varphi(0)|^2 + 2\|\varphi\|_{k,\alpha}^2. \end{aligned}$$

Yet,  $|\varphi(0)| \leq \|\varphi\|_{\mathcal{C}^0,\alpha} \leq C\|\varphi\|_{k,\alpha}$  by (31) and  $\int_{\mathbb{R}_+} 1/(1+|x|^{2\alpha}) dx < +\infty$ , for any fixed  $\alpha > 1/2$ , so that  $\|R\varphi\|_{k,\alpha}^2 \leq C\|\varphi\|_{k,\alpha}^2$ .  $\square$

**Lemma IV.3.** *For any fixed  $\alpha$  in  $\mathbb{R}_+$  and  $x, y$  in  $\mathbb{R}$ , the mappings  $\delta_x$  and  $D_{x,y} : \mathcal{W}_0^{1,\alpha} \rightarrow \mathbb{R}$ , defined by  $\delta_x(\varphi) := \varphi(x)$  and  $D_{x,y}(\varphi) := \varphi(x) - \varphi(y)$  are linear continuous. In particular, for all  $\alpha$  in  $\mathbb{R}_+$ , there exist some positive constants  $C_1$  and  $C_2$  such that, if  $x$  and  $y$  are bounded by some constant  $M$ , i.e.  $|x| \leq M$  and  $|y| \leq M$ , then*

$$\begin{cases} \|\delta_x\|_{-2,\alpha} \leq \|\delta_x\|_{-1,\alpha} \leq C_1(1+M^\alpha), \\ \|D_{x,y}\|_{-2,\alpha} \leq \|D_{x,y}\|_{-1,\alpha} \leq C_2(1+M^\alpha). \end{cases} \quad (34)$$

*Proof.* Remark that  $|D_{x,y}(\varphi)| \leq |\varphi(x)| + |\varphi(y)| = |\delta_x(\varphi)| + |\delta_y(\varphi)|$ . Hence, it suffices to show that there exists some positive constant  $C$  such that  $\|\delta_x\|_{-1,\alpha} \leq C(1+|x|^\alpha)$ . Yet,  $|\delta_x(\varphi)| = |\varphi(x)| \leq \|\varphi\|_{C^{0,\alpha}}(1+|x|^\alpha) \leq C\|\varphi\|_{1,\alpha}(1+|x|^\alpha)$  by (31).  $\square$

**Remark IV.4.** *At this point, let us mention two reasons why weighted Sobolev spaces are more appropriate than standard (non-weighted) Sobolev spaces of functions on  $\mathbb{R}_+$ :*

- *we want to be able to consider functions of  $\mathcal{C}_b^k$  as test functions: indeed,  $\Psi$  must be considered as a test function, in Equation (55) below for instance, yet we do not want  $\Psi$  to be compactly supported with respect to the age  $s$  or even to rapidly decrease when  $s$  goes to infinity. The natural space to which  $\Psi$  belongs is some  $\mathcal{C}_b^k$  space,*
- *in order to ensue criterion (A<sub>1'</sub>), a compact embedding is required but Maurin's theorem does not apply for standard Sobolev spaces on  $\mathbb{R}_+$  (see [1, Theorem 6.37]).*

In order to apply Lemma IV.2 and to satisfy the first point in the remark above, the weight  $\alpha$  is assumed to be greater than 1/2 in all the next sections so that (30) holds true.

## IV.2 Decomposition of the fluctuations

Here, we give a semi-martingale representation of  $\eta^n$  used to simplify the study of tightness (recall that  $R$  is defined above in Lemma IV.2).

**Proposition IV.5.** *Under Assumption (A<sub>LLN</sub>), for every test function  $\varphi$  in  $\mathcal{C}_b^1$  and  $t \geq 0$ ,*

$$\langle \eta_t^n, \varphi \rangle - \langle \eta_0^n, \varphi \rangle = \int_0^t (\langle \eta_z^n, L_z \varphi \rangle + A_z^n(\varphi)) dz + M_t^n(\varphi), \quad (35)$$

with  $L_z \varphi(s) = \varphi'(s) + \Psi(s, \bar{\gamma}(z))R\varphi(s)$  for all  $z \geq 0$  and  $s$  in  $\mathbb{R}$ , where  $\bar{\gamma}$  is defined by (11), and

$$\begin{cases} M_t^n(\varphi) := n^{-1/2} \sum_{i=1}^n \int_0^t R\varphi(S_{z-}^{n,i}) (N^{n,i}(dz) - \lambda_z^{n,i} dz), \\ A_z^n(\varphi) := n^{-1/2} \sum_{i=1}^n R\varphi(S_{z-}^{n,i}) (\lambda_z^{n,i} - \Psi(S_{z-}^{n,i}, \bar{\gamma}(z))). \end{cases} \quad (36)$$

Furthermore, for any  $\varphi$  in  $\mathcal{C}_b^1$ ,  $(M_t^n(\varphi))_{t \geq 0}$  is a real valued  $\mathbb{F}$ -martingale with angle bracket given by

$$\langle M^n(\varphi) \rangle_t = \frac{1}{n} \sum_{i=1}^n \int_0^t R\varphi(S_{z-}^{n,i})^2 \lambda_z^{n,i} dz. \quad (37)$$

**Remark IV.6.** *To avoid confusion, let us mention that (36) defines  $M_t^n$  and  $A_z^n$  as distributions acting on test functions. More precisely, we show below that they can be seen as distributions in  $\mathcal{W}_0^{-2,\alpha}$  (Proposition IV.7). However, we do not use the notation for the dual action  $\langle \cdot, \cdot \rangle$  to avoid tricky notation involving several angle brackets in (37) for instance.*

The proof of Proposition IV.5 relies on the integrability properties of the stochastic intensity and is given in Appendix A.2.

### IV.3 Estimates in dual spaces

Below are stated estimates of the terms  $\eta^n$ ,  $A^n$  and  $M^n$  - appearing in (35) - regarded as distributions. More precisely, the estimates given in this section are stated in terms of the norm on either  $\mathcal{W}_0^{-1,\alpha}$  or  $\mathcal{W}_0^{-2,\alpha}$  for any  $\alpha > 1/2$  (in comparison with  $\mathcal{W}_0^{-2,2}$  and  $\mathcal{W}_0^{-4,1}$  in [23] for instance). Usually, like in [16, 23, 26, 28], the weight is linked to the maximal order of the moment estimates obtained on the positions of the particles. Here, the age processes are bounded in finite time horizon (remind (8)) so the weight  $\alpha$  of the Sobolev space can be taken as large as wanted. The weighted Sobolev spaces are nevertheless interesting here since, in particular, the distribution  $\eta_t^n$  belongs to  $\mathcal{W}_0^{-1,\alpha}$  for all  $t \geq 0$  (see Proposition IV.7 below). We refer to the introductory discussion in Section I for complements on the usefulness of the weights.

We first give estimates in the smaller space  $\mathcal{W}_0^{-1,\alpha}$ . This is later used in order to prove tightness (remember condition (A<sub>1'</sub>) of the Aldous type criterion stated on page 14).

**Proposition IV.7.** *Under Assumption (A<sub>LLN</sub>), for any  $\alpha > 1/2$  and  $\theta \geq 0$ , the following statements hold true:*

(i) *the sequence  $(\eta^n)_{n \geq 1}$  is such that,*

$$\sup_{n \geq 1} \sup_{t \in [0, \theta]} \mathbb{E} \left[ \|\eta_t^n\|_{-1, \alpha}^2 \right] < +\infty, \quad (38)$$

(ii) *the process  $(M_t^n)_{t \geq 0}$ , defined by (36), is an  $\mathbb{F}$ -martingale which belongs to  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-1,\alpha})$  almost surely. Furthermore, for any  $\theta \geq 0$ ,*

$$\sup_{n \geq 1} \mathbb{E} \left[ \sup_{t \in [0, \theta]} \|M_t^n\|_{-1, \alpha}^2 \right] < +\infty. \quad (39)$$

(iii) *the sequence  $(A^n)_{n \geq 1}$ , defined by (36), is such that,*

$$\sup_{n \geq 1} \sup_{t \in [0, \theta]} \mathbb{E} \left[ \|A_t^n\|_{-2, \alpha}^2 \right] < +\infty. \quad (40)$$

(iv) *under (A<sub>s, c<sub>b</sub><sup>Ψ</sup></sub>), for any  $z$  in  $\mathbb{R}_+$ , the application  $L_z$  defined in Proposition IV.5 is a linear continuous mapping from  $\mathcal{W}_0^{2,\alpha}$  to  $\mathcal{W}_0^{1,\alpha}$  and, for all  $\varphi$  in  $\mathcal{W}_0^{2,\alpha}$ ,*

$$\sup_{z \in [0, \theta]} \frac{\|L_z \varphi\|_{1, \alpha}^2}{\|\varphi\|_{2, \alpha}^2} < +\infty. \quad (41)$$

The proof of Proposition IV.7 is given in Appendix A.3 and mainly relies on the estimates given in Lemma IV.3. However, let us mention that:

- the following expansion is used in the proof of (iii) as well as in Section V.1: using that  $\lambda_t^{n,i} = \Psi(S_{t-}^{n,i}, \gamma_t^n)$  and (A<sub>y, c<sub>2</sub><sup>Ψ</sup></sub>), it follows from Taylor's inequality that for  $\varphi$  in  $\mathcal{W}_0^{2,\alpha}$ ,

$$A_t^n(\varphi) = \frac{1}{n} \sum_{i=1}^n R\varphi(S_{t-}^{n,i}) \frac{\partial \Psi}{\partial y}(S_{t-}^{n,i}, \bar{\gamma}(t)) \left( \sqrt{n}(\gamma_t^n - \bar{\gamma}(t)) + \sqrt{nr}r_t^{n,i} \right), \quad (42)$$

with the rests satisfying  $|r_t^{n,i}| \leq \sup_{s,y} |\frac{\partial^2 \Psi}{\partial y^2}(s,y)| |\gamma_t^n - \bar{\gamma}(t)|^2/2$ . This upper-bound does not depend on  $\varphi$ . Let us denote  $\Gamma_{t-}^n := \sqrt{n}(\gamma_t^n - \bar{\gamma}(t))$  and

$$R_t^{n,(1)}(\varphi) := \frac{1}{n} \sum_{i=1}^n (R\varphi(S_{t-}^{n,i}) \frac{\partial \Psi}{\partial y}(S_{t-}^{n,i}, \bar{\gamma}(t)) \sqrt{n} r_t^{n,i}),$$

so that (42) rewrites as

$$A_t^n(\varphi) = \left\langle \bar{\mu}_{S_t}^n, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t)) R\varphi \right\rangle \Gamma_{t-}^n + R_t^{n,(1)}(\varphi). \quad (43)$$

- Lemma IV.1 and the following properties are used to prove point (iv): under Assumption  $(\mathcal{A}_{s,C_b^2}^\Psi)$ , the functions

$$t \mapsto \|\Psi(\cdot, \bar{\gamma}(t))\|_{C_b^2} \text{ and } t \mapsto \left\| \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t)) \right\|_{C_b^1} \text{ are locally bounded,} \quad (44)$$

since  $t \mapsto \bar{\gamma}(t)$  is locally bounded. In the same way, under Assumption  $(\mathcal{A}_{s,C_b^4}^\Psi)$ , the function

$$t \mapsto \|\Psi(\cdot, \bar{\gamma}(t))\|_{C_b^4} \text{ is locally bounded.} \quad (45)$$

Proposition IV.7, combined with the first line of Equation (33), gives that  $\eta^n$ ,  $A^n$  and  $M^n$  belong to  $\mathcal{W}_0^{-2,\alpha}$ . Hence, we may consider the following decomposition in  $\mathcal{W}_0^{-2,\alpha}$ ,

$$\eta_t^n - \eta_0^n = \int_0^t L_z^* \eta_z^n dz + \int_0^t A_z^n dz + M_t^n, \quad (46)$$

where  $L_z^*$  is the adjoint operator of  $L_z$ .

**Remark IV.8.** As a corollary of Proposition IV.7-(iv), one has, for all  $\alpha > 1/2$ , all  $w$  in  $\mathcal{W}_0^{-1,\alpha}$  and all  $\theta \geq 0$ ,

$$\sup_{z \in [0, \theta]} \frac{\|L_z^* w\|_{-2,\alpha}^2}{\|w\|_{-1,\alpha}^2} < +\infty. \quad (47)$$

Indeed, both  $\|L_z^* w\|_{-2,\alpha}^2 \leq \sup_{\|\varphi\|_{2,\alpha}=1} \|L_z \varphi\|_{1,\alpha}^2 \|w\|_{-1,\alpha}^2$  and Equation (41) give the result.

Furthermore, the Doob-Meyer process  $(\ll M^n \gg_t)_{t \geq 0}$  associated with the square integrable  $\mathbb{F}$ -martingale  $(M_t^n)_{t \geq 0}$  satisfies the following: for any  $t \geq 0$ ,  $\ll M^n \gg_t$  is the linear continuous mapping from  $\mathcal{W}_0^{2,\alpha}$  to  $\mathcal{W}_0^{-2,\alpha}$  given, for all  $\varphi_1, \varphi_2$  in  $\mathcal{W}_0^{2,\alpha}$ , by

$$\langle \ll M^n \gg_t(\varphi_1), \varphi_2 \rangle = \frac{1}{n} \sum_{i=1}^n \int_0^t R\varphi_1(S_{z-}^{n,i}) R\varphi_2(S_{z-}^{n,i}) \lambda_z^{n,i} dz.$$

This last equation can be retrieved thanks to the polarization identity from (37).

Yet, to give sense to Equation (46), we need the lemma stated below.

**Lemma IV.9.** Under  $(\mathcal{A}_{TGN})$ , the integrals  $\int_0^t L_z^* \eta_z^n dz$  and  $\int_0^t A_z^n dz$  are almost surely well defined as Bochner integrals in  $\mathcal{W}_0^{-2,\alpha}$  for any  $\alpha > 1/2$ . In particular, the functions  $t \mapsto \int_0^t L_z^* \eta_z^n dz$  and  $t \mapsto \int_0^t A_z^n dz$  are almost surely strongly continuous in  $\mathcal{W}_0^{-2,\alpha}$ .

*Proof.* Since  $\mathcal{W}_0^{-2,\alpha}$  is separable, it suffices to verify that (see Yosida [42, p. 133]):

- (i) for every  $\varphi$  in  $\mathcal{W}_0^{2,\alpha}$ , the functions  $z \mapsto \langle L_z^* \eta_z^n, \varphi \rangle = \langle \eta_z^n, L_z \varphi \rangle$  and  $z \mapsto A_z^n(\varphi)$  are measurable,
- (ii) the integrals  $\int_0^t \|L_z^* \eta_z^n\|_{-2,\alpha} dz$  and  $\int_0^t \|A_z^n\|_{-2,\alpha} dz$  are finite almost surely.

The first condition is immediate. The second one follows from the controls we have shown.

Indeed, on the one hand, it follows from Equation (47) that  $\int_0^t \|L_z^* \eta_z^n\|_{-2,\alpha} dz \lesssim_t \int_0^t \|\eta_z^n\|_{-1,\alpha} dz$  and Proposition IV.7-(i) implies  $\mathbb{E}[\int_0^t \|\eta_z^n\|_{-1,\alpha+1} dz] < +\infty$  so that  $\int_0^t \|L_z^* \eta_z^n\|_{-2,\alpha} dz$  is finite a.s.

On the other hand, Proposition IV.7-(iii) gives that  $\mathbb{E}[\int_0^t \|A_z^n\|_{-2,\alpha} dz]$  is finite and so  $\int_0^t \|A_z^n\|_{-2,\alpha} dz$  is finite a.s.  $\square$

Now, using the decomposition (46) we are able to somehow exchange the expectation with the supremum in the control of  $\eta$ , i.e. Equation (38).

**Proposition IV.10.** *Under  $(\mathcal{A}_{TGN})$ , for every  $\alpha > 1/2$  and  $\theta \geq 0$ ,*

$$\sup_{n \geq 1} \mathbb{E} \left[ \sup_{t \in [0, \theta]} \|\eta_t^n\|_{-2,\alpha}^2 \right] < +\infty, \quad (48)$$

and  $t \mapsto \eta_t^n$  belongs to  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha})$  almost surely.

*Proof.* Starting from (46), we have by convexity of the square function

$$\sup_{t \in [0, \theta]} \|\eta_t^n\|_{-2,\alpha}^2 \leq 4 \left[ \|\eta_0^n\|_{-2,\alpha}^2 + \theta \int_0^\theta (\|L_z^* \eta_z^n\|_{-2,\alpha}^2 + \|A_z^n\|_{-2,\alpha}^2) dz + \sup_{t \in [0, \theta]} \|M_t^n\|_{-2,\alpha}^2 \right].$$

We deduce from Equation (41) that  $\int_0^\theta \mathbb{E}[\|L_z^* \eta_z^n\|_{-2,\alpha}^2] dz \lesssim_\theta \sup_{z \in [0, \theta]} \mathbb{E}[\|\eta_z^n\|_{-1,\alpha}^2]$ . Hence, taking the expectation in both sides of the inequality above and applying Proposition IV.7 (remind (33)), we get (48). Starting from (46) and using that the integrals are continuous from Lemma IV.9 and  $M^n$  is càdlàg from Proposition IV.7-(ii), it follows that  $\eta^n$  is càdlàg.  $\square$

#### IV.4 Tightness result

Using the estimates proved in Section IV.3, the tightness criterion stated on page 13 can be checked.

**Theorem IV.11.** *Under  $(\mathcal{A}_{TGN})$ , for any  $\alpha > 1/2$ , the sequences of the laws of  $(M^n)_{n \geq 1}$  and of  $(\eta^n)_{n \geq 1}$  are tight in the space  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha})$ .*

*Proof.* Condition  $(A_1')$  with  $H_0 = \mathcal{W}_0^{-1,\alpha+1}$  and  $H = \mathcal{W}_0^{-2,\alpha}$  is satisfied for both processes as a consequence of embedding (33) (remind that Hilbert-Schmidt operators are compact) and Proposition IV.7.

On the one hand, condition  $(A_2)$  holds for  $(M^n)_{n \geq 1}$  as soon as it holds for the trace of the processes  $(\ll M^n \gg)_{n \geq 1}$  given below (46) [22, Rebolledo's theorem, p. 40]. Let  $(\varphi_k)_{k \geq 1}$  be an orthonormal basis of  $\mathcal{W}_0^{2,\alpha}$ . Let  $\theta \geq 0$ ,  $\delta_0 > 0$  and  $\delta \leq \delta_0$ .

Furthermore, let  $\tau_n$  be an  $\mathbb{F}$ -stopping time smaller than  $\theta$ .

$$\begin{aligned} & |\mathrm{Tr} \ll M^n \gg_{\tau_n + \delta} - \mathrm{Tr} \ll M^n \gg_{\tau_n}| \\ &= \left| \sum_{k \geq 1} \langle \ll M^n \gg_{\tau_n + \delta}(\varphi_k), \varphi_k \rangle - \langle \ll M^n \gg_{\tau_n}(\varphi_k), \varphi_k \rangle \right| \\ &\leq \sum_{k \geq 1} \frac{1}{n} \sum_{i=1}^n \int_{\tau_n}^{\tau_n + \delta} [R\varphi_k(S_{z-}^{n,i})]^2 \lambda_z^{n,i} dz \leq \|\Psi\|_\infty \frac{1}{n} \sum_{i=1}^n \int_{\tau_n}^{\tau_n + \delta} \sum_{k \geq 1} R\varphi_k(S_{z-}^{n,i})^2 dz. \end{aligned}$$

Noticing that  $R\varphi_k(S_{z-}^{n,i}) = D_{0, S_{z-}^{n,i}}(\varphi_k)$  and then using Lemma IV.3 and the fact that the ages  $S_{z-}^{n,i}$  are upper bounded by  $M_{S_0} + z + \leq M_{S_0} + \theta + \delta_0$  (thanks to  $(\mathcal{A}_\infty^{u_0})$ , remind (8)), it follows that

$$\mathbb{E} [|\mathrm{Tr} \ll M^n \gg_{\tau_n + \delta} - \mathrm{Tr} \ll M^n \gg_{\tau_n}|] \leq \delta_0 \|\Psi\|_\infty (C_2)^2 (1 + (M_{S_0} + \theta + \delta_0)^\alpha)^2.$$

This last bound is arbitrarily small for  $\delta_0$  small enough which gives condition  $(\mathbf{A}_2)$  thanks to Markov's inequality.

On the other hand, using decomposition (46) and the fact that  $(M^n)_{n \geq 1}$  is tight, it suffices to show the tightness of the remaining terms  $(R_t^n = \eta_0^n + \int_0^t L_z^* \eta_z^n dz + \int_0^t A_z^n dz)_{n \geq 1}$  in order to show tightness of  $(\eta^n)_{n \geq 1}$ . Yet, using Equation (47), we have

$$\begin{aligned} & \|R_{\tau_n + \delta}^n - R_{\tau_n}^n\|_{-2, \alpha}^2 = \left\| \int_{\tau_n}^{\tau_n + \delta} L_z^* \eta_z^n + A_z^n dz \right\|_{-2, \alpha}^2 \\ &\leq 2\delta \int_{\tau_n}^{\tau_n + \delta} (\|L_z^* \eta_z^n\|_{-2, \alpha}^2 + \|A_z^n\|_{-2, \alpha}^2) dz \leq 2\delta_0 \int_0^{\theta + \delta_0} (C \|\eta_z^n\|_{-1, \alpha+1}^2 + \|A_z^n\|_{-2, \alpha}^2) dz, \end{aligned}$$

where  $C$  depends on  $\theta$  and  $\delta_0$ . Then, Proposition IV.7 implies that  $\sup_{n \geq 1} \mathbb{E} [\|R_{\tau_n + \delta}^n - R_{\tau_n}^n\|_{-2, \alpha}^2] \leq C\delta_0$  for  $\delta_0$  small enough. Finally, Markov's inequality gives condition  $(\mathbf{A}_2)$  for  $(R^n)_{n \geq 1}$  and so the tightness of  $(\eta^n)_{n \geq 1}$ .  $\square$

**Remark IV.12.** For any  $\alpha > 1/2$ , every limit (with respect to the convergence in law)  $M$  (respectively  $\eta$ ) in  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha})$  of the sequence  $(M^n)_{n \geq 1}$  (resp.  $(\eta^n)_{n \geq 1}$ ) satisfies

$$\mathbb{E} \left[ \sup_{t \in [0, \theta]} \|M_t\|_{-2, \alpha}^2 \right] < +\infty \quad \left( \text{resp. } \mathbb{E} \left[ \sup_{t \in [0, \theta]} \|\eta_t\|_{-2, \alpha}^2 \right] < +\infty \right). \quad (49)$$

Moreover, the limit laws are supported in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha})$ .

*Proof.* Let us first show that the limit points are continuous. According to [5, Theorem 13.4.], it suffices to prove that for all  $\theta \geq 0$ , the maximal jump size of  $M^n$  and  $\eta^n$  on  $[0, \theta]$  converge to 0 almost surely in order to prove the last point. Yet, for all  $\varphi$  in  $\mathcal{W}_0^{2, \alpha}$ ,

$$\Delta M_t^n(\varphi) := |M_t^n(\varphi) - M_{t-}^n(\varphi)| = \frac{1}{\sqrt{n}} \sum_{i=1}^n D_{0, S_{t-}^{n,i}}(\varphi) \mathbf{1}_{t \in N^{n,i}},$$

where we use the definition of  $M_t^n(\varphi)$  given by (36) for  $\varphi$  in  $\mathcal{C}_b^1$  and a density argument to extend it to  $\varphi$  in  $\mathcal{W}_0^{2, \alpha}$ , and

$$\langle \Delta \eta_t^n, \varphi \rangle := |\langle \eta_t^n, \varphi \rangle - \langle \eta_{t-}^n, \varphi \rangle| = \frac{1}{\sqrt{n}} \sum_{i=1}^n D_{0, S_{t-}^{n,i}}(\varphi) \mathbf{1}_{t \in N^{n,i}}$$

where we used the fact that  $(u_t)_{t \geq 0}$  is continuous in  $\mathcal{W}_0^{-2,\alpha}$  (see Lemma B.2). Since almost surely there is no common point to any two of the point processes  $(N^{n,i})_{i=1,\dots,n}$ , there is, almost surely, for all  $t \geq 0$ , at most one of the  $\mathbb{1}_{t \in N^{n,i}}$  which is non null. Then, Lemma IV.3 implies

$$\begin{cases} \sup_{t \in [0, \theta]} \|\Delta M_t^n\|_{-2,\alpha} \leq \frac{1}{\sqrt{n}} C_2 (1 + (M_{S_0} + \theta)^\alpha), \\ \sup_{t \in [0, \theta]} \|\Delta \eta_t^n\|_{-2,\alpha} \leq \frac{1}{\sqrt{n}} C_2 (1 + (M_{S_0} + \theta)^\alpha), \end{cases}$$

which gives the desired convergence to 0.

Finally, the two statements of Equation (49) are consequences of Propositions IV.7-(ii) (remind (33)) and IV.10 where we use the previous step and the fact that the mapping  $g \mapsto \sup_{t \in [0, \theta]} \|g_t\|_{-2,\alpha}^2$  from  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha})$  to  $\mathbb{R}$  is continuous at every point  $g^0$  in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha})$ .  $\square$

## V Characterization of the limit

The aim of this section is to prove convergence of the sequence  $(\eta^n)_{n \geq 1}$  by identifying the limit fluctuation process  $\eta$  as the unique solution of a SDE in infinite dimension. We first prove, in Section V.1, that every possible limit process  $\eta$  satisfies a certain SDE (Theorem V.6). Then, we show, in Section V.2, that this SDE uniquely characterizes the limit law, which completes the proof of the convergence in law of  $(\eta^n)_{n \geq 1}$  to  $\eta$ .

### V.1 Candidate for the limit equation

In this section, the limit version of Equation (46) is stated. Apart from  $\eta^n$ , there are two random processes in (46) that are  $A^n$  and  $M^n$ . The following notation encompasses the source of the stochasticity of both  $A^n$  and  $M^n$  and is mainly used in order to track the correlations between those two quantities: for all  $n \geq 1$ , let  $W^n$  be the  $\mathcal{W}_0^{-1,\alpha}$ -valued martingale defined, for all  $t \geq 0$  and  $\varphi$  in  $\mathcal{W}_0^{1,\alpha}$ , by

$$W_t^n(\varphi) := \frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^t \varphi(S_{z-}^{n,i})(N^{n,i}(dz) - \lambda_z^{n,i} dz).$$

Notice that  $M_t^n(\varphi) = W_t^n(R\varphi)$ . Furthermore, as for  $M^n$ , the Doob-Meyer process  $(\ll W^n \gg_t)_{t \geq 0}$  associated with  $(W_t^n)_{t \geq 0}$  satisfies the following: for any  $t \geq 0$ ,  $\ll W^n \gg_t$  is the linear continuous mapping from  $\mathcal{W}_0^{2,\alpha}$  to  $\mathcal{W}_0^{-2,\alpha}$  given, for all  $\varphi_1$  and  $\varphi_2$  in  $\mathcal{W}_0^{2,\alpha}$ , by

$$\langle \ll W^n \gg_t(\varphi_1), \varphi_2 \rangle = \frac{1}{n} \sum_{i=1}^n \int_0^t \varphi_1(S_{z-}^{n,i}) \varphi_2(S_{z-}^{n,i}) \lambda_z^{n,i} dz. \quad (50)$$

All the results given for  $M^n$  in the previous section can be extended to  $W^n$ . In particular,

$$\text{the sequence } (W^n)_{n \geq 1} \text{ is tight in } \mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha}). \quad (51)$$

Next, we prove that it converges towards the Gaussian process  $W$  defined below.

**Definition V.1.** For any  $\alpha > 1/2$ , let  $W$  be a continuous centred Gaussian process with values in  $\mathcal{W}_0^{-2,\alpha}$  with covariance given, for all  $\varphi_1$  and  $\varphi_2$  in  $\mathcal{W}_0^{2,\alpha}$ , for all  $t$  and



$t' \geq 0$ , by

$$\begin{aligned} \mathbb{E} [W_t(\varphi_1)W_{t'}(\varphi_2)] &= \int_0^{t \wedge t'} \langle u_z, \varphi_1 \varphi_2 \Psi(\cdot, \bar{\gamma}(z)) \rangle dz \\ &= \int_0^{t \wedge t'} \int_0^{+\infty} \varphi_1(s) \varphi_2(s) \Psi(s, \bar{\gamma}(z)) u(z, s) ds dz, \end{aligned} \quad (52)$$

where  $u$  is the unique solution of (4).

**Remark V.2.** We refer to the PhD manuscript of the author [10] for the existence and uniqueness in law of such a process  $W$ . Yet, let us mention here that the process  $W$  defined above does not depend on the weight  $\alpha$  in the sense that the definition is consistent with respect to the weights. Indeed, say  $W^\alpha$  and  $W^\beta$  are two processes in the sense of Definition V.1 with values in  $\mathcal{W}_0^{-2,\alpha}$  and  $\mathcal{W}_0^{-2,\beta}$  respectively. Assume for instance that  $\beta > \alpha$ . Then,  $W^\beta$  can be seen as a process with values in  $\mathcal{W}_0^{-2,\alpha}$  via the canonical embedding  $\mathcal{W}_0^{-2,\beta} \hookrightarrow \mathcal{W}_0^{-2,\alpha}$ . Yet, the covariance structure (52) does not depend on the weights  $\alpha$  and  $\beta$  so  $W^\beta$  is also a Gaussian process with values in  $\mathcal{W}_0^{-2,\alpha}$  with the prescribed covariance and the uniqueness in law guarantees the equality of the laws of  $W^\alpha$  and  $W^\beta$  as  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha})$ -valued random variables.

**Proposition V.3.** Under  $(\mathcal{A}_{TGN})$ , for any  $\alpha > 1/2$ , the sequence  $(W^n)_{n \geq 1}$  of processes in  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha})$  converges in law to  $W$ .

The proof of Proposition V.3 is given in Appendix A.4. It relies on the convergence of the bracket (50) towards the covariance (52) and an application of Rebolledo's central limit theorem (the maximum size of the jumps is bounded up to a constant by  $n^{-1/2}$  and so goes to 0).

Denote by  $\mathbf{1} : \mathbb{R}_+ \rightarrow \mathbb{R}$  the constant function equal to 1 (which belongs to  $\mathcal{W}_0^{2,\alpha}$  since we assume  $\alpha > 1/2$ ) and note that  $W_t^n(\mathbf{1})$  is the rescaled canonical martingale associated with the system of age-dependent Hawkes processes, namely

$$W_t^n(\mathbf{1}) = \sqrt{n} \left( \frac{1}{n} \sum_{i=1}^n N_t^{n,i} - \int_0^t \lambda_z^{n,i} dz \right).$$

Now, let us expand the decomposition (46) in order to get a closed equation. Let us recall the expansion of  $A^n$  given by (43), that is

$$A_t^n(\varphi) = \left\langle \bar{\mu}_{S_t}^n, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t)) R\varphi \right\rangle \Gamma_{t-}^n + R_t^{n,(1)}(\varphi),$$

with  $\Gamma_{t-}^n = \sqrt{n}(\gamma_t^n - \bar{\gamma}(t))$  and the rest term:

$$R_t^{n,(1)}(\varphi) := \frac{1}{n} \sum_{i=1}^n (R\varphi(S_{t-}^{n,i}) \frac{\partial \Psi}{\partial y}(S_{t-}^{n,i}, \bar{\gamma}(t)) \sqrt{nr_t^{n,i}}).$$

Below, we use the fact that this rest term converges to 0 in  $L^1$  norm: indeed, recall that

$$|r_t^{n,i}| \lesssim |\gamma_t^n - \bar{\gamma}(t)|^2 \quad (53)$$

and, thanks to Proposition III.1,

$$\mathbb{E} [|\gamma_t^n - \bar{\gamma}(t)|^2] \lesssim_t n^{-1}.$$

Since  $\Gamma_{t-}^n$  (as part of  $A_t^n(\varphi)$ ) only appears in (46) as an integrand and is only discontinuous on a set of Lebesgue measure equal to zero, we can replace it by its

càdlàg version denoted by  $\Gamma_t^n$ . Let us consider the decomposition  $\Gamma_t^n = \Upsilon_t^1 + \Upsilon_t^2 + \Upsilon_t^3$ , with

$$\begin{cases} \Upsilon_t^1 := \sqrt{n} \int_0^t h(t-z) \left( \frac{1}{n} \sum_{i=1}^n N^{n,i}(dz) - \lambda_z^{n,i} dz \right) = \int_0^t h(t-z) dW_z^n(\mathbf{1}), \\ \Upsilon_t^2 := \sqrt{n} \int_0^t h(t-z) \frac{1}{n} \sum_{i=1}^n (\lambda_z^{n,i} - \Psi(S_{z-}^{n,i}, \bar{\gamma}(z))) dz, \\ \Upsilon_t^3 := \sqrt{n} \int_0^t h(t-z) \frac{1}{n} \sum_{i=1}^n (\Psi(S_{z-}^{n,i}, \bar{\gamma}(z)) - \bar{\lambda}(z)) dz = \int_0^t h(t-z) \langle \eta_z^n, \Psi(\cdot, \bar{\gamma}(z)) \rangle dz, \end{cases}$$

where we used, in the last line, the fact that  $\bar{\mu}_{S_{z-}}^n = \bar{\mu}_{S_z}^n$  for almost every  $z$  in  $\mathbb{R}_+$ , and  $\bar{\lambda}(z) = \langle u_z, \Psi(\cdot, \bar{\gamma}(z)) \rangle$ .

Based on Assumption  $(\mathcal{A}_{y, \mathcal{C}^2}^\Psi)$ , as for Equation (42), one can give the Taylor expansion of the term

$$\Upsilon_t^2 = \sqrt{n} \int_0^t h(t-z) \frac{1}{n} \sum_{i=1}^n (\Psi(S_{z-}^{n,i}, \gamma_z^n) - \Psi(S_{z-}^{n,i}, \bar{\gamma}(z))) dz.$$

On the one hand, gathering the decomposition (35) with (43) and on the other hand gathering  $\Gamma_t^n = \Upsilon_t^1 + \Upsilon_t^2 + \Upsilon_t^3$  with the Taylor expansion of  $\Upsilon_t^2$  give that  $(\eta^n, \Gamma^n)$  satisfies the following closed system for all  $\varphi$  in  $\mathcal{W}_0^{2,\alpha}$ ,

$$\begin{aligned} \langle \eta_t^n, \varphi \rangle - \langle \eta_0^n, \varphi \rangle &= \int_0^t \langle \eta_z^n, L_z \varphi \rangle dz + \int_0^t \left\langle \bar{\mu}_{S_z}^n, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) R \varphi \right\rangle \Gamma_z^n dz \\ &\quad + \int_0^t R_z^{n,(1)}(\varphi) dz + W_t^n(R\varphi), \end{aligned} \quad (54)$$

$$\begin{aligned} \Gamma_t^n &= \int_0^t h(t-z) \left\langle \bar{\mu}_{S_z}^n, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \Gamma_z^n dz + \int_0^t h(t-z) R_z^{n,(2)} dz \\ &\quad + \int_0^t h(t-z) \langle \eta_z^n, \Psi(\cdot, \bar{\gamma}(z)) \rangle dz + \int_0^t h(t-z) dW_z^n(\mathbf{1}), \end{aligned} \quad (55)$$

where the rest term  $R_z^{n,(2)}$  is defined by

$$R_z^{n,(2)} := \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{\partial \Psi}{\partial y}(S_{t-}^{n,i}, \bar{\gamma}(t)) r_t^{n,i}.$$

Once again, notice that  $\Gamma_{z-}^n$ , which naturally appears in the first integral term of (55), is replaced by its càdlàg version  $\Gamma_z^n$  since they are equal except on a null measure set.

Let us denote  $V_t^n := \int_0^t h(t-z) dW_z^n(\mathbf{1})$  and  $V_t := \int_0^t h(t-z) dW_z(\mathbf{1})$ . The convergence of the sources of stochasticity in the system (54)-(55) is stated in the following corollary of Proposition V.3.

**Corollary V.4.** *Under  $(\mathcal{A}_{TGN})$  and  $(\mathcal{A}_{HöL}^h)$ , the following convergence in law holds true in  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha} \times \mathbb{R})$ ,*

$$\left( R^* W_t^n, V_t^n \right)_{t \geq 0} \Rightarrow \left( R^* W_t, V_t \right)_{t \geq 0},$$

where  $R^*$  denotes the adjoint of  $R$ .

The proof of Corollary V.4 uses Billingsley tightness criterion for real-valued stochastic processes and is given in Appendix A.5.

Before taking the limit  $n \rightarrow +\infty$  in the system (54)-(55), we state the tightness of  $(\Gamma^n)_{n \geq 1}$ . Nevertheless, let us first mention that we use the following estimates: as a consequence of Proposition III.1, for all  $k \geq 0$  and  $\theta \geq 0$ ,

$$\sup_{t \in [0, \theta]} \mathbb{E} \left[ |\Gamma_t^n|^k \right] < +\infty, \quad (56)$$

since  $\sup_{t \in [0, \theta]} \mathbb{E} [|\Gamma_t^n|^k] = \sup_{t \in [0, \theta]} \mathbb{E} [|\Gamma_{t-}^n|^k]$  because the underlying point processes admit intensities so that there is almost surely no jump at time  $\theta$ .

**Proposition V.5.** *Under  $(\mathcal{A}_{TGN})$  and  $(\mathcal{A}_{HöI}^h)$ , the sequence of the laws of  $(\Gamma^n)_{n \geq 1}$  is tight in  $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$ . Furthermore, the possible limit laws are supported in  $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$  and satisfy, for all  $k \geq 0$ ,*

$$\sup_{t \in [0, \theta]} \mathbb{E} \left[ |\Gamma_t|^k \right] < +\infty, \quad (57)$$

The proof of Proposition V.5 uses Aldous tightness criterion for real-valued stochastic processes and is given in Appendix A.6.

Both sequences  $(\eta^n)_{n \geq 1}$  and  $(\Gamma^n)_{n \geq 1}$  are tight with continuous limit trajectories. Tightness of  $(\eta^n, \Gamma^n)_{n \geq 1}$  hence follows and we are now in position to give the system satisfied by any limit  $(\eta, \Gamma)$ .

**Theorem V.6.** *Under  $(\mathcal{A}_{TGN})$  and  $(\mathcal{A}_{HöI}^h)$ , for all  $\alpha > 1/2$ , any limit  $(\eta, \Gamma)$  of the sequence  $(\eta^n, \Gamma^n)_{n \geq 1}$  is a solution in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha} \times \mathbb{R})$  of the following system (formulated in  $\mathcal{W}_0^{-3, \alpha} \times \mathbb{R}$ ),*

$$\forall \varphi \in \mathcal{W}_0^{3, \alpha}, \quad \langle \eta_t, \varphi \rangle - \langle \eta_0, \varphi \rangle = \int_0^t \langle \eta_z, L_z \varphi \rangle dz + \int_0^t \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) R \varphi \right\rangle \Gamma_z dz + W_t(R\varphi), \quad (58)$$

$$\Gamma_t = \int_0^t h(t-z) \langle \eta_z, \Psi(\cdot, \bar{\gamma}(z)) \rangle dz + \int_0^t h(t-z) \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \Gamma_z dz + \int_0^t h(t-z) dW_z(\mathbf{1}). \quad (59)$$

The proof of Theorem V.6 consists in proving continuity properties to apply the continuous mapping theorem. It is given in Appendix A.7.

**Remark V.7.** *The linear operator  $L_z$  appearing in (54) and (58) reduces the regularity of the test functions by 1. Hence, if we consider Equation (54) for test functions  $\varphi$  in  $\mathcal{W}_0^{2, \alpha}$  then we must consider  $\eta^n$  as taking values in  $\mathcal{W}_0^{-1, \alpha}$  when dealing with the integral term  $\int_0^t \langle \eta_z^n, L_z \varphi \rangle dz$ . Yet  $(\eta^n)_{n \geq 1}$  is not tight in this space. Thus we consider (54) for test functions in  $\mathcal{W}_0^{3, \alpha}$  so that every term is tight. That is why the limit equation (58) is formulated in  $\mathcal{W}_0^{-3, \alpha}$ . However, the limit process  $\eta$  takes values in the smaller space  $\mathcal{W}_0^{-2, \alpha}$ .*

**Remark V.8.** *The initial condition  $\eta_0$  of the system (58)-(59) is determined by the initial density  $u_0$ . It is an infinite dimensional gaussian random variable. Indeed,  $\eta_0$  is well defined as the limit in  $\mathcal{W}_0^{-3, \alpha}$  of  $\eta_0^n$ . The sequence  $(\eta_0^n)_{n \geq 1}$  is tight in  $\mathcal{W}_0^{-3, \alpha}$*

(it is tight in  $\mathcal{W}_0^{-2,\alpha}$  and there is a continuous embedding of  $\mathcal{W}_0^{-2,\alpha}$  into  $\mathcal{W}_0^{-3,\alpha}$ ) and for any  $\varphi$  in  $\mathcal{W}_0^{3,\alpha}$ , we have the convergence of the real-valued random variables  $\langle \eta_0^n, \varphi \rangle = \sqrt{n} \langle \bar{\mu}_{S_0}^n - u_0, \varphi \rangle$  by applying the standard central limit theorem since the initial conditions are i.i.d.

## V.2 Uniqueness of the limit law

The next step in order to prove convergence of the sequence  $(\eta^n, \Gamma^n)_{n \geq 1}$  is to prove uniqueness of the solutions of the limit system (58)-(59). Since the system is linear, the standard argument is to consider the system satisfied by the difference between two solutions and show that its unique solution is trivial. Let  $(\eta, \Gamma)$  and  $(\hat{\eta}, \hat{\Gamma})$  be two solutions associated with the same “noise”  $W$  and the same initial condition  $\eta_0$ . Denote by  $\tilde{\eta} := \eta - \hat{\eta}$  and  $\tilde{\Gamma} := \Gamma - \hat{\Gamma}$  the differences. Then,  $(\tilde{\eta}, \tilde{\Gamma})$  is a solution of the following system

$$\forall \varphi \in \mathcal{W}_0^{3,\alpha}, \quad \langle \tilde{\eta}_t, \varphi \rangle - \int_0^t \langle \tilde{\eta}_z, L_z \varphi \rangle dz - \int_0^t \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) R \varphi \right\rangle \tilde{\Gamma}_z dz = 0, \quad (60)$$

$$\tilde{\Gamma}_t - \int_0^t h(t-z) \langle \tilde{\eta}_z, \Psi(\cdot, \bar{\gamma}(z)) \rangle dz - \int_0^t h(t-z) \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \tilde{\Gamma}_z dz = 0. \quad (61)$$

The standard follow-up is to use Grönwall’s lemma. Let us show here why it is not sufficient in our case. For instance, assume we want to prove that  $\|\tilde{\eta}\|_{-3,\alpha} = 0$ : heuristically, when applied to (61), Grönwall’s argument gives that  $\|\tilde{\Gamma}_t\|$  is bounded by some locally bounded function of  $t$  times the integral  $\int_0^t \|\tilde{\eta}_z\|_{-3,\alpha} dz$ . However, even if we use this bound for  $\tilde{\Gamma}$  in (60), Grönwall’s argument cannot be applied since the term  $\int_0^t \langle \tilde{\eta}_z, L_z \varphi \rangle dz$  involves  $\|\tilde{\eta}_z\|_{-2,\alpha}$  which is greater than the desired norm  $\|\tilde{\eta}_z\|_{-3,\alpha}$ . This problem cannot be bypassed by upgrading the regularity as we have done before to deal with the fact that the operator  $L_z$  reduces the regularity of the test functions.

Since the main limitation comes from the differential part of the operator  $L_z$ , let us consider  $L_z$  as the sum of the first order differential operator plus a perturbation. More precisely, let  $\mathcal{L} : \varphi \mapsto \varphi'$  and  $G_t : \varphi \mapsto \Psi(\cdot, \bar{\gamma}(t)) R \varphi$  so that  $L_t = \mathcal{L} + G_t$ . Let us present here the heuristics behind the argument we use to bypass the issue induced by the differential operator  $\mathcal{L}$ : instead of studying the time derivative  $\frac{d}{dt} \langle \tilde{\eta}_t, \varphi \rangle$  in (60), the idea is to find some family of test functions  $(\varphi_t)_{t \geq 0}$  such that  $\langle \tilde{\eta}_t, \frac{d}{dt} \varphi_t \rangle = -\langle \tilde{\eta}_t, \mathcal{L} \varphi_t \rangle$ ; thus the differential operator  $\mathcal{L}$  vanishes in  $\frac{d}{dt} \langle \tilde{\eta}_t, \varphi_t \rangle$  and Grönwall’s argument can be applied.

More precisely, let us introduce the shift operators  $\tau_t : \varphi \mapsto \varphi(\cdot + t)$  for all  $t \geq 0$ . Notice that these shift operators are linked with the method of characteristics applied to a transport equation with constant speed equal to 1 which is exactly the dynamics described by the differential operator  $\mathcal{L}$ . Below are given some bounds for the operators  $\mathcal{L}$ ,  $G_t$  and  $\tau_t$  when acting on the space  $\mathcal{C}_b^4$ .

**Lemma V.9.** *Let  $\varphi$  be in  $\mathcal{C}_b^4$ . Assume that  $t \mapsto \|\Psi(\cdot, \bar{\gamma}(t))\|_{\mathcal{C}_b^4}$  is locally bounded. Then,  $\|\mathcal{L}\varphi\|_{\mathcal{C}_b^3} \leq \|\varphi\|_{\mathcal{C}_b^4}$ , for all  $t \geq 0$ ,  $\|\tau_t \varphi\|_{\mathcal{C}_b^4} = \|\varphi\|_{\mathcal{C}_b^4}$  and*

$$t \mapsto \frac{\|G_t \varphi\|_{\mathcal{C}_b^4}}{\|\varphi\|_{\mathcal{C}_b^4}} \quad \text{is locally bounded.}$$

*Proof.* The first two assertions follow from the definition of the norms  $\|\cdot\|_{\mathcal{C}_b^k}$ . The third and last one follows from Leibniz rule.  $\square$

**Remark V.10.** From now on, the test functions are considered in  $\mathcal{C}_b^4$ . Thus, we prove that  $\eta$  is characterized by the limit equation as a process with values in the dual space  $\mathcal{C}_b^{-4}$ . Nevertheless, since  $\mathcal{C}_b^4$  is dense in  $\mathcal{W}_0^{3,\alpha}$ , it is also characterized by the limit equation as a process with values in  $\mathcal{W}_0^{-3,\alpha}$  for instance.

Let  $t \geq t'$  and  $s$  in  $\mathbb{R}$ . Then,

$$\int_{t'}^t \tau_{t-z} \varphi'(s) dz = \int_{t'}^t \varphi'(s+t-z) dz = \varphi(s+t-t') - \varphi(s) = \tau_{t-t'} \varphi(s) - \varphi(s).$$

Moreover, since  $\tau_t$  and  $\mathcal{L}$  commute, one has

$$\tau_{t-t'} \varphi(s) - \varphi(s) = \int_{t'}^t \mathcal{L}(\tau_{t-z} \varphi)(s) dz.$$

Yet, Lemma V.9 gives that  $\|\mathcal{L}(\tau_{t-z} \varphi)\|_{\mathcal{C}_b^3} \leq \|\varphi\|_{\mathcal{C}_b^4}$  thus  $\int_{t'}^t \mathcal{L} \tau_{t-z} \varphi dz$  makes sense as a Bochner integral in  $\mathcal{C}_b^3$  as soon as  $\varphi$  is in  $\mathcal{C}_b^4$ . Hence, in the proof below we use the following statement: for all  $\varphi$  in  $\mathcal{C}_b^4$ ,

$$\tau_{t-t'} \varphi - \varphi = \int_{t'}^t \mathcal{L}(\tau_{t-z} \varphi) dz, \quad \text{as points in } \mathcal{C}_b^3. \quad (62)$$

**Proposition V.11.** Under  $(\mathcal{A}_{CLT})$ , the system (58)-(59) has no more than one solution in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha} \times \mathbb{R})$  once the initial condition  $\eta_0$  and the “noise”  $W$  are fixed.

*Proof.* Let  $(\eta, \Gamma)$  and  $(\hat{\eta}, \hat{\Gamma})$  be two solutions of (58)-(59) in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha} \times \mathbb{R})$  associated with the the same “noise”  $W$  and the same initial condition  $\eta_0$ . Denote by  $\tilde{\eta} := \eta - \hat{\eta}$  and  $\tilde{\Gamma} := \Gamma - \hat{\Gamma}$  the differences. Since  $\alpha > 1/2$ , we have  $\mathcal{W}_0^{-2,\alpha} \subset \mathcal{C}_b^{-4}$  (remind (33)) so  $\tilde{\eta}$  belongs to  $\mathcal{C}_b^{-4}$  and we will prove that  $\|\tilde{\eta}\|_{\mathcal{C}_b^{-4}} = 0$ .

Starting from (61), one has

$$|\tilde{\Gamma}_t| \leq h_\infty(t) \|\Psi(\cdot, \bar{\gamma}(t))\|_{\mathcal{C}_b^4} \int_0^t \|\tilde{\eta}_z\|_{\mathcal{C}_b^{-4}} dz + h_\infty(t) \text{Lip}(\Psi) \int_0^t |\tilde{\Gamma}|_z dz,$$

and Lemma B.1 gives  $|\tilde{\Gamma}_t| \lesssim_t \int_0^t \|\tilde{\eta}_z\|_{\mathcal{C}_b^{-4}} dz$ . Now, let  $\varphi$  be in  $\mathcal{C}_b^4$  and use (62) and the fact that  $\tilde{\eta}$  is in  $\mathcal{W}_0^{-2,\alpha} \subset \mathcal{C}_b^{-3}$  to get  $\langle \tilde{\eta}_t, \varphi \rangle = D_1 - D_2$  where

$$\begin{cases} D_1 := \int_0^t \langle \tilde{\eta}_{t'}, (\mathcal{L} + G_{t'}) (\tau_{t-t'} \varphi) \rangle dt' + \int_0^t \left\langle u_{t'}, R \tau_{t-t'} \varphi \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t')) \right\rangle \tilde{\Gamma}_{t'} dt'. \\ D_2 := \int_0^t \left\langle \tilde{\eta}_{t'}, (\mathcal{L} + G_{t'}) \left( \int_{t'}^t \mathcal{L}(\tau_{t-z} \varphi) dz \right) \right\rangle dt' \\ \quad + \int_0^t \left\langle u_{t'}, R \int_{t'}^t \mathcal{L}(\tau_{t-z} \varphi) dz \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t')) \right\rangle \tilde{\Gamma}_{t'} dt'. \end{cases} \quad (63)$$

The linearity of the operators allows to write  $D_2 = D_{2,A} + D_{2,B}$  with

$$\begin{cases} D_{2,A} := \int_0^t \int_{t'}^t \langle \tilde{\eta}_{t'}, (\mathcal{L} + G_{t'}) (\mathcal{L}(\tau_{t-z} \varphi)) \rangle dz dt' \\ D_{2,B} := \int_0^t \int_{t'}^t \left\langle u_{t'}, R \mathcal{L} \tau_{t-z} \varphi \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t')) \right\rangle \tilde{\Gamma}_{t'} dz dt'. \end{cases}$$

Then, the idea is to use Fubini’s theorem to exchange the two integrals  $\int_0^t$  and  $\int_{t'}^t$ .

On the one hand,

$$\int_0^t \int_{t'}^t |\langle \tilde{\eta}_{t'}, (\mathcal{L} + G_{t'}) (\mathcal{L}(\tau_{t-z}\varphi)) \rangle| dz dt' \leq \int_0^t \int_{t'}^t \|\tilde{\eta}_{t'}\|_{\mathcal{C}_b^{-2}} \|(\mathcal{L} + G_{t'}) (\mathcal{L}(\tau_{t-z}\varphi))\|_{\mathcal{C}_b^2} dz dt'.$$

Notice that  $\sup_{t' \in [0, t]} \|\tilde{\eta}_{t'}\|_{\mathcal{C}_b^{-2}} \leq C(\sup_{t' \in [0, t]} \|\eta_{t'}\|_{-2, \alpha} + \sup_{t' \in [0, t]} \|\hat{\eta}_{t'}\|_{-2, \alpha}) < +\infty$  since  $\eta$  and  $\hat{\eta}$  takes values in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha})$  and that, thanks to Lemma V.9, for all  $t' \leq t$ ,  $\|(\mathcal{L} + G_{t'}) (\mathcal{L}(\tau_{t-z}\varphi))\|_{\mathcal{C}_b^2} \lesssim_t \|\varphi\|_{\mathcal{C}_b^4} < +\infty$ . Hence, Fubini's theorem gives

$$D_{2,A} = \int_0^t \int_0^z \langle \tilde{\eta}_{t'}, (\mathcal{L} + G_{t'}) (\mathcal{L}(\tau_{t-z}\varphi)) \rangle dt' dz. \quad (64)$$

On the other hand,

$$\int_0^t \int_{t'}^t \left| \left\langle u_{t'}, R\mathcal{L}\tau_{t-z}\varphi \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t')) \right\rangle \right| |\tilde{\Gamma}_{t'}| dz dt' \leq 2\text{Lip}(\Psi) \int_0^t \int_{t'}^t \|\mathcal{L}(\tau_{t-z}\varphi)\|_{\infty} |\tilde{\Gamma}_{t'}| dz dt'.$$

Remark that  $\|\mathcal{L}(\tau_{t-z}\varphi)\|_{\infty} \leq \|\varphi\|_{\mathcal{C}_b^4}$  and  $\sup_{t' \in [0, t]} |\tilde{\Gamma}_{t'}| \leq \sup_{t' \in [0, t]} |\Gamma_{t'}| + \sup_{t' \in [0, t]} |\hat{\Gamma}_{t'}| < +\infty$  since  $\Gamma$  and  $\hat{\Gamma}$  takes values in  $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ . Hence, Fubini's theorem gives

$$D_{2,B} = \int_0^t \int_0^z \left\langle u_{t'}, R\mathcal{L}\tau_{t-z}\varphi \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t')) \right\rangle \tilde{\Gamma}_{t'} dt' dz. \quad (65)$$

Now, for any  $z$  in  $[0, t]$ , Equation (60) with  $\varphi = \mathcal{L}(\tau_{t-z}\varphi)$  (it is a valid test function since it belongs to  $\mathcal{C}_b^3 \subset \mathcal{W}_0^{3, \alpha}$ ) gives

$$\begin{aligned} \langle \tilde{\eta}_z, \mathcal{L}(\tau_{t-z}\varphi) \rangle &= \int_0^z \langle \tilde{\eta}_{t'}, (\mathcal{L} + G_{t'}) (\mathcal{L}(\tau_{t-z}\varphi)) \rangle dt' \\ &\quad + \int_0^z \left\langle u_{t'}, R\mathcal{L}\tau_{t-z}\varphi \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t')) \right\rangle \tilde{\Gamma}_{t'} dt'. \end{aligned}$$

Gathering the equation above with (64) and (65) gives

$$D_2 = \int_0^t \langle \tilde{\eta}_z, \mathcal{L}(\tau_{t-z}\varphi) \rangle dz,$$

which is exactly the term driven by  $\mathcal{L}$  in the definition of  $D_1$  (63) so that, coming back to  $D_1 - D_2$ , we have

$$D_1 - D_2 = \langle \tilde{\eta}_t, \varphi \rangle = \int_0^t \langle \tilde{\eta}_{t'}, G_{t'}(\tau_{t-t'}\varphi) \rangle dt' + \int_0^t \left\langle u_{t'}, R\tau_{t-t'}\varphi \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t')) \right\rangle \tilde{\Gamma}_{t'} dt'.$$

Hence, using the bound we proved on  $\tilde{\Gamma}$ , we have for all  $\varphi$  in  $\mathcal{C}_b^4$ ,

$$|\langle \tilde{\eta}_t, \varphi \rangle| \lesssim_t \int_0^t \|\tilde{\eta}_{t'}\|_{\mathcal{C}_b^{-4}} \|G_{t'}(\tau_{t-t'}\varphi)\|_{\mathcal{C}_b^4} dt' + 2\|\varphi\|_{\mathcal{C}_b^4} \text{Lip}(\Psi) \int_0^t \int_0^{t'} \|\tilde{\eta}_z\|_{\mathcal{C}_b^{-4}} dz dt',$$

and so  $\|\tilde{\eta}_t\|_{\mathcal{C}_b^{-4}} \lesssim_t \int_0^t \|\tilde{\eta}_{t'}\|_{\mathcal{C}_b^{-4}} dt'$  and Lemma B.1 gives that for all  $t \geq 0$ ,  $\|\tilde{\eta}_t\|_{\mathcal{C}_b^{-4}} = 0$  thus  $|\tilde{\Gamma}_t| = 0$  thanks to the bound we proved on  $\tilde{\Gamma}$ . Finally, since  $\mathcal{C}_b^4$  is dense in  $\mathcal{W}_0^{2, \alpha}$ , we have  $\|\tilde{\eta}_t\|_{-2, \alpha} = 0$ . Thus, we have  $(\eta, \Gamma) = (\hat{\eta}, \hat{\Gamma})$  in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha} \times \mathbb{R})$ .  $\square$

We are now in position to conclude with the convergence of  $(\eta^n, \Gamma^n)_{n \geq 1}$ .

**Theorem V.12.** *Under  $(\mathcal{A}_{CLT})$ , for any  $\alpha > 1/2$ , the sequence  $(\eta^n, \Gamma^n)_{n \geq 1}$  converges in law in  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha} \times \mathbb{R})$  to the unique solution of the system (58)-(59) in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha} \times \mathbb{R})$ .*

*Proof.* Since  $(\eta^n, \Gamma^n)_{n \geq 1}$  is tight (Theorem IV.11 and Proposition V.5), let  $(\eta, \Gamma)$  be a limit point. According to Theorem V.6,  $(\eta, \Gamma)$  is a solution of the limit system (58)-(59) in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha} \times \mathbb{R})$ . Finally, the law of  $(\eta, \Gamma)$  is uniquely characterized by the limit system (Proposition V.11 gives path-wise uniqueness and so Yamada-Watanabe theorem gives weak uniqueness by the same argument as [35, Theorem IX.1.7(i)]) and uniqueness of the limit law implies convergence of  $(\eta^n, \Gamma^n)_{n \geq 1}$ .  $\square$

**Remark V.13.** *As mentioned in the introduction, considering processes over finite time horizons would have lead to equivalent results. This claim is based on the fact that the limit equation (58) is independent of the values of the test function  $\varphi$  outside the support  $K_t$  of  $\eta_t^n$ . Indeed,*

- *on the one hand, the test function  $\varphi$  appears in the drift term, more precisely  $\int_0^t \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) R \varphi \right\rangle \Gamma_z dz$ , evaluated against the measure  $u_z$  which is supported in  $K_t$ ;*
- *on the other hand, the covariance structure of the Gaussian process  $W$  implies this independence property for  $W_t(R\varphi)$ .*

*In that sense, the convergence stated for the whole positive time line  $\mathbb{R}_+$  in Theorem V.12 implies that the central limit theorem also holds true for the process  $(\eta_t^n)_{0 \leq t \leq \theta}$  as taking values in the dual of a standard Sobolev space of functions supported by  $K_\theta$ . Conversely, the limit equation is consistent in time in the sense that one can recover our result by sticking together the CLTs obtained for the finite time horizon processes  $(\eta_t^n)_{0 \leq t \leq \theta}$ .*

## VI Application to the “almost” derivation of an SPDE

This section focuses on a system of stochastic partial differential equations (SPDE), introduced and studied in [13] where some qualitative properties are discussed. The SPDE is a noisy version of the PDE system (4) and is expected to be a more precise approximation of the age-dependent Hawkes processes in a mean-field framework. The SPDE system associated with the system size  $n$  is the following

$$\begin{cases} \frac{\partial \tilde{u}^n(t, s)}{\partial t} + \frac{\partial \tilde{u}^n(t, s)}{\partial s} + \Psi(s, X_t^n) \tilde{u}^n(t, s) + \sqrt{\frac{\Psi(s, X_t^n) \tilde{u}^n(t, s)}{n}} \zeta(t, s) = 0, \\ \tilde{u}^n(t, 0) = \int_{s \in \mathbb{R}} \Psi(s, X_t^n) \tilde{u}^n(t, s) + \sqrt{\frac{\Psi(s, X_t^n) \tilde{u}^n(t, s)}{n}} \zeta(t, s) ds, \end{cases} \quad (66)$$

where for all  $t \geq 0$ ,  $X_t^n = \int_0^t h(t-z) \tilde{u}^n(z, 0) dz$  and  $\zeta(t, s)$  is a Gaussian space-time white-noise. The important thing to note about  $\zeta$  is that the  $\mathcal{W}_0^{-2,\alpha}$ -valued process defined by, for all  $t \geq 0$  and  $\varphi$  in  $\mathcal{W}_0^{2,\alpha}$ ,

$$\int_0^t \int_0^{+\infty} \varphi(s) \sqrt{\Psi(s, \bar{\gamma}(z))} u(z, s) \zeta(z, s) ds dz,$$

is a Gaussian process with the same law as  $W$  defined in Definition V.1.

Hence, at a first sight, there are some similarities between the system above and the limit system obtained for the fluctuation process, i.e. (58)-(59). Let us give here some heuristics: assume that  $u$ , the solution of the PDE system (4), and  $\tilde{u}^n$  are close to each other, and similarly for the auxiliary variables  $X(t)$  and  $X_t^n$ , then



- the “non-noisy” spiking dynamics term  $\Psi(s, X_t^n) \tilde{u}^n(t, s)$  is close to the mean-field spiking dynamics, appearing in the operator  $L_t$  in the limit system (58)-(59), modulo an error term which is expected to appear, in a linear approximation, as the term involving the derivative  $\frac{\partial \Psi}{\partial y}$  in (58);
- the covariance structure of the Gaussian process  $W$ , appearing in the limit system (58)-(59), is close to the covariance structure of the noise term appearing in the SPDE system above.

What is proposed in this section is to consider the second-order approximation of the empirical measure given by the central limit theorem, namely  $\hat{u}_t^n = u_t + n^{-1/2} \eta_t$  where  $u_t = u(t, \cdot)$  is the probability distribution solution of (4), and show that is an “almost” solution of the SPDE system in some sense defined below. Up to our knowledge, this kind of result is novel and deserves to be developed in this article.

Let us remind that this section is devoted to an application of the CLT so Assumption ( $\mathcal{A}_{\text{CLT}}$ ) is supposed to hold true below. Furthermore, the stronger assumption that  $\Psi$  is in  $\mathcal{C}_b^4$  is made.

## VI.1 Theoretical frame for the SPDE

Up to our knowledge, there is no theoretical frame well established for the SPDE system (66). A pathwise notion of solution seems to be hard to handle because of the square root term appearing in front of the Gaussian white noise. In particular, it is not trivial to show that the argument of the square root remains non negative. That is why we propose the following notion of solution.

**Definition VI.1.** *The measure-valued process  $(\tilde{u}_t^n)_{t \geq 0}$  is a solution of (66) if it satisfies: for all  $\varphi$  in  $\mathcal{C}_b^\infty$ ,*

$$\langle \tilde{u}_t^n, \varphi \rangle - \langle \tilde{u}_0^n, \varphi \rangle = \int_0^t \langle \tilde{u}_z^n, \tilde{L}_z^n \varphi \rangle dz + \tilde{W}_t^n(R\varphi), \quad (67)$$

where  $\tilde{W}^n$  is a Gaussian process with Doob-Meyer process given by,

$$\langle \llbracket \tilde{W}^n \rrbracket_t(\varphi_1), \varphi_2 \rangle = \frac{1}{n} \int_0^t \langle \tilde{u}_z^n, \varphi_1 \varphi_2 \Psi(\cdot, \tilde{\gamma}_z^n) \rangle dz, \quad (68)$$

and  $\tilde{L}_z^n \varphi := \varphi' + \Psi(\cdot, \tilde{\gamma}_z^n) \varphi$  with

$$\tilde{\gamma}_t^n = \int_0^t h(t-z) \langle \tilde{u}_z^n, \Psi(\cdot, \tilde{\gamma}_z^n) \rangle dz + \int_0^t h(t-z) d\tilde{W}_z^n(\mathbf{1}). \quad (69)$$

The well-posedness of such definition is not addressed here. We only stress the fact that  $\hat{u}^n$  is an “almost” solution in some sense related to Definition VI.1.

## VI.2 Weak sense dynamics for the second-order approximation

To catch the dynamics of  $\hat{u}^n$ , we somehow want to add up the dynamics of  $u$ , given by the PDE system (4), and the dynamics of  $\eta$ , given by the limit system. These two are formulated by different means, the main difference being that the PDE formulation involves (in the weak sense) bivariate test functions whereas the formulation for  $\eta$  involves univariate test functions: we turn to the second one in order to get a system like (67)-(69).

On the one hand, the dynamics of  $\eta$  is given by the limit equation (58) that we remind here: for all  $\varphi$  in  $\mathcal{W}_0^{3,\alpha}$ ,

$$\langle \eta_t, \varphi \rangle - \langle \eta_0, \varphi \rangle = \int_0^t \langle \eta_z, L_z \varphi \rangle dz + \int_0^t \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) R \varphi \right\rangle \Gamma_z dz + W_t(R\varphi).$$

On the other hand, the dynamics of  $u$  is given (see [9, Theorem III.5.]) by the weak sense formulation of the PDE system (4) (which is driven by the generator  $L_t$ ): for all  $\phi$  in  $\mathcal{C}_{c,b}^\infty(\mathbb{R}_+^2)$ ,

$$\int_{\mathbb{R}_+^2} \left( \frac{\partial \phi}{\partial t}(t, s) + L_t(\phi(t, \cdot)) \right) u(t, s) dt ds + \int_{\mathbb{R}_+} \phi(0, s) u_0(s) ds = 0 \quad (70)$$

where the test function space  $\mathcal{C}_{c,b}^\infty(\mathbb{R}_+^2)$  is defined as follows,

$$\mathcal{C}_{c,b}^\infty(\mathbb{R}_+^2) \left\{ \begin{array}{l} \text{The function } \phi \text{ belongs to } \mathcal{C}_{c,b}^\infty(\mathbb{R}_+^2) \text{ if} \\ \bullet \phi \text{ is continuous, uniformly bounded,} \\ \bullet \phi \text{ has uniformly bounded derivatives of every order,} \\ \bullet \text{ there exists } T > 0 \text{ such that } \phi(t, s) = 0 \text{ for all } t > T \text{ and } s \geq 0. \end{array} \right.$$

Then, taking  $\phi(t, s)$  that converges to a product function of the form  $\varphi(s)\mathbb{1}_{t \leq T}$ , we get that for all  $\varphi$  in  $\mathcal{C}_b^\infty(\mathbb{R}_+)$ ,

$$\langle u_T, \varphi \rangle - \langle u_0, \varphi \rangle = \int_0^T \langle u_t, L_t \varphi \rangle dt. \quad (71)$$

Combining (71) with the limit equation (58), we prove that  $\hat{u}^n$  satisfies: for all  $\varphi$  in  $\mathcal{C}_b^\infty(\mathbb{R}_+)$ ,

$$\begin{aligned} \langle \hat{u}_t^n, \varphi \rangle - \langle \hat{u}_0^n, \varphi \rangle &= \int_0^t \langle \hat{u}_z^n, L_z \varphi \rangle dz + n^{-1/2} \int_0^t \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) R \varphi \right\rangle \Gamma_z dz \\ &\quad + n^{-1/2} W_t(R\varphi). \end{aligned}$$

This last equation can be rewritten, with anything new but some notation, as a system in the flavour of (67)-(69) as stated in the proposition below.

**Proposition VI.2.** *The process  $\hat{u}^n$  satisfies: for all  $\varphi$  in  $\mathcal{C}_b^\infty$ ,*

$$\langle \hat{u}_t^n, \varphi \rangle - \langle \hat{u}_0^n, \varphi \rangle = \int_0^t \left\langle \hat{u}_z^n, \hat{L}_z^n \varphi \right\rangle dz + \hat{W}_t^n(R\varphi) + r_t^n(\varphi), \quad (72)$$

where  $\hat{W}^n$  is a Gaussian process with Doob-Meyer process given by,

$$\text{DM}_t(\varphi_1, \varphi_2) := \left\langle \ll \hat{W}^n \gg_t(\varphi_1), \varphi_2 \right\rangle = \frac{1}{n} \int_0^t \langle u_z, \varphi_1 \varphi_2 \Psi(\cdot, \bar{\gamma}(z)) \rangle dz, \quad (73)$$

and  $\hat{L}_z^n \varphi := \varphi' + \Psi(\cdot, \hat{\gamma}_t^n) \varphi$  with

$$\hat{\gamma}_t^n = \int_0^t h(t-z) \langle \hat{u}_z^n, \Psi(\cdot, \hat{\gamma}_z^n) \rangle dz + \int_0^t h(t-z) d\hat{W}_z^n(\mathbf{1}). \quad (74)$$

The following notation is used above:  $\hat{W}^n := n^{-1/2} W$  where  $W$  is the Gaussian process of Definition V.1 and

$$r_t^n(\varphi) := \int_0^t \left\langle \hat{u}_z^n, (L_z - \hat{L}_z^n) \varphi \right\rangle + \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) R \varphi \right\rangle \frac{\Gamma_z}{\sqrt{n}} dz. \quad (75)$$

**Remark VI.3.** *The process  $(\hat{\gamma}_t^n)_{t \geq 0}$  is characterized by the fixed point equation (74) (see Lemma A.1).*

Comparing (67)-(69) with (72)-(74), the only differences between the two systems are, the additional term  $r_t^n(\varphi)$  in (72), and the substitution of the Doob-Meyer process associated with the noise term: it should be given by

$$\hat{\text{DM}}_t(\varphi_1, \varphi_2) := \frac{1}{n} \int_0^t \langle \hat{u}_z^n, \varphi_1 \varphi_2 \Psi(\cdot, \hat{\gamma}_z^n) \rangle dz. \quad (76)$$

As stated below, these two differences are negligible, as  $n \rightarrow +\infty$ , with respect to the other terms of the system that are at least of order  $n^{-1/2}$ .

**Proposition VI.4.** *The distribution of the process  $\hat{u}^n$  is an “almost” solution of (67)-(69) in the sense that: the rest term in (72) is negligible since*

$$\mathbb{E} [|r_t^n(\varphi)|] \lesssim_t n^{-1} \|\varphi\|_{2,1},$$

and the covariance structures are “almost” the same since

$$\mathbb{E} \left[ \left| \text{DM}_t(\varphi_1, \varphi_2) - \hat{\text{DM}}_t(\varphi_1, \varphi_2) \right| \right] \lesssim_t n^{-3/2} \|\varphi_1\|_{3,1} \|\varphi_2\|_{3,1}.$$

The proof of Proposition VI.4 relies on some refined versions of already established estimates and is given in Appendix A.8. The main difficulty and difference in comparison with the preceding sections is that  $\hat{u}_t^n$  is not a probability measure, unlike  $u_t$  and  $\bar{\mu}_t^n$ . However, this difficulty is bypassed in the proof by means of the following estimates.

**Lemma VI.5.** *For any  $\varphi$  in  $\mathcal{C}_b^2(\mathbb{R}_+)$ , we have*

$$|\langle \hat{u}_t^n, \varphi \rangle| \leq \left( 1 + \frac{C}{\sqrt{n}} \|\eta_t\|_{-2,1} \right) \|\varphi\|_{\mathcal{C}_b^2}.$$

We mainly use this estimate with the intensity function  $\Psi$ : for all  $y_1, y_2$ ,

$$\begin{cases} |\langle \hat{u}_t^n, \Psi(\cdot, y_1) \rangle| \leq \left( 1 + \frac{C}{\sqrt{n}} \|\eta_t\|_{-2,1} \right) \|\Psi\|_{\mathcal{C}_b^2}, \\ |\langle \hat{u}_t^n, \Psi(\cdot, y_1) - \Psi(\cdot, y_2) \rangle| \leq \left( 1 + \frac{C}{\sqrt{n}} \|\eta_t\|_{-2,1} \right) \|\Psi\|_{\mathcal{C}_b^3} |y_1 - y_2|. \end{cases} \quad (77)$$

The second line of (77) is very useful to use some kind of Lipschitz control even when integrating with respect to  $\hat{u}_t^n$ , which is not as direct as in the case when the integration is done with respect to a probability measure.

*Proof.* The first assertion is a direct consequence of the definition of  $\hat{u}^n$  and the embedding (30). Since  $\Psi$  is in  $\mathcal{C}_b^4(\mathbb{R}_+ \times \mathbb{R})$ , the next ones are direct from

$$\begin{cases} \|\Psi(\cdot, y_1)\|_{\mathcal{C}_b^2(\mathbb{R}_+)} \leq \|\Psi\|_{\mathcal{C}_b^2(\mathbb{R}_+ \times \mathbb{R})}, \\ \|\Psi(\cdot, y_1) - \Psi(\cdot, y_2)\|_{\mathcal{C}_b^2(\mathbb{R}_+)} \leq \|\Psi\|_{\mathcal{C}_b^3(\mathbb{R}_+ \times \mathbb{R})} |y_1 - y_2|. \end{cases}$$

□

To conclude, let us remind that we have shown how a second-order approximation of the empirical measure  $\bar{\mu}_t^n$ , namely  $\hat{u}_t^n$  since  $\sqrt{n}(\bar{\mu}_t^n - \hat{u}_t^n)$  goes to 0 as a consequence of Theorem V.12, can be considered as an “almost” solution of the SPDE (66). The next step would be to prove that the/any solution of the SPDE is a second order approximation of  $\bar{\mu}_t^n$ . To address such a question, we would first need a suitable theoretical framework to treat the well-posedness of the SPDE. This is the subject of a future work.

## A Proofs

### A.1 Proofs linked with Proposition III.1

**Proof of (26).** For simplicity, we show that, for every  $m \leq n$ , there exists a constant  $C$  which is independent of  $n$ ,  $p$  and  $m$  such that

$$\mathbb{E} \left[ \left( \frac{1}{m} \sum_{j=1}^m (\Delta_{t-}^{n,j})^p \right)^k \right] \leq C \left( \sum_{k'=1}^{k-1} m^{k'-k} \varepsilon_n^{(k',pk)}(t) + \varepsilon_n^{(k,p)}(t) \right), \quad (78)$$

from which (26) follows by choosing  $m = \lfloor \frac{n}{k} \rfloor$ .

Let us recall the multinomial formula using multi-indices  $\mathbf{q} = (q_1, \dots, q_m)$ ,

$$\left( \frac{1}{m} \sum_{i=1}^m x_i \right)^k = \frac{1}{m^k} \sum_{|\mathbf{q}|=k} \binom{k}{\mathbf{q}} \prod_{i=1}^m x_i^{q_i},$$

where  $|\mathbf{q}| = \sum_{i=1}^m q_i$ . Denote by  $k(\mathbf{q})$  the number of strictly positive indices in  $\mathbf{q}$ . Since the  $q_i$ 's are integers,  $|\mathbf{q}| = k$  implies  $k(\mathbf{q}) \leq k$ . First, let us remark that, for all  $k' = 1, \dots, k$ , the number of multi-indices  $\mathbf{q}$  such that  $k(\mathbf{q}) = k'$  and  $|\mathbf{q}| = k$  is bounded by  $p(k', k) m^{k'}$  with  $p(k', k) := \binom{k'-1}{k-1}$  being the number of partitions of  $k$  into exactly  $k'$  parts. Indeed, the vector consisting in the  $k'$  strictly positive indices forms a partition of  $k$  and there are at most  $m^{k'}$  ways to complete it by  $m - k'$  zeros to build a vector of length  $m$ .

Then, using the exchangeability of the processes  $\Delta^{n,j}$ , we have

- if  $k(\mathbf{q}) = k$ , then all the positive  $q_i$ 's are equal to one and  $\mathbb{E}[\prod_{i=1}^n ((\Delta_{t-}^{n,i})^p)^{q_i}] = \varepsilon_n^{(k,p)}(t)$ ,
- if  $k(\mathbf{q}) < k$ , we can bound all the positive  $q_i$ 's by  $k$  so that  $\mathbb{E}[\prod_{i=1}^n ((\Delta_{t-}^{n,i})^p)^{q_i}] \leq \varepsilon_n^{(k(\mathbf{q}),pk)}(t)$ .

Hence, using that  $\binom{k}{\mathbf{q}} \leq k!$ , (78) holds with  $C = \max_{k'=1, \dots, k} p(k', k) k!$  for instance.

**Proof of (27).** Let us first recall that  $\xi_n^{(k)}(t) = \mathbb{E}[|\gamma_t^n - \bar{\gamma}(t)|^k]$  where  $\gamma_t^n$  and  $\bar{\gamma}(t)$  are respectively defined below (12) and in (11). By convexity of the function  $x \mapsto |x|^k$  (remind that  $k \geq 2$ ), let us consider the decomposition

$$\xi_n^{(k)}(t) \leq 4^{k-1} (A^n(t) + B^n(t) + C^n(t) + D^n(t)) \quad (79)$$

where

$$\begin{cases} A^n(t) := \mathbb{E} \left[ \left| \int_0^t h(t-z) \frac{1}{n} \sum_{j=1}^n (N^{n,j}(dz) - \lambda_z^{n,j} dz) \right|^k \right], \\ B^n(t) := \mathbb{E} \left[ \left| \int_0^t h(t-z) \frac{1}{n} \sum_{j=1}^n (\Psi(S_{z-}^{n,j}, \gamma_z^n) - \Psi(\bar{S}_{z-}^j, \gamma_z^n)) dz \right|^k \right], \\ C^n(t) := \mathbb{E} \left[ \left| \int_0^t h(t-z) \frac{1}{n} \sum_{j=1}^n (\Psi(\bar{S}_{z-}^j, \gamma_z^n) - \Psi(\bar{S}_{z-}^j, \bar{\gamma}(z))) dz \right|^k \right], \\ D^n(t) := \mathbb{E} \left[ \left| \int_0^t h(t-z) \frac{1}{n} \sum_{j=1}^n (\bar{\lambda}_z^j - \bar{\lambda}(z)) dz \right|^k \right]. \end{cases}$$

Recall that  $\lambda_z^{n,j} = \Psi(S_{z-}^{n,j}, \gamma_z^n)$  and  $\bar{\lambda}_z^j = \Psi(\bar{S}_{z-}^j, \bar{\gamma}(z))$ .

- Study of  $A^n(t)$ . Fix  $t$  and consider the martingale  $(M_x^{n,t})_{x \geq 0}$  defined, for all  $x \geq 0$ , by

$$M_x^{n,t} := \int_0^x h(t-z) \frac{1}{n} \sum_{j=1}^n (N^{n,j}(dz) - \lambda_z^{n,j} dz).$$

Its quadratic variation is  $[M^{n,t}]_x = n^{-2} \int_0^x h(t-z)^2 \sum_{j=1}^n N^{n,j}(dz)$ . Yet, Assumption  $(\mathcal{A}_\infty^h)$  implies that

$$[M^{n,t}]_t \leq n^{-2} h_\infty(t)^2 \sum_{j=1}^n N_t^{n,j}.$$

Using the convexity of the power function (since  $k/2 \geq 1$ ) and exchangeability, one has

$$\mathbb{E} \left[ [M^{n,t}]_t^{k/2} \right] \lesssim_{(t,k)} n^{-k} \mathbb{E} \left[ \left| \sum_{j=1}^n N_t^{n,j} \right|^{k/2} \right] \lesssim_{(t,k)} n^{-k} n^{k/2} \mathbb{E} \left[ |N_t^{n,1}|^{k/2} \right].$$

Yet, the intensity of  $N^{n,1}$  is bounded by  $\|\Psi\|_\infty$  so  $N^{n,1}$  is stochastically dominated by a Poisson process with intensity  $\|\Psi\|_\infty$ . Hence,  $\mathbb{E}[|N_t^{n,1}|^{k/2}] \leq \mathbb{E}[\text{Pois}(t|\|\Psi\|_\infty)^{k/2}]$  where  $\text{Pois}(t|\|\Psi\|_\infty)$  is a Poisson variable with parameter  $t|\|\Psi\|_\infty$ . This last expectation is bounded uniformly in  $n$  by a locally bounded function of the time  $t$ . Then, Burkholder-Davis-Gundy inequality [38, p. 894] gives

$$A_1^n(t) = \mathbb{E} \left[ |M_t^{n,t}|^k \right] \leq \mathbb{E} \left[ [M^{n,t}]_t^{k/2} \right] \lesssim_{(t,k)} n^{-k/2}.$$

- Study of  $B^n(t)$ . Here, we use the fact that  $S_{z-}^{n,j} = \bar{S}_{z-}^j$  with high probability and more precisely we recover the quantities  $\varepsilon_n^{(k,p)}$  that we want to control. Using the convexity of the power function, Assumption  $(\mathcal{A}_\infty^h)$  and denoting  $x_z^j = \Psi(S_{z-}^{n,j}, \gamma_z^n) - \Psi(\bar{S}_{z-}^j, \gamma_z^n)$  we have

$$B^n(t) \leq h_\infty(t)^k t^{k-1} \int_0^t \mathbb{E} \left[ \left| \frac{1}{n} \sum_{j=1}^n x_z^j \right|^k \right] dz.$$

Yet,  $|x_z^j|$  is bounded by  $\|\Psi\|_\infty \mathbb{1}_{\{S_{z-}^{n,j} \neq \bar{S}_{z-}^j\}} \leq \|\Psi\|_\infty \Delta_{z-}^{n,j}$ . Hence, using (78) with  $p = 1$  and  $m = n$ ,

$$\begin{aligned} B^n(t) &\leq \|\Psi\|_\infty^k h_\infty(t)^k t^{k-1} \int_0^t \mathbb{E} \left[ \left| \frac{1}{n} \sum_{j=1}^n \Delta_{z-}^{n,j} \right|^k \right] dz \\ &\leq \|\Psi\|_\infty^k h_\infty(t)^k C t^{k-1} \int_0^t \left( \sum_{k'=1}^{k-1} n^{k'-k} \varepsilon_n^{(k',k)}(z) + \varepsilon_n^{(k,1)}(z) \right) dz \\ &\leq \|\Psi\|_\infty^k h_\infty(t)^k C t^k \left( \sum_{k'=1}^{k-1} n^{k'-k} \varepsilon_n^{(k',k)}(t) + \varepsilon_n^{(k,1)}(t) \right), \end{aligned} \quad (80)$$

where the last line comes from the fact that the  $\varepsilon_n^{(k',k)}$ 's are non-decreasing functions of  $t$ . Hence,  $B_t^n \lesssim_{(t,k)} \sum_{k'=1}^{k-1} n^{k'-k} \varepsilon_n^{(k',k)}(t) + \varepsilon_n^{(k,1)}(t)$ .

- Study of  $C^n(t)$ . Using the Lipschitz continuity of  $\Psi$ , Assumption  $(\mathcal{A}_\infty^h)$ , one has

$$C^n(t) \leq \text{Lip}(\Psi)^k h_\infty(t)^k t^{k-1} \int_0^t \mathbb{E} \left[ |\gamma_z^n - \bar{\gamma}(z)|^k \right] dz \lesssim_{(t,k)} \int_0^t \xi_n^{(k)}(z) dz. \quad (81)$$

- Study of  $D^n(t)$ . First remark that using Assumption  $(\mathcal{A}_\infty^h)$ , we have

$$D^n(t) \leq h_\infty(t)^k t^{k-1} \int_0^t \mathbb{E} \left[ \left| \frac{1}{n} \sum_{j=1}^n \bar{\lambda}_z^j - \bar{\lambda}(z) \right|^k \right] dz.$$

Yet, the  $\bar{\lambda}_z^j$ 's are i.i.d. with mean  $\bar{\lambda}(z)$  and they are bounded by  $\|\Psi\|_\infty$ . Hence, Rosenthal inequality [29] gives the existence of a constant  $C(k)$  which depends only on  $k$  and  $\|\Psi\|_\infty$  such that

$$\mathbb{E} \left[ \left| \frac{1}{n} \sum_{j=1}^n \bar{\lambda}_z^j - \bar{\lambda}(z) \right|^k \right] \leq C(k) n^{-k/2}.$$

It then follows that  $D^n(t) \lesssim_{(t,k)} n^{-k/2}$ .

One deduces from the decomposition (79) and the four bounds on  $A^n$ ,  $B^n$ ,  $C^n$  and  $D^n$  that

$$\xi_n^{(k)}(t) \lesssim_{(t,k)} \left( n^{-k/2} + \sum_{k'=1}^{k-1} n^{k'-k} \varepsilon_n^{(k',k)}(t) + \varepsilon^{(k,1)}(t) \right) + \int_0^t \xi_n^{(k)}(z) dz,$$

and so Lemma B.1 below gives the desired bound.

## A.2 Proof of Proposition IV.5

By definition of  $\eta^n$  (Equation (15)),

$$\langle \eta_t^n, \varphi \rangle - \langle \eta_0^n, \varphi \rangle = \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n \left( \langle \delta_{S_t^{n,i}}, \varphi \rangle - \langle \delta_{S_0^{n,i}}, \varphi \rangle \right) - (\langle u_t, \varphi \rangle - \langle u_0, \varphi \rangle) \right].$$

Since, for all  $i = 1, \dots, n$ , the age process  $(S_t^{n,i})_{t \geq 0}$  is piece-wise continuous, increasing with rate 1 and jumps from  $S_{t-}^{n,i}$  to 0 when  $N_t^{n,i} - N_{t-}^{n,i} = 1$ , we have

$$\langle \delta_{S_t^{n,i}}, \varphi \rangle - \langle \delta_{S_0^{n,i}}, \varphi \rangle = \int_0^t \varphi'(S_z^{n,i}) dz + \int_0^t R\varphi(S_{z-}^{n,i}) N^{n,i}(dz)$$

and so

$$\langle \bar{\mu}_{S_t}^n, \varphi \rangle - \langle \bar{\mu}_{S_0}^n, \varphi \rangle = \int_0^t \langle \bar{\mu}_{S_z}^n, \varphi' \rangle dz + \frac{1}{n} \sum_{i=1}^n \int_0^t R\varphi(S_{z-}^{n,i}) N^{n,i}(dz). \quad (82)$$

Now, we have in the same way

$$\langle \delta_{\bar{S}_t^1}, \varphi \rangle - \langle \delta_{\bar{S}_0^1}, \varphi \rangle = \int_0^t \varphi'(\bar{S}_z^1) dz + \int_0^t R\varphi(\bar{S}_{z-}^1) \bar{N}^1(dz)$$

and, by definition of  $(u_t)_{t \geq 0}$ ,

$$\langle u_t, \varphi \rangle - \langle u_0, \varphi \rangle = \mathbb{E} \left[ \int_0^t \varphi'(\bar{S}_z^1) dz \right] + \mathbb{E} \left[ \int_0^t R\varphi(\bar{S}_{z-}^1) \bar{N}^1(dz) \right]. \quad (83)$$

Yet, since  $\varphi'$  is bounded, Fubini's theorem gives that  $\mathbb{E}[\int_0^t \varphi'(\bar{S}_z^1) dz] = \int_0^t \langle u_z, \varphi' \rangle dz$ . Moreover, remind that the intensity of  $\bar{N}^1$  is  $\bar{\lambda}_t^1 = \Psi(\bar{S}_{t-}^1, \bar{\gamma}(t))$ . Yet, since  $\varphi$  and  $\Psi$  are bounded,

$$\mathbb{E} \left[ \int_0^t |R\varphi(\bar{S}_{z-}^1)| \bar{\lambda}_z^1 dz \right] < +\infty,$$

and so  $\mathbb{E}[\int_0^t R\varphi(\overline{S}_{z-}^1)\overline{N}^1(dz)] = \mathbb{E}[\int_0^t R\varphi(\overline{S}_{z-}^1)\overline{\lambda}_z^1 dz]$  since  $(\overline{S}_{t-}^1)_{t \geq 0}$  is a predictable process (see [6, II. T8]). Using once again Fubini's theorem, we end up with

$$\mathbb{E} \left[ \int_0^t R\varphi \left( \overline{S}_{z-}^1 \right) \overline{N}^1(dz) \right] = \int_0^t \langle u_z, \Psi(\cdot, \overline{\gamma}(z)) R\varphi \rangle dz$$

and so (83) becomes

$$\langle u_t, \varphi \rangle - \langle u_0, \varphi \rangle = \int_0^t \langle u_z, \varphi' \rangle dz + \int_0^t \langle u_z, \Psi(\cdot, \overline{\gamma}(z)) R\varphi \rangle dz. \quad (84)$$

Gathering (82) and (84) gives

$$\begin{aligned} \langle \eta_t^n, \varphi \rangle - \langle \eta_0^n, \varphi \rangle &= \int_0^t \langle \eta_z^n, \varphi' \rangle dz \\ &+ \sqrt{n} \left( \frac{1}{n} \sum_{i=1}^n \int_0^t R\varphi \left( S_{z-}^{n,i} \right) N^{n,i}(dz) - \int_0^t \langle u_z, \Psi(\cdot, \overline{\gamma}(z)) R\varphi \rangle dz \right) \end{aligned}$$

and so

$$\begin{aligned} \langle \eta_t^n, \varphi \rangle - \langle \eta_0^n, \varphi \rangle &= \int_0^t \langle \eta_z^n, L_z \varphi \rangle dz \\ &+ \sqrt{n} \left( \frac{1}{n} \sum_{i=1}^n \int_0^t R\varphi \left( S_{z-}^{n,i} \right) N^{n,i}(dz) - \int_0^t \langle \overline{\mu}_{S_{z-}^n}, \Psi(\cdot, \overline{\gamma}(z)) R\varphi \rangle dz \right), \quad (85) \end{aligned}$$

where we used that, almost surely,  $\overline{\mu}_{S_{z-}^n} = \overline{\mu}_{S_z^n}$  for almost every  $z$  in  $\mathbb{R}_+$ . Then, the second term in the right-hand side of (85) rewrites as  $M_t^n(\varphi) + \int_0^t A_z^n(\varphi) dz$ .

It remains to show that  $(M_t^n(\varphi))_{t \geq 0}$  is an  $\mathbb{F}$ -martingale. Yet, for all  $i = 1, \dots, n$ ,

$$\mathbb{E} \left[ \int_0^t \left| R\varphi \left( S_{z-}^{n,i} \right) \lambda_z^{n,i} dz \right| \right] \leq 2\|\varphi\|_\infty \mathbb{E} \left[ \int_0^t \lambda_z^{n,i} dz \right] = 2\|\varphi\|_\infty \mathbb{E} \left[ N_t^{n,i} \right] < +\infty,$$

and the  $\mathbb{F}$ -predictability of the age processes  $(S_{t-}^{n,i})_{t \geq 0}$  gives the result (see [6, II. T8]). Finally, the expression of the angle bracket (37) follows from standard computations for point processes (see [18, Proposition II.4.1.]).

### A.3 Proof of Proposition IV.7

**Proof of (i).** Let  $(\varphi_k)_{k \geq 1}$  be an orthonormal basis of  $\mathcal{W}_0^{1,\alpha}$  so that, in particular  $\|\eta_t^n\|_{-1,\alpha}^2 = \sum_{k \geq 1} \langle \eta_t^n, \varphi_k \rangle^2$ . Using the coupling (12)-(13), we have for every  $k$  and  $t \leq \theta$ ,

$$\langle \eta_t^n, \varphi_k \rangle = \sqrt{n} \left( \frac{1}{n} \sum_{i=1}^n \varphi_k(S_t^{n,i}) - \mathbb{E} \left[ \varphi_k(\overline{S}_t^i) \right] \right) = S_t^n(\varphi_k) + T_t^n(\varphi_k),$$

where

$$\begin{cases} S_t^n(\varphi_k) := n^{-1/2} \sum_{i=1}^n \varphi_k(S_t^{n,i}) - \varphi_k(\overline{S}_t^i) \\ T_t^n(\varphi_k) := n^{-1/2} \sum_{i=1}^n \varphi_k(\overline{S}_t^i) - \mathbb{E}[\varphi_k(\overline{S}_t^i)]. \end{cases}$$



On the one hand, using the independence of the age processes  $(\bar{S}_t^i)_{t \geq 0}$ , we have

$$\begin{aligned} \mathbb{E} \left[ \sum_{k \geq 1} T_t^n(\varphi_k)^2 \right] &= \sum_{k \geq 1} \mathbb{E} \left[ \frac{1}{n} \left( \sum_{i=1}^n \varphi_k(\bar{S}_t^i) - \mathbb{E} [\varphi_k(\bar{S}_t^i)] \right)^2 \right] \\ &= \sum_{k \geq 1} \mathbb{E} \left[ \left( \varphi_k(\bar{S}_t^1) - \mathbb{E} [\varphi_k(\bar{S}_t^1)] \right)^2 \right] \leq \sum_{k \geq 1} \mathbb{E} \left[ \left( \varphi_k(\bar{S}_t^1) \right)^2 \right] \\ &\leq \mathbb{E} \left[ \sum_{k \geq 1} (\delta_{\bar{S}_t^1}(\varphi_k))^2 \right] = \mathbb{E} \left[ \|\delta_{\bar{S}_t^1}\|_{-1, \alpha}^2 \right]. \end{aligned}$$

Then, using Lemma IV.3 and the fact that the age  $\bar{S}_t^1$  is upper bounded by  $M_{S_0} + t \leq M_{S_0} + \theta$  (thanks to  $(\mathcal{A}_\infty^{u_0})$ , remind (8)), it follows that

$$\mathbb{E} \left[ \sum_{k \geq 1} T_t^n(\varphi_k)^2 \right] \leq (C_1)^2 (1 + (M_{S_0} + \theta)^\alpha)^2$$

and so  $\sup_{n \geq 1} \sup_{t \in [0, \theta]} \mathbb{E} [\sum_{k \geq 1} T_t^n(\varphi_k)^2] < +\infty$ .

On the other hand, expanding the square and using exchangeability of the age processes  $(S_t^{n,i})_{t \geq 0}$ , one has

$$\begin{aligned} \mathbb{E} \left[ \sum_{k \geq 1} S_t^n(\varphi_k)^2 \right] &= n^{-1} \mathbb{E} \left[ \sum_{k \geq 1} \left( \sum_{i=1}^n \varphi_k(S_t^{n,i}) - \varphi_k(\bar{S}_t^i) \right)^2 \right] \\ &= (n-1) \mathbb{E} \left[ \sum_{k \geq 1} (\varphi_k(S_t^{n,1}) - \varphi_k(\bar{S}_t^1)) (\varphi_k(S_t^{n,2}) - \varphi_k(\bar{S}_t^2)) \right] \\ &\quad + \mathbb{E} \left[ \sum_{k \geq 1} (\varphi_k(S_t^{n,1}) - \varphi_k(\bar{S}_t^1))^2 \right]. \end{aligned} \quad (86)$$

Since the ages  $S_t^{n,1}$ ,  $\bar{S}_t^1$ ,  $S_t^{n,2}$  and  $\bar{S}_t^2$  are upper bounded by  $M_{S_0} + \theta$  and  $(\varphi_k(x_1) - \varphi_k(x_2))(\varphi_k(y_1) - \varphi_k(y_2)) = 0$  as soon as  $x_1 = x_2$  or  $y_1 = y_2$ , we have

$$\begin{aligned} \mathbb{E} \left[ \sum_{k \geq 1} (\varphi_k(S_t^{n,1}) - \varphi_k(\bar{S}_t^1)) (\varphi_k(S_t^{n,2}) - \varphi_k(\bar{S}_t^2)) \right] \\ \leq \chi_n^{(2)}(\theta) \sup_{x, y \leq M_{S_0} + \theta} \sum_{k \geq 1} |\varphi_k(x) - \varphi_k(y)|^2, \end{aligned} \quad (87)$$

where  $\chi_n^{(2)}(\theta)$  is defined by (16). Yet, since  $(\varphi_k)_{k \geq 1}$  is an orthonormal basis of  $\mathcal{W}_0^{1, \alpha}$ , we have

$$\begin{cases} \sum_{k \geq 1} (\varphi_k(S_t^{n,1}) - \varphi_k(\bar{S}_t^1))^2 = \sum_{k \geq 1} \langle D_{S_t^{n,1}, \bar{S}_t^1}, \varphi_k \rangle^2 = \|D_{S_t^{n,1}, \bar{S}_t^1}\|_{-1, \alpha}^2 \\ \sup_{x, y \leq M_{S_0} + \theta} \sum_{k \geq 1} |\varphi_k(x) - \varphi_k(y)|^2 = \sup_{x, y \leq M_{S_0} + \theta} \|D_{x, y}\|_{-1, \alpha}^2. \end{cases} \quad (88)$$

Hence, using Lemma IV.3 and once again the fact that the ages  $S_t^{n,1}$  and  $\bar{S}_t^1$  are upper bounded by  $M_{S_0} + \theta$ , we have, by gathering (86)-(88),

$$\mathbb{E} \left[ \sum_{k \geq 1} S_t^n(\varphi_k)^2 \right] \leq (n-1) \chi_n^{(2)}(\theta) (C_2)^2 (1 + (M_{S_0} + \theta)^\alpha)^2 + (C_2)^2 (1 + (M_{S_0} + \theta)^\alpha)^2,$$

and it follows from Proposition III.1 that  $\sup_{n \geq 1} \sup_{t \in [0, \theta]} \mathbb{E}[\sum_{k \geq 1} S_t^n(\varphi_k)^2] < +\infty$ .

Finally, by convexity of the square function,  $\|\eta_t^n\|_{-1, \alpha}^2 \leq 2 \sum_{k \geq 1} S_t^n(\varphi_k)^2 + T_t^n(\varphi_k)^2$  so that (38) follows from the two steps above.

**Proof of (ii).** We first show (39) and then use it in order to prove that  $(M_t^n)_{t \geq 0}$  is càdlàg. Let  $(\varphi_k)_{k \geq 1}$  be an orthonormal basis of  $\mathcal{W}_0^{1, \alpha}$  composed of  $\mathcal{C}^\infty$  functions with compact support. For all  $k \geq 1$ , the test function  $\varphi_k$  belongs to  $\mathcal{C}_b^1$  so that  $(M_t^n(\varphi_k))_{t \geq 0}$  is an  $\mathbb{F}$ -martingale (Proposition IV.5). Using Doob's inequality for real-valued martingales [21, Theorem 1.43.] and Equation (37), one has

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \in [0, \theta]} \|M_t^n\|_{-1, \alpha}^2 \right] &\leq \sum_{k \geq 1} \mathbb{E} \left[ \sup_{t \in [0, \theta]} M_t^n(\varphi_k)^2 \right] \\ &\leq C \sum_{k \geq 1} \mathbb{E} [M_\theta^n(\varphi_k)^2] \leq C \|\Psi\|_\infty \mathbb{E} \left[ \int_0^\theta \sum_{k \geq 1} R\varphi_k(S_{z-}^{n,1})^2 dz \right], \end{aligned}$$

where the last inequality comes from exchangeability and boundedness of the intensity. Noticing that  $R\varphi_k(S_{z-}^{n,1}) = D_{0, S_{z-}^{n,1}}(\varphi_k)$  and then using Lemma IV.3 as we have done in the proof of (i), it follows that

$$\mathbb{E} \left[ \int_0^\theta \sum_{k \geq 1} R\varphi_k(S_{z-}^{n,1})^2 dz \right] \leq (C_2)^2 \int_0^\theta (1 + (M_{S_0} + \theta)^\alpha)^2 dz,$$

which does not depend on  $n$  and gives (39). Moreover, gathering the integrability property given by (39) and the fact that, for all  $k \geq 1$ , the process  $(M_t^n(\varphi_k))_{t \geq 0}$  is an  $\mathbb{F}$ -martingale, we have that  $M^n$  is a  $\mathcal{W}_0^{-1, \alpha}$ -valued  $\mathbb{F}$ -martingale.

It remains to show that  $(M_t^n)_{t \geq 0}$  is càdlàg. First remark that for any  $k$ , the  $\mathbb{F}$ -martingale  $(M_t^n(\varphi_k))_{t \geq 0}$  is càdlàg. Let  $\varepsilon > 0$  and  $t_0 > 0$ . For any  $n \geq 1$ ,

$$\mathbb{E} \left[ \sum_{k \geq 1} \sup_{t \in [0, t_0+1]} M_t^n(\varphi_k)^2 \right] < +\infty,$$

so there exists a set  $\Omega^n$  such that  $\mathbb{P}(\Omega^n) = 1$  and for all  $\omega$  in  $\Omega^n$ ,

$$\sum_{k \geq 1} \sup_{t \in [0, t_0+1]} \langle M_t^n(\omega), \varphi_k \rangle^2 < +\infty.$$

Once  $\omega$  is fixed in  $\Omega^n$ , there exists an integer  $k_0$  (which depends on  $\omega$ ) such that  $\sum_{k > k_0} \sup_{t \in [0, t_0+1]} \langle M_t^n(\omega), \varphi_k \rangle^2 < \varepsilon$ . Let  $t$  be such that  $t_0 < t \leq t_0 + 1$ , using the right continuity of  $t \mapsto \langle M_t^n(\omega), \varphi_k \rangle$ , we have, dropping  $\omega$  for simplicity of notations,

$$\begin{aligned} \|M_t^n - M_{t_0}^n\|_{-1, \alpha}^2 &= \sum_{k \geq 1} (M_t^n(\varphi_k) - M_{t_0}^n(\varphi_k))^2 \\ &\leq \sum_{k=1}^{k_0} (M_t^n(\varphi_k) - M_{t_0}^n(\varphi_k))^2 + 2 \sum_{k > k_0} [M_t^n(\varphi_k)^2 + M_{t_0}^n(\varphi_k)^2] \\ &\leq \sum_{k=1}^{k_0} \varepsilon + 4\varepsilon = (k_0 + 4)\varepsilon, \end{aligned}$$

as soon as  $|t - t_0|$  is small enough. Hence,  $t \mapsto M_t^n(\omega)$  is right continuous with values in  $\mathcal{W}_0^{-1,\alpha}$ . In the same way, let  $(t_m)_{m \geq 1}$  be a sequence such that  $t_m < t_0$  and  $t_m \rightarrow t_0$ . For any integers  $m$  and  $\ell$ , we have, dropping  $\omega$  for simplicity of notations,

$$\begin{aligned} \|M_{t_m}^n - M_{t_\ell}^n\|_{-1,\alpha}^2 &= \sum_{k \geq 1} (M_{t_m}^n(\varphi_k) - M_{t_\ell}^n(\varphi_k))^2 \\ &\leq \sum_{k=1}^{k_0} (M_{t_m}^n(\varphi_k) - M_{t_\ell}^n(\varphi_k))^2 + 4\varepsilon. \end{aligned}$$

Yet, for all  $k = 1, \dots, k_0$ , the sequence  $(M_{t_m}^n(\varphi_k))_{m \geq 1}$  is convergent hence Cauchy. It follows that  $(M_{t_m}^n(\omega))_{m \geq 1}$  is a Cauchy sequence and so converges in  $\mathcal{W}_0^{-1,\alpha}$ . Hence,  $t \mapsto M_t^n(\omega)$  admits left limits in  $\mathcal{W}_0^{-1,\alpha}$ . Finally,  $t \mapsto M_t^n$  belongs to  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-1,\alpha})$  almost surely.

**Proof of (iii).** Starting from (43), we have, by convexity of the square function,

$$A_t^n(\varphi)^2 \leq 2 \left( \left\langle \bar{\mu}_{S_t}^n, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t)) R\varphi \right\rangle^2 (\Gamma_{t-}^n)^2 + R_t^{n,(1)}(\varphi)^2 \right)$$

Let  $(\varphi_k)_{k \geq 1}$  be an orthonormal basis of  $\mathcal{W}_0^{2,\alpha}$  so that  $\|A_t^n\|_{-2,\alpha}^2 = \sum_{k \geq 1} A_t^n(\varphi_k)^2$ . Noticing that  $R\varphi_k(S_{t-}^{n,i}) = D_{0,S_{t-}^{n,i}}(\varphi_k)$  and then using Lemma IV.3 as we have done in the proof of (i), it follows that

$$\begin{aligned} \sum_{k \geq 1} \left\langle \bar{\mu}_{S_t}^n, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t)) R\varphi_k \right\rangle^2 &\leq \text{Lip}(\Psi)^2 \frac{1}{n} \sum_{i=1}^n \left( \sum_{k \geq 1} R\varphi_k(S_{t-}^{n,i})^2 \right) \\ &\leq \text{Lip}(\Psi)^2 (C_2)^2 (1 + (M_{S_0} + \theta)^\alpha)^2, \end{aligned}$$

and in the same way,

$$\sum_{k \geq 1} R_t^{n,(1)}(\varphi_k)^2 \leq \text{Lip}(\Psi)^2 (C_2)^2 (1 + (M_{S_0} + \theta)^\alpha)^2 \frac{1}{n} \sum_{i=1}^n (\sqrt{nr_t}^{n,i})^2.$$

Hence,

$$\sum_{k \geq 1} A_t^n(\varphi_k)^2 \leq 2 \text{Lip}(\Psi)^2 (C_2)^2 (1 + (M_{S_0} + \theta)^\alpha)^2 \left( (\Gamma_{t-}^n)^2 + \frac{1}{n} \sum_{i=1}^n (\sqrt{nr_t}^{n,i})^2 \right).$$

Yet, as a consequence of Proposition III.1,  $\xi_n^{(2)}(t) = \mathbb{E}[|\Gamma_{t-}^n|^2]/n \lesssim_\theta n^{-1}$  and  $\xi_n^{(4)}(t) = \mathbb{E}[|\Gamma_{t-}^n|^4]/n^2 \lesssim_\theta n^{-2}$ . In particular, uniformly in  $t \leq \theta$ , the  $L^1$  norm of  $(\Gamma_{t-}^n)^2$  is of order 1 while the  $L^1$  norm of the rest term satisfies

$$\frac{1}{n} \sum_{i=1}^n (\sqrt{nr_t}^{n,i})^2 \leq n \left( \sup_{s,y} \left| \frac{\partial^2 \Psi}{\partial y^2}(s,y) \right| \right)^2 |\gamma_t^n - \bar{\gamma}(t)|^4 / 4$$

and so vanishes to 0 as  $n$  goes to infinity. Hence,

$$\sup_{n \geq 1} \sup_{t \in [0,\theta]} \mathbb{E} [\|A_t^n\|_{-2,\alpha}^2] = \sup_{n \geq 1} \sup_{t \in [0,\theta]} \mathbb{E} \left[ \sum_{k \geq 1} A_t^n(\varphi_k)^2 \right] < +\infty.$$

**Proof of (iv).** By definition of  $L_z$  and the triangular inequality,

$$\|L_z \varphi\|_{1,\alpha}^2 \leq 2(\|\varphi'\|_{1,\alpha}^2 + \|\Psi(\cdot, \bar{\gamma}(z))R\varphi\|_{1,\alpha}^2).$$

Firstly,  $\|\varphi'\|_{1,\alpha}^2 \leq \|\varphi\|_{2,\alpha}^2$ . Secondly, by Lemma IV.1, for all  $z \leq \theta$ ,

$$\|\Psi(\cdot, \bar{\gamma}(z))R\varphi\|_{1,\alpha}^2 \leq C \sup_{z \in [0, \theta]} \|\Psi(\cdot, \bar{\gamma}(z))\|_{C_b^2}^2 \|\varphi\|_{1,\alpha}^2.$$

Finally, (41) follows from (44) and the continuity of the mapping  $R$  (Lemma IV.2).

#### A.4 Proof of Proposition V.3

As stated in Equation (51), the sequence  $(W^n)_{n \geq 1}$  is tight. Then, let us consider the following decomposition, for any  $\varphi_1$  and  $\varphi_2$  in  $\mathcal{W}_0^{2,\alpha}$ ,

$$\langle \ll W^n \gg_t(\varphi_1), \varphi_2 \rangle - \int_0^t \langle u_z, \varphi_1 \varphi_2 \Psi(\cdot, \bar{\gamma}(z)) \rangle dz = B_t^n + C_t^n,$$

with

$$\begin{cases} B_t^n := \frac{1}{n} \sum_{i=1}^n \int_0^t \varphi_1(S_{z-}^{n,i}) \varphi_2(S_{z-}^{n,i}) \left( \lambda_z^{n,i} - \Psi(S_{z-}^{n,i}, \bar{\gamma}(z)) \right) dz, \\ C_t^n := \int_0^t \langle \bar{\mu}_{S_z}^n - u_z, \varphi_1 \varphi_2 \Psi(\cdot, \bar{\gamma}(z)) \rangle dz, \end{cases}$$

where we used the fact that, almost surely,  $\bar{\mu}_{S_{z-}}^n = \bar{\mu}_{S_z}^n$  for almost every  $z$  in  $\mathbb{R}_+$ . The first term  $B_t^n$  converges in  $L^1$  to 0 by using the Lipschitz continuity of  $\Psi$  and the convergence of  $\gamma^n$  to  $\bar{\gamma}$  given by Proposition III.1. From the convergence

$$\frac{1}{n} \sum_{i=1}^n \delta_{(S_t^{n,i})_{t \geq 0}} \xrightarrow[n \rightarrow \infty]{} \mathcal{L}((\bar{S}_t^1)_{t \geq 0}),$$

which is given in [9, Corollary IV.4], one can deduce that for almost every  $z$ ,

$$\frac{1}{n} \sum_{i=1}^n \delta_{S_z^{n,i}} \xrightarrow[n \rightarrow \infty]{} u_z$$

(see for instance [21, Proposition VI.3.14 and Lemma VI.3.12]). Then, dominated convergence implies that the second term  $C_t^n$  converges in expectation to 0. Hence, the bracket of  $W^n$  (50) converges to the covariance (52) for  $t' = t$ .

Furthermore, as for  $M^n$  (see the proof of Remark IV.12), the maximum jump size of  $W^n$  converges to 0. Hence, Rebolledo's central limit theorem for local martingales [34] gives, for every  $\varphi_1, \dots, \varphi_k$  in  $\mathcal{W}_0^{2,\alpha}$  and  $t_1, \dots, t_k \geq 0$ , the convergence of  $(W_{t_1}^n(\varphi_1), \dots, W_{t_k}^n(\varphi_k))$  to a Gaussian vector with the prescribed covariance (52). The limit law of  $(W^n)_{n \geq 1}$  is then characterized as the law of a continuous Gaussian process with covariance (52).

#### A.5 Proof of Corollary V.4

First, the tightness (and convergence) of  $(R^* W^n)_{n \geq 1}$  comes from the continuity of  $R^*$  as a mapping from  $\mathcal{W}_0^{-2,\alpha}$  to  $\mathcal{W}_0^{-2,\alpha}$  which comes from the continuity of  $R$  as a mapping from  $\mathcal{W}_0^{2,\alpha}$  to  $\mathcal{W}_0^{2,\alpha}$  (Lemma IV.2). Then, let us show that  $(V^n)_{n \geq 1}$  is tight in  $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$ .

Assume that  $h(0) = 0$  and extend the function  $h$  to the whole real line by the value 0 on the negative real numbers.

-(i) For all  $n \geq 1$ ,  $V_0^n = 0$  a.s. so  $(V_0^n)_{n \geq 1}$  is clearly tight.

For any  $t > r \geq 0$ , since  $h(r - z) = 0$  as soon as  $z \geq r$ , one has

$$V_t^n - V_r^n = \int_0^t [h(t - z) - h(r - z)] dW_z^n(\mathbf{1}).$$

Let us denote, for all  $x \geq 0$ ,  $V_{r,t}^n(x) = \int_0^x [h(t - z) - h(r - z)] dW_z^n(\mathbf{1})$ . It is a martingale with respect to  $x$ . Burkholder-Davis-Gundy inequality [38, p. 894] gives the existence of a universal constant  $C_p$  such that

$$\mathbb{E} \left[ \sup_{x \leq t} |V_{r,t}^n(x)|^{2p} \right] \leq C_p \mathbb{E} \left[ [V_{r,t}^n]_t^p \right].$$

Yet, the quadratic variation of  $V_{r,t}^n$  is given by

$$[V_{r,t}^n]_x = \frac{1}{n} \sum_{j=1}^n \int_0^x [h(t - z) - h(r - z)]^2 N^{n,j}(dz) \leq \text{Höl}(h)^2 |t - r|^{2\beta(h)} \frac{1}{n} \sum_{j=1}^n N_x^{n,j}. \quad (89)$$

So, using the exchangeability, we have for all  $p \geq 0$ ,

$$\mathbb{E} [|V_t^n - V_r^n|^{2p}] \leq C_p \text{Höl}(h)^{2p} |t - r|^{2\beta(h)p} \mathbb{E} [|N_t^{n,1}|^p]. \quad (90)$$

Yet, the intensity of  $N^{n,1}$  is bounded so that  $N^{n,1}$  is stochastically dominated by a Poisson process with intensity  $\|\Psi\|_\infty$ . Hence,  $\mathbb{E}[|N_t^{n,1}|^p] \leq \mathbb{E}[\text{Pois}(t|\|\Psi\|_\infty)^p]$  where  $\text{Pois}(t|\|\Psi\|_\infty)$  is a Poisson variable with parameter  $t|\|\Psi\|_\infty$ . This implies that  $\mathbb{E}[|N_t^{n,1}|^p]$  is bounded uniformly in  $n$  by a locally bounded function of the time  $t$ , say  $\tilde{C}_p(t)$  (which can be assumed to be increasing continuous without any loss of generality). Then, taking  $p = 1$ ,  $t = \delta$  and  $r = 0$  and using Markov's inequality gives

-(ii) for all  $\varepsilon > 0$ ,  $\lim_{\delta \rightarrow 0} \limsup_n \mathbb{P}(|V_\delta^n - V_0^n| > \varepsilon) = 0$ .

Finally, taking  $p = 1/\beta(h)$  and using Markov's inequality gives

-(iii) for all  $\nu > 0$ ,  $\mathbb{P}(|V_t^n - V_r^n| > \nu) \leq \nu^{-2/\beta(h)} |F(t) - F(s)|^2$ ,

where  $F(t) := (C_{\frac{1}{\beta(h)}} \tilde{C}_{\frac{1}{\beta(h)}}(t) \text{Höl}(h)^{2/\beta(h)})^{1/2} t$  defines an increasing continuous function.

Hence, (i), (ii) and (iii) allow to apply Billingsley's criterion for tightness [21, Theorem VI.4.1] to deduce that  $(V^n)_{n \geq 1}$  is tight.

Now, if  $h(0) \neq 0$ , one can use the following decomposition,

$$V_t^n = \int_0^t [h(t - z) - h(0)] dW_z^n(\mathbf{1}) + h(0) W_t^n(\mathbf{1}).$$

The first term is tight thanks to what we have done in the case  $h(0) = 0$  whereas the second term is converging since  $(W^n)_n$  is converging, whence  $(V^n)_{n \geq 1}$  is tight.

Now, since the limit trajectories of  $R^*W^n$  are continuous,  $(R^*W^n, V^n)_{n \geq 1}$  is tight in  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha} \times \mathbb{R})$ . It now suffices to characterize the limiting finite dimensional distributions. Recall that  $V_t = \int_0^t h(t - z) dW_z(\mathbf{1})$  and denote by  $(t_1, \dots, t_k)$  a  $k$ -tuple of positive times.

First, suppose that  $h$  is piecewise constant. In that case, the convergence of  $W^n$  towards  $W$  easily implies the convergence of  $((R^*W_{t_1}^n, V_{t_1}^n), \dots, (R^*W_{t_k}^n, V_{t_k}^n))$  to

$((R^*W_{t_1}, V_{t_1}), \dots, (R^*W_{t_k}, V_{t_k}))$  (use the fact that  $h$  is a piecewise function to write  $V_{t_i}^n$  as a sum of increments of  $W^n(\mathbf{1})$ ).

Then, since  $h$  is continuous, one can find, for each  $\varepsilon > 0$ , a piecewise constant function  $h^\varepsilon$  such that  $\|h - h^\varepsilon\|_\infty \leq \varepsilon$ . Denote  $V_t^{n,\varepsilon} := \int_0^t h^\varepsilon(t-z) dW_z^n(\mathbf{1})$  and notice that  $\mathbb{E}[|V_t^n - V_t^{n,\varepsilon}|^2] \leq 2\varepsilon^2 \mathbb{E}[\langle W^n(\mathbf{1}) \rangle_t] \leq 2\varepsilon^2 \|\Psi\|_\infty t \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . In the same way, denote  $V_t^\varepsilon := \int_0^t h^\varepsilon(t-z) dW_z(\mathbf{1})$  and remark that  $\mathbb{E}[|V_t - V_t^\varepsilon|^2] \leq 2\varepsilon^2 \|\Psi\|_\infty t \rightarrow 0$  as  $\varepsilon \rightarrow 0$ . Yet, the previous point gives the convergence, in terms of finite dimensional distributions, of  $V^{n,\varepsilon}$  to  $V^\varepsilon$  for all  $\varepsilon > 0$  so the convergence, in terms of finite dimensional distributions, of  $V^n$  to  $V$  follows which ends the proof.

## A.6 Proof of Proposition V.5

The idea is to use (55). The first step is to simplify (55) by using the following convergences

$$\begin{cases} \mathbb{E} \left[ \left| \sup_{t \in [0, \theta]} \int_0^t h(t-z) R_z^{n,(2)} dz \right| \right] \rightarrow 0, \\ \mathbb{E} \left[ \left| \sup_{t \in [0, \theta]} \int_0^t h(t-z) \left\langle \bar{\mu}_{S_z}^n - u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \Gamma_z^n dz \right| \right] \rightarrow 0. \end{cases} \quad (91)$$

These two convergences follow from the two following claims: by (56),

$$\sup_{z \in [0, \theta]} \mathbb{E} \left[ \left| R_z^{n,(2)} \right| \right] \leq \text{Lip}(\Psi) C n^{-1/2} \sup_{z \in [0, \theta]} \mathbb{E} [|\Gamma_z^n|^2] \rightarrow 0,$$

and, by Cauchy-Schwarz inequality,

$$\sup_{z \in [0, \theta]} \mathbb{E} \left[ \left| \left\langle \bar{\mu}_{S_z}^n - u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \Gamma_z^n \right| \right] \rightarrow 0. \quad (92)$$

Indeed,

$$\begin{aligned} \mathbb{E} \left[ \left| \left\langle \bar{\mu}_{S_z}^n - u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \Gamma_z^n \right| \right] &\leq \mathbb{E} \left[ \left| \frac{1}{\sqrt{n}} \left\langle \eta_z^n, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \right|^2 \right]^{1/2} \mathbb{E} [|\Gamma_z^n|^2]^{1/2} \\ &\leq \frac{1}{\sqrt{n}} \mathbb{E} [|\eta_z^n|_{-1, \alpha}^2]^{1/2} \left\| \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\|_{1, \alpha} \mathbb{E} [|\Gamma_z^n|^2]^{1/2}, \end{aligned}$$

for any  $\alpha > 1/2$  and (92) follows from Proposition IV.7-(i) and Equations (30), (44) and (56).

Return to (55). The right-hand side is tight since it is convergent (Corollary V.4) and the last term in the left hand side is tight since  $(\eta^n)_{n \geq 1}$  is tight (with continuous limit points) and  $\eta \mapsto \int_0^t h(t-z) \langle \eta_z, \Psi(\cdot, \bar{\gamma}(z)) \rangle dz$  is continuous at every point  $\eta$  in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha})$  thanks to Lemma B.3 (remind (30) and (44)). Moreover, the term in the middle may be simplified by means of (91). Hence it remains to prove the tightness of the sequence of continuous processes  $(I^n)_{n \geq 1}$  defined, for all  $t \geq 0$ , by

$$I_t^n := \int_0^t h(t-z) \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \Gamma_z^n dz.$$

We use Aldous criterion [5, Theorem 16.10.], that is the simplified version of the one stated on page 13 but for real valued processes. First, for all  $\theta \geq 0$ ,

$$\mathbb{E} \left[ \sup_{t \in [0, \theta]} |I_t^n| \right] \leq h_\infty(\theta) \text{Lip}(\Psi) \int_0^\theta \mathbb{E} [|\Gamma_z^n|] dz,$$

is bounded uniformly with respect to  $n$  thanks to Equation (56). And Markov's inequality implies that, for every  $\theta \geq 0$  and  $\varepsilon > 0$ , there exists  $a > 0$  such that

$$\sup_{n \geq 1} \mathbb{P} \left( \sup_{t \in [0, \theta]} |I_t^n| \geq a \right) \leq \varepsilon,$$

which is the standard compactness condition.

Then, for the Aldous criterion, let us consider  $\delta_0 > 0$ ,  $\delta \leq \delta_0$  and for all  $n \geq 1$ , an  $\mathbb{F}$ -stopping time smaller than  $\theta$  denoted by  $\tau_n$ . Assume for a while that  $h(0) = 0$  and extend the function  $h$  to the whole real line by setting 0 on the negative real numbers. As for Equation (89), we have

$$|I_{\tau_n + \delta}^n - I_{\tau_n}^n| \leq \text{Höl}(h) \delta_0^{\beta(h)} \text{Lip}(\Psi) \int_0^{\theta + \delta_0} |\Gamma_z^n| dz.$$

Hence, as before, (56) implies that  $\sup_{n \geq 1} \mathbb{E}[|I_{\tau_n + \delta}^n - I_{\tau_n}^n|] \leq C(\theta + \delta_0) \delta_0^{\beta(h)}$  which is arbitrary small for  $\delta_0$  small enough and Markov's inequality gives that, for any  $\varepsilon_1, \varepsilon_2 > 0$ , there exists  $\delta_0$  such that  $\sup_{n \geq 1} \sup_{\delta \leq \delta_0} \mathbb{P}(|I_{\tau_n + \delta}^n - I_{\tau_n}^n| \geq \varepsilon_1) \leq \varepsilon_2$ , that is Aldous criterion. Hence,  $(I^n)_{n \geq 1}$  is tight in  $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$ .

Now, if  $h(0) \neq 0$ , one can use the following decomposition,

$$I^n(t) = \int_0^t (h(t-z) - h(0)) \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \Gamma_z^n dz + h(0) \int_0^t \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \Gamma_z^n dz.$$

The first term is tight thanks to what we have done in the case  $h(0) = 0$  whereas the tightness of the second one is simpler and left to the reader (use Equation (56)).

It only remains to check that the limit points are continuous. The idea is the same as for Remark IV.12. According to [5, Theorem 13.4.], it suffices to prove that for all  $\theta \geq 0$ , the maximal jump size of  $\Gamma^n$  on  $[0, \theta]$  converges to 0. Yet, using the continuity of  $\bar{\gamma}$  and the reverse triangle inequality, we have

$$\Delta \Gamma_t^n := |\Gamma_t^n - \Gamma_{t-}^n| \leq \sqrt{n} |\gamma_{t+}^n - \gamma_t^n|,$$

where we remind the definition  $\gamma_t^n = n^{-1} \sum_{j=1}^n \int_0^{t-} h(t-z) N^{n,j}(dz)$  and we define  $\gamma_{t+}^n := n^{-1} \sum_{j=1}^n \int_0^t h(t-z) N^{n,j}(dz)$ . Now, Assumption ( $\mathcal{A}_{\text{Höl}}^h$ ) implies that  $h$  is continuous and so we deduce that

$$\Delta \Gamma_t^n \leq \sqrt{n} \frac{1}{n} \sum_{i=1}^n h(0) \mathbf{1}_{t \in N^{n,i}}.$$

Since almost surely there is no common point to any two of the point processes  $(N^{n,i})_{i=1, \dots, n}$ , there is, almost surely, for all  $t \geq 0$ , at most one of the  $\mathbf{1}_{t \in N^{n,i}}$  which is non null. Hence,  $\sup_{t \in [0, \theta]} \Delta \Gamma_t^n \leq h(0) n^{-1/2}$  a.s., which gives the desired convergence to 0.

## A.7 Proof of Theorem V.6

First, let us notice that we use the two following statements whose proofs are similar to those of Proposition IV.7-(iv) and Remark IV.8: for any  $\alpha > 1/2$  and  $\theta \geq 0$ ,

$$\text{for all } \varphi \text{ in } \mathcal{W}_0^{3, \alpha}, \quad \sup_{z \in [0, \theta]} \frac{\|L_z \varphi\|_{2, \alpha}^2}{\|\varphi\|_{3, \alpha}^2} < +\infty, \quad (93)$$



$$\text{and, for all } w \text{ in } \mathcal{W}_0^{-2,\alpha}, \quad \sup_{z \in [0, \theta]} \frac{\|L_z^* w\|_{-3,\alpha}^2}{\|w\|_{-2,\alpha}^2} < +\infty. \quad (94)$$

As a consequence of tightness and continuity of the limit trajectories, we have tightness of the process  $(\eta^n, \Gamma^n, W^n, V^n)_{n \geq 1}$  in  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}^{-2,\alpha} \times \mathbb{R} \times \mathcal{W}_0^{-2,\alpha} \times \mathbb{R})$ . Hence, let us assume without loss of generality that the sequence converges to  $(\eta, \Gamma, W, V)$  in  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}^{-2,\alpha} \times \mathbb{R} \times \mathcal{W}_0^{-2,\alpha} \times \mathbb{R})$ .

Then, let  $(\varphi_k)_{k \geq 1}$  be an orthonormal basis of  $\mathcal{W}_0^{3,\alpha}$  and define the following applications: for all  $k \geq 1$ ,  $F_k : \mathcal{D}(\mathbb{R}_+, \mathcal{W}^{-2,\alpha} \times \mathbb{R} \times \mathcal{W}_0^{-2,\alpha}) \rightarrow \mathcal{D}(\mathbb{R}_+, \mathbb{R})$  satisfy for all  $t \geq 0$ ,

$$\begin{aligned} F_k(f^1, f^2, f^3)(t) &:= \langle f_t^1, \varphi_k \rangle - \langle f_0^1, \varphi_k \rangle - \int_0^t \langle f_z^1, L_z \varphi_k \rangle dz \\ &\quad - \int_0^t \left\langle u_z, R \varphi_k \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle f_z^2 dz - f_t^3(R \varphi_k), \end{aligned}$$

and  $G : \mathcal{D}(\mathbb{R}_+, \mathcal{W}^{-2,\alpha} \times \mathbb{R} \times \mathbb{R}) \rightarrow \mathcal{D}(\mathbb{R}_+, \mathbb{R})$  satisfy for all  $t \geq 0$ ,

$$\begin{aligned} G(g^1, g^2, g^3)(t) &:= g_t^2 - \int_0^t h(t-z) \langle g_z^1, \Psi(\cdot, \bar{\gamma}(z)) \rangle dz \\ &\quad - \int_0^t h(t-z) \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle g_z^2 dz - g_t^3. \end{aligned}$$

Notice that the system (58)-(59) is equivalent to

$$\begin{cases} \forall k \geq 1, & F_k(\eta, \Gamma, W) = 0 \\ G(\eta, \Gamma, V) = 0. \end{cases} \quad (95)$$

**Step one.** Let us show that the first line of (95) is satisfied. First, we prove that for all  $k \geq 1$ ,  $F_k$  is continuous at every point  $(\tilde{f}^1, \tilde{f}^2, \tilde{f}^3)$  in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}^{-2,\alpha} \times \mathbb{R} \times \mathcal{W}_0^{-2,\alpha})$ . To state continuity of  $F_k$  at a continuous trajectory, it suffices to show continuity with respect to each coordinate  $f^1$ ,  $f^2$  and  $f^3$ .

- Equation (93) implies that  $z \mapsto \|L_z \varphi_k\|_{2,\alpha}$  is locally bounded and Lemma B.3 gives the following:

$$f^1 \mapsto \left( t \mapsto \langle f^1(t), \varphi_k \rangle - \langle f^1(0), \varphi_k \rangle - \int_0^t \langle f^1(z), L_z \varphi_k \rangle dz \right)$$

is a mapping from  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha})$  into  $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$  which is continuous at every point  $\tilde{f}^1$  in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha})$ .

- Then, notice that  $|\langle u_z, R \varphi_k \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \rangle| \leq \text{Lip}(\Psi) \mathbb{E}[|D_{0, \bar{S}_z^1}(\varphi_k)|]$  and Lemma IV.3 gives that

$$z \mapsto \left\langle u_z, R \varphi_k \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \text{ is locally bounded,}$$

so that applying Lemma B.4 gives the continuity of  $F_k$  with respect to  $f^2$ .

- Finally,  $F_k$  is clearly continuous with respect to  $f^3$ .

Notice that (54) gives for any  $k \geq 1$ ,

$$\forall t \geq 0, F_k(\eta^n, \Gamma^n, W^n)(t) - \int_0^t R_z^{n,(1)}(\varphi_k) dz - \int_0^t \left\langle \bar{\mu}_{S_z}^n - u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) R \varphi_k \right\rangle \Gamma_z^n dz = 0. \quad (96)$$

Yet, we have, on the one hand, for all  $\theta \geq 0$ ,

$$\begin{aligned} \sup_{z \in [0, \theta]} \mathbb{E} \left[ |R_z^{n,(1)}(\varphi_k)| \right] &\leq \sqrt{n} \text{Lip}(\Psi) \sup_{z \in [0, \theta]} \mathbb{E} \left[ \frac{1}{n} \sum_{i=1}^n |D_{0, S_z^{n,i}}(\varphi_k)| |r_z^{n,i}| \right] \\ &\leq n^{-1/2} C \text{Lip}(\Psi) (1 + (M_{T_0} + \theta)^\alpha) \|\varphi_k\|_{2, \alpha} \sup_{z \in [0, \theta]} \mathbb{E} [|\Gamma_z^n|^2] \rightarrow 0, \end{aligned}$$

where we used Lemma IV.3, Equations (53) and (56), and on the other hand,

$$\sup_{z \in [0, \theta]} \mathbb{E} \left[ \left| \left\langle \bar{\mu}_{S_z}^n - u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) R \varphi_k \right\rangle \Gamma_z^n \right| \right] \rightarrow 0,$$

which follows from Cauchy-Schwarz inequality as we have done for (92).

These two convergences above imply

$$\begin{cases} \mathbb{E} \left[ \left| \sup_{t \in [0, \theta]} \int_0^t R_z^{n,(1)}(\varphi_k) dz \right| \right] \rightarrow 0, \\ \mathbb{E} \left[ \left| \sup_{t \in [0, \theta]} \int_0^t \left\langle \bar{\mu}_{S_z}^n - u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) R \varphi_k \right\rangle \Gamma_z^n dz \right| \right] \rightarrow 0. \end{cases} \quad (97)$$

On the one hand, gathering (96) and (97) gives the convergence of  $F_k(\eta^n, \Gamma^n, W^n)$  to 0 in probability and, on the other hand, applying the continuous mapping theorem gives the convergence in law of  $F_k(\eta^n, \Gamma^n, W^n)$  to  $F_k(\eta, \Gamma, W)$ . Identifying the limits gives  $F_k(\eta, \Gamma, W) = 0$  which ends this step.

**Step two.** Let us show that the second line of (95) is satisfied. First, we prove that  $G$  is continuous at every point  $(\tilde{g}^1, \tilde{g}^2, \tilde{g}^3)$  in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}^{-2, \alpha} \times \mathbb{R} \times \mathbb{R})$ . To state continuity of  $G$  at a continuous trajectory, it suffices to show continuity with respect to each coordinate  $g^1$ ,  $g^2$  and  $g^3$ .

- Equations (30), (44) and Lemma B.3 give the following:

$$g^1 \mapsto \left( t \mapsto \int_0^t h(t-z) \langle g_z^1, \Psi(\cdot, \bar{\gamma}(z)) \rangle dz \right)$$

is a mapping from  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha})$  into  $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$  which is continuous at every point  $\tilde{g}^1$  in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha})$ .

- Then, notice that  $|\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \rangle| \leq \text{Lip}(\Psi)$  so that

$$z \mapsto \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \text{ is locally bounded,}$$

so that applying Lemma B.4 gives the continuity of  $G$  with respect to  $g^2$ .

- Finally,  $G$  is clearly continuous with respect to  $g^3$ .

Notice that (55) gives

$$\begin{aligned} \forall t \geq 0, \quad G(\eta^n, \Gamma^n, V^n)(t) - \int_0^t h(t-z) R_z^{n,(2)} dz \\ - \int_0^t h(t-z) \left\langle \bar{\mu}_{S_z}^n - u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \Gamma_z^n dz = 0. \end{aligned}$$

Finally, the argument used to end the previous step also applies here.

To conclude, the two steps above give (95) which gives that the process  $(\eta, \Gamma)$  is a solution of (58)-(59). Finally, its trajectories are supported in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha} \times \mathbb{R})$  since  $\eta$  is supported in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2,\alpha})$  and  $\Gamma$  is supported in  $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$  as a solution of (59) (remind that  $h$  is Hölder continuous).

## A.8 Proof of Proposition VI.4

First, the following Lemma states the well-posedness of  $\hat{\gamma}^n$  as defined by (74).

**Lemma A.1.** *Under  $(\mathcal{A}_{CLT})$ , assume furthermore that  $\Psi$  is in  $\mathcal{C}_b^4$ . For all  $T \geq 0$ , there exists a pathwise unique of solution  $(\hat{\gamma}_t^n)_{t \in [0, T]}$  of Equation (74). Furthermore, the solution  $\hat{\gamma}^n$  has continuous paths.*

*Proof.* We first deal with the continuity of  $U_t^n := \int_0^t h(t-z) d\hat{W}_z^n(\mathbf{1})$  appearing in (74). Following the arguments given in the proof of Corollary V.4 to get the control (90), one can prove that, for all  $r < t < T$  and  $p \geq 0$ , there exists a universal constant  $C_p$  such that

$$\mathbb{E} [|U_t^n - U_r^n|^{2p}] \leq C_p \text{Höl}(h)^{2p} |t-r|^{2\beta(h)p} \int_0^t \frac{\langle u_z, \Psi(\cdot, \bar{\gamma}_z) \rangle}{n} dz.$$

The integral above being clearly bounded by  $T \|\Psi\|_\infty$ , one can for instance take  $p = 1/\beta(h)$  in the equation above in order to apply Kolmogorov continuity theorem.

Hence, without loss of generality, one can deduce that any solution of (74) admits a modification with continuous paths. Let  $t_0 < T$  and  $g$  be in  $\mathcal{C}([0, T], \mathbb{R})$  and consider the application  $F_{t_0, g} : \mathcal{C}([0, t_0], \mathbb{R}) \rightarrow \mathcal{C}([0, t_0], \mathbb{R})$  defined by

$$F_{t_0, g}(\gamma)(t) := \int_0^t h(t-z) \langle \hat{u}_z^n, \Psi(\cdot, \gamma(z)) \rangle dz + g(t).$$

Remind that  $\hat{u}_t^n = u_t + n^{-1/2} \eta_t$  and remark that

$$\begin{aligned} |\langle \eta_t, \Psi(\cdot, \gamma_1(t)) - \Psi(\cdot, \gamma_2(t)) \rangle| &\leq C \|\eta_t\|_{-2,1} \|\Psi(\cdot, \gamma_1(t)) - \Psi(\cdot, \gamma_2(t))\|_{\mathcal{C}_b^2} \\ &\leq C \|\eta_t\|_{-2,1} |\gamma_1(t) - \gamma_2(t)|, \end{aligned}$$

since  $\Psi$  is  $\mathcal{C}_b^4$ , where the  $C$ 's are deterministic constants. Hence, we have

$$\begin{aligned} \|F_{t_0, g}(\gamma_1) - F_{t_0, g}(\gamma_2)\|_{\mathcal{C}([0, t_0])} &\leq t_0 h_\infty(t_0) \|\gamma_1 - \gamma_2\|_{\mathcal{C}([0, t_0])} (\text{Lip}(\Psi) \\ &\quad + C \sup_{t \in [0, T]} \|\eta_t\|_{-2,1}). \end{aligned}$$

Yet,  $\sup_{t \in [0, T]} \|\eta_t\|_{-2,1}$  is almost surely finite as a consequence of (49) so one can find  $t_0$  small enough<sup>4</sup> such that  $F_{t_0, g}$  is a contraction. The proof is then completed by iteration. Let us mention how goes the second step: fix  $t_0$  such that  $F_{t_0, g}$  is

<sup>4</sup>The time  $t_0$  is random and depends on the time horizon  $T$ .

a contraction and consider the application  $G_{t_0, g} : \mathcal{C}([t_0, 2t_0], \mathbb{R}) \rightarrow \mathcal{C}([t_0, 2t_0], \mathbb{R})$  defined by

$$G_{t_0, g}(\gamma)(t) := \int_0^t h(t-z) \langle \hat{u}_z^n, \Psi(\cdot, \gamma(z)) \rangle dz + g(t),$$

where the value of  $\gamma(t)$  for  $t$  in  $[0, t_0]$  is given by the unique Banach-Picard fixed point of  $F_{t_0, g}$ . The same kind of computations as before gives

$$\begin{aligned} \|G_{t_0, g}(\gamma_1) - G_{t_0, g}(\gamma_2)\|_{\mathcal{C}([t_0, 2t_0])} &\leq t_0 h_\infty(t_0) \|\gamma_1 - \gamma_2\|_{\mathcal{C}([t_0, 2t_0])} (\text{Lip}(\Psi) \\ &\quad + C \sup_{t \in [0, T]} \|\eta_t\|_{-2,1}), \end{aligned}$$

so that  $G_{t_0, g}$  is also a contraction.  $\square$

Then, the following Lemma relates  $\hat{\gamma}_t^n$  with its approximation appearing through the CLT, namely  $\check{\gamma}_t^n := \bar{\gamma}(t) + n^{-1/2}\Gamma_t$ .

**Lemma A.2.** *We have*

$$\mathbb{E} [|\hat{\gamma}_t^n - \check{\gamma}_t^n|] \lesssim_t n^{-1}. \quad (98)$$

Furthermore, we show within the proof that

$$\mathbb{E} [|\hat{\gamma}_t^n - \bar{\gamma}(t)|^2] \lesssim_t n^{-1}. \quad (99)$$

*Proof.* Let us prove the a priori rough bound,

$$\mathbb{E} \left[ |\hat{\gamma}_t^n - \check{\gamma}_t^n|^2 \right]^{1/2} \lesssim_t n^{-1/2}. \quad (100)$$

We use the decomposition  $\hat{\gamma}_t^n - \check{\gamma}_t^n = (\hat{\gamma}_t^n - \bar{\gamma}(t)) - n^{-1/2}\Gamma_t$ . On the one hand,  $\mathbb{E} [ |n^{-1/2}\Gamma_t|^2 ]^{1/2} \lesssim_t n^{-1/2}$  thanks to (57). On the other hand, we use the decomposition

$$\mathbb{E} \left[ |\hat{\gamma}_t^n - \bar{\gamma}(t)|^2 \right] \leq 3(A_1^n(t) + A_2^n(t) + A_3^n(t)),$$

where

$$\begin{cases} A_1^n(t) := \mathbb{E} \left[ \left| \int_0^t h(t-z) \langle \hat{u}_z^n - u_z, \Psi(\cdot, \hat{\gamma}_z^n) \rangle dz \right|^2 \right], \\ A_2^n(t) := \mathbb{E} \left[ \left| \int_0^t h(t-z) \langle u_z, \Psi(\cdot, \hat{\gamma}_z^n) - \Psi(\cdot, \bar{\gamma}(z)) \rangle dz \right|^2 \right], \\ A_3^n(t) := \mathbb{E} \left[ \left| \int_0^t h(t-z) d\hat{W}_z^n(\mathbf{1}) \right|^2 \right]. \end{cases}$$

- Study of  $A_1^n(t)$ . Using that  $\hat{u}_z^n - u_z = n^{-1/2}\eta_z$  and  $\Psi$  belongs to  $\mathcal{C}_b^2$ , we have

$$A_1^n(t) \leq t^2 h_\infty(t)^2 n^{-1} \mathbb{E} \left[ \sup_{z \in [0, t]} \|\eta_z\|_{-2,1}^2 \right] \|\Psi\|_{\mathcal{C}_b^2}^2 \lesssim_t n^{-1},$$

where we used (49).

- Study of  $A_2^n(t)$ . Using the Lipschitz continuity of  $\Psi$ , we have

$$A_2^n(t) \leq t h_\infty(t)^2 \text{Lip}(\Psi) \int_0^t \mathbb{E} \left[ |\hat{\gamma}_z^n - \bar{\gamma}(z)|^2 \right] dz \lesssim_t \int_0^t \mathbb{E} \left[ |\hat{\gamma}_z^n - \bar{\gamma}(z)|^2 \right] dz,$$

which is convenient to apply the Grönwall-type Lemma B.1.

- Study of  $A_3^n(t)$ . By definition of the bracket of  $\hat{W}^n$ , we have

$$A_3^n(t) = \int_0^t h(t-z)^2 n^{-1} \langle u_z, \Psi(\cdot, \bar{\gamma}(z)) \rangle dz \lesssim_t n^{-1}.$$

As expected, applying Lemma B.1 gives  $\mathbb{E}[|\hat{\gamma}_t^n - \bar{\gamma}(t)|^2] \lesssim_t n^{-1}$ , that is Equation (99). Then, gathering the two steps above proves (100).

Then, let us show how to use (100) in order to prove (98). We have,

$$\begin{aligned} \tilde{\gamma}_t^n = \int_0^t h(t-z) \left[ \langle \hat{u}_z^n, \Psi(\cdot, \bar{\gamma}(z)) \rangle + \left\langle u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle n^{-1/2} \Gamma_z \right] dz \\ + \int_0^t h(t-z) d\hat{W}_z^n(\mathbf{1}). \end{aligned}$$

Hence we use the decomposition

$$\mathbb{E}[|\hat{\gamma}_t^n - \tilde{\gamma}_t^n|] \leq 3(B_1^n(t) + B_2^n(t) + B_3^n(t)),$$

where

$$\begin{cases} B_1^n(t) := \mathbb{E} \left[ \left| \int_0^t h(t-z) \langle \hat{u}_z^n, \Psi(\cdot, \hat{\gamma}_z^n) - \Psi(\cdot, \tilde{\gamma}_z^n) \rangle dz \right| \right], \\ B_2^n(t) := \mathbb{E} \left[ \left| \int_0^t h(t-z) \left\langle \hat{u}_z^n, \Psi(\cdot, \tilde{\gamma}_z^n) - \Psi(\cdot, \bar{\gamma}(z)) - \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \frac{\Gamma_z}{\sqrt{n}} \right\rangle dz \right| \right], \\ B_3^n(t) := \mathbb{E} \left[ \left| \int_0^t h(t-z) \left\langle \hat{u}_z^n - u_z, \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\rangle \frac{\Gamma_z}{\sqrt{n}} dz \right| \right]. \end{cases}$$

- Study of  $B_1^n(t)$ . Using Lemma VI.5, Cauchy-Schwarz inequality and finally (49) and the a priori rough bound (100), we have

$$\begin{aligned} B_1^n(t) &\leq h_\infty(t) \int_0^t \mathbb{E} \left[ \left( 1 + \frac{C}{\sqrt{n}} \|\eta_t\|_{-2,1} \right) \|\Psi\|_{\mathcal{C}_b^3} |\hat{\gamma}_z^n - \tilde{\gamma}_z^n| \right] dz, \\ &\lesssim_t \int_0^t \mathbb{E}[|\hat{\gamma}_z^n - \tilde{\gamma}_z^n|] dz + n^{-1/2} \int_0^t \mathbb{E}[\|\eta_t\|_{-2,1}^2]^{1/2} \mathbb{E}[|\hat{\gamma}_z^n - \tilde{\gamma}_z^n|^2]^{1/2} dz \\ &\lesssim_t \int_0^t \mathbb{E}[|\hat{\gamma}_z^n - \tilde{\gamma}_z^n|] dz + n^{-1} \end{aligned}$$

where we used (49).

- Study of  $B_2^n(t)$ . From Taylor's inequality it follows that

$$\left\| \Psi(\cdot, \tilde{\gamma}_z^n) - \Psi(\cdot, \bar{\gamma}(z)) - \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \frac{\Gamma_z}{\sqrt{n}} \right\|_{\mathcal{C}_b^2} \leq \|\Psi\|_{\mathcal{C}_b^4} \frac{|\Gamma_z|^2}{n},$$

and so, using Lemma VI.5, Cauchy-Schwarz inequality and finally (49) and (57), we have

$$\begin{aligned} B_2^n(t) &\leq h_\infty(t) \int_0^t \mathbb{E} \left[ \left( 1 + \frac{C}{\sqrt{n}} \|\eta_t\|_{-2,1} \right) \|\Psi\|_{\mathcal{C}_b^4} \frac{|\Gamma_z|^2}{n} \right] dz, \\ &\lesssim_t n^{-1} \int_0^t \mathbb{E}[|\Gamma_z|^2] dz + n^{-3/2} \int_0^t \mathbb{E}[\|\eta_t\|_{-2,1}^2]^{1/2} \mathbb{E}[|\Gamma_z|^4]^{1/2} dz, \\ &\lesssim_t n^{-1}. \end{aligned}$$

- Study of  $B_3^n(t)$ . By definition of  $\hat{u}^n = u - n^{-1/2}\eta$  and doing as above, we have

$$B_3^n(t) \leq h_\infty(t) n^{-1} \int_0^t \mathbb{E} \left[ \|\eta_t\|_{-2,1} \left\| \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(z)) \right\|_{\mathcal{C}_b^2} |\Gamma_z| \right] dz \lesssim_t n^{-1}.$$

Finally, applying Lemma B.1 gives (98) and ends the proof.  $\square$

We are now in position to prove Proposition VI.4. Let us use the decomposition

$$\mathbb{E} [|r_t^n(\varphi)|] \leq \int_0^t A_1^n(z) + A_2^n(z) + A_3^n(z) dz,$$

where

$$\begin{cases} A_1^n(t) := \mathbb{E} \left[ \left| \left\langle \hat{u}_t^n - u_t, (L_t - \hat{L}_t^n) \varphi \right\rangle \right| \right], \\ A_2^n(t) := \mathbb{E} \left[ \left| \left\langle u_t, \left[ \Psi(\cdot, \bar{\gamma}(t)) + \frac{\partial \Psi}{\partial y}(\cdot, \bar{\gamma}(t)) n^{-1/2} \Gamma_t - \Psi(\cdot, \check{\gamma}_t^n) \right] R \varphi \right\rangle \right| \right], \\ A_3^n(t) := \mathbb{E} \left[ \left| \left\langle u_t, [\Psi(\cdot, \hat{\gamma}_t^n) - \Psi(\cdot, \check{\gamma}_t^n)] R \varphi \right\rangle \right| \right]. \end{cases}$$

- Study of  $A_1^n(t)$ . Using that  $\hat{u}_t^n - u_t = n^{-1/2} \eta_t$ ,  $\Psi$  belongs to  $\mathcal{C}_b^3$ , and an inequality similar to the second line of (77), we have

$$A_1^n(t) \lesssim_t n^{-1/2} \mathbb{E} \left[ \|\eta_t\|_{-2,1} \|\Psi\|_{\mathcal{C}_b^3} |\bar{\gamma}(t) - \hat{\gamma}_t^n| \right] \|\varphi\|_{2,1}.$$

Cauchy-Schwarz inequality with Equations (49) and (99) gives  $A_1^n(t) \lesssim_t n^{-1}$ .

- Study of  $A_2^n(t)$ . From Taylor's inequality and then (31) and (57), it follows that

$$A_2^n(t) \lesssim_t n^{-1} \|\Psi\|_{\mathcal{C}_b^2} \mathbb{E} [|\Gamma_t|^2] \|\varphi\|_{\infty} \lesssim_t n^{-1} \|\varphi\|_{2,1}.$$

- Study of  $A_3^n(t)$ . Finally, using (98) and then (31), we have

$$A_3^n(t) \lesssim_t \text{Lip}(\Psi) \mathbb{E} [|\hat{\gamma}_t^n - \check{\gamma}_t^n|] \|\varphi\|_{\infty} \lesssim_t n^{-1} \|\varphi\|_{2,1}.$$

Gathering the computations above, we prove the first part of Proposition VI.4. For the second part, let us use the decomposition,

$$\mathbb{E} \left[ \left| \text{DM}_t(\varphi_1, \varphi_2) - \hat{\text{DM}}_t(\varphi_1, \varphi_2) \right| \right] \leq \frac{1}{n} \int_0^t B_1^n(z) + B_2^n(z) dz,$$

where

$$\begin{cases} B_1^n(t) := \mathbb{E} \left[ \left| \left\langle u_t, [\Psi(\cdot, \hat{\gamma}_t^n) - \Psi(\cdot, \bar{\gamma}(t))] \varphi_1 \varphi_2 \right\rangle \right| \right], \\ B_2^n(t) := n^{-1/2} \mathbb{E} \left[ \left| \left\langle \eta_t, \Psi(\cdot, \hat{\gamma}_t^n) \varphi_1 \varphi_2 \right\rangle \right| \right]. \end{cases}$$

- Study of  $B_1^n(t)$ . Using the Lipschitz continuity of  $\Psi$  and then (99), we have

$$B_1^n(t) \leq \text{Lip}(\Psi) \|\varphi_1\|_{\infty} \|\varphi_2\|_{\infty} \mathbb{E} [|\hat{\gamma}_t^n - \bar{\gamma}(t)|] \lesssim_t n^{-1/2} \|\varphi_1\|_{\infty} \|\varphi_2\|_{\infty}.$$

- Study of  $B_2^n(t)$ . Using Lemma IV.1, we have,

$$|\langle \eta_t, \Psi(\cdot, \hat{\gamma}_t^n) \varphi_1 \varphi_2 \rangle| \leq C \|\eta_t\|_{-2,1} \|\Psi(\cdot, \hat{\gamma}_t^n)\|_{\mathcal{C}_b^2(\mathbb{R}_+)} \|\varphi_1 \varphi_2\|_{2,1}.$$

Yet,  $\|\Psi(\cdot, \hat{\gamma}_t^n)\|_{\mathcal{C}_b^2(\mathbb{R}_+)} \leq \|\Psi\|_{\mathcal{C}_b^2(\mathbb{R}_+ \times \mathbb{R})}$  and, by combining Lemma IV.1 and (31),  $\|\varphi_1 \varphi_2\|_{2,1} \leq C \|\varphi_1\|_{\mathcal{C}_b^2} \|\varphi_2\|_{2,1} \leq C \|\varphi_1\|_{3,1} \|\varphi_2\|_{3,1}$ . Hence, Equation (49) gives  $B_2^n(t) \lesssim_t n^{-1/2} \|\varphi_1\|_{3,1} \|\varphi_2\|_{3,1}$ . Gathering the computations above ends the proof.

## B Lemmas

The following lemma is a generalization of the standard Grönwall lemma.

**Lemma B.1.** *Let  $f, g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be two locally bounded non-negative measurable functions. Assume that for all  $t \geq 0$ ,*

$$f(t) \lesssim_t g(t) + \int_0^t f(s) ds. \quad (101)$$

*Then, for any  $\theta \geq 0$ ,  $\sup_{t \in [0, \theta]} f(t) \lesssim_{\theta} \sup_{t \in [0, \theta]} g(t)$ .*

*Proof.* For a fixed  $\theta$ , Equation (101) implies that there exists a constant  $C$  such that for all  $t \leq \theta$ ,  $f(t) \leq C(\sup_{t \in [0, \theta]} g(t) + \int_0^t f(s) ds)$ . Hence, standard Grönwall's inequality gives  $\sup_{t \in [0, \theta]} f(t) \leq C \sup_{t \in [0, \theta]} g(t) e^{C\theta}$  which ends the proof.  $\square$

The next lemma proves continuity in time for the law of the age process associated with a point process.

**Lemma B.2.** *Assume that  $N$  admits the bounded  $\mathbb{F}$ -intensity  $\lambda_t$  and satisfy Assumption  $(\mathcal{A}_\infty^{u_0})$ . Denote by  $(S_t)_{t \geq 0}$  its associated age process. Then, the law of  $S_t$  denoted by  $w_t$  is such that  $t \mapsto w_t$  belongs to  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha})$  for any  $\alpha > 1/2$ .*

*Proof.* This continuity result comes from the fact that the probability that  $N$  has a point in an interval goes to 0 as the size of the interval goes to 0. Fix  $\alpha > 1/2$  and let  $t, t'$  be positive real numbers. First, remark that  $S_{t+t'} = S_t + t'$  as soon as there is no point of  $N$  in the interval  $[t, t + t']$  and so one has for all  $\varphi$  in  $\mathcal{W}_0^{2, \alpha}$ ,

$$|\varphi(S_{t+t'}) - \varphi(S_t)| \leq \|D_{S_{t+t'}, S_t}\|_{-2, \alpha} \|\varphi\|_{2, \alpha} + (|\varphi(S_{t+t'})| + |\varphi(S_t)|) \mathbf{1}_{N([t, t+t']) \neq 0}.$$

The bound obtained in Lemma IV.3 for the operator  $D_{x, y}$  is too rough here. We need a finer bound: it holds that there exists a constant  $C$  such that  $\|D_{x, y}\|_{-2, \alpha} \leq C|x - y|(1 + \max(|x|^\alpha, |y|^\alpha))$ . Indeed, by density, let us assume that  $\varphi$  is  $\mathcal{C}^\infty$  with compact support and remark that

$$\begin{aligned} |\varphi(x) - \varphi(y)| &\leq |x - y| \sup_{z, |z| \leq \max(|x|, |y|)} |\varphi'(z)| \leq |x - y|(1 + \max(|x|^\alpha, |y|^\alpha)) \|\varphi\|_{\mathcal{C}^{1, \alpha}} \\ &\leq C|x - y|(1 + \max(|x|^\alpha, |y|^\alpha)) \|\varphi\|_{2, \alpha}, \end{aligned}$$

where we used (31) in the last inequality. Since  $(\mathcal{A}_\infty^{u_0})$  is satisfied,  $S_{t+t'}$  and  $S_t$  are upper bounded by  $M_{S_0} + t + t'$  so that

$$\begin{cases} \|D_{S_{t+t'}, S_t}\|_{-2, \alpha} \leq C t' (1 + (M_{S_0} + t + t')^\alpha) \\ (|\varphi(S_{t+t'})| + |\varphi(S_t)|) \leq 2(1 + (M_{S_0} + t + t')^\alpha) \|\varphi\|_{\mathcal{C}^{0, \alpha}}. \end{cases}$$

Hence, (31) gives

$$|\varphi(S_{t+t'}) - \varphi(S_t)| \leq C(t' + \mathbf{1}_{N([t, t+t']) \neq 0}) \|\varphi\|_{2, \alpha}.$$

Yet,  $\mathbb{P}(N([t, t + t']) \neq 0) \leq \mathbb{E}[N([t, t + t'])] = \mathbb{E}[\int_t^{t+t'} \lambda_z dz]$  goes to 0 as  $t'$  goes to 0. The same argument for  $t' < 0$  gives continuity.  $\square$

The three lemmas below are used to get the limit equation satisfied by the fluctuations.

**Lemma B.3.** *Let  $h$  be a locally bounded function and  $(\varphi_t)_{t \geq 0}$  be a family of test functions in  $\mathcal{W}_0^{2, \alpha}$  such that  $t \mapsto \|\varphi_t\|_{2, \alpha}$  is locally bounded. Then,  $F : g \mapsto \int_0^t h(t - z) \langle g(z), \varphi_z \rangle dz$  is a mapping from  $\mathcal{D}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha})$  to  $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$  which is continuous at every point  $g_0$  in  $\mathcal{C}(\mathbb{R}_+, \mathcal{W}_0^{-2, \alpha})$ .*

*Proof.* Let  $(g_n)_{n \geq 1}$  be any sequence such that  $g_n \rightarrow g_0$  for the Skorokhod topology. Since  $g_0$  is continuous, the convergence also holds true for the local uniform topology [21, Proposition VI.1.17.]. We have for all  $\theta \geq 0$ ,

$$\sup_{t \in [0, \theta]} |F(g_n)(t) - F(g_0)(t)| \leq \sup_{z \in [0, \theta]} h(z) \sup_{z \in [0, \theta]} \|g_n(z) - g_0(z)\|_{-2, \alpha} \sup_{z \in [0, \theta]} \|\varphi_z\|_{2, \alpha}. \quad (102)$$

Yet, the right hand side of (102) goes to 0 as  $n$  goes to infinity, which ends the proof.  $\square$



**Lemma B.4.** *Assume that  $(g^n)_{n \geq 1}$  converges to  $g$  for the Skorokhod topology in  $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$ . If  $h$  satisfies  $(\mathcal{A}_{\text{H\"ol}}^h)$  and  $f$  is locally bounded, then*

$$\int_0^t h(t-z)f(z)g^n(z)dz \xrightarrow{n \rightarrow +\infty} \int_0^t h(t-z)f(z)g(z)dz,$$

as functions of  $t$  in  $\mathcal{C}(\mathbb{R}_+, \mathbb{R})$  for the local uniform topology. In particular, the application  $F$  from  $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$  to  $\mathcal{D}(\mathbb{R}_+, \mathbb{R})$  defined by

$$F(g)(t) := \int_0^t h(t-z)f(z)g(z)dz,$$

is continuous.

*Proof.* Let  $c^n(t) := \int_0^t h(t-z)f(z)g^n(z)dz$  and  $c(t) := \int_0^t h(t-z)f(z)g(z)dz$ . Assume for a while that  $h(0) = 0$  and extend the function  $h$  to the whole real line by setting 0 on the negative real numbers. Then, for all  $t, \delta \geq 0$ ,

$$\begin{aligned} |c^n(t+\delta) - c^n(t)| &\leq \int_0^{t+\delta} |h(t+\delta-z) - h(t-z)||f(z)||g^n(z)|dz \\ &\leq (t+\delta)\text{H\"ol}(h) \sup_{z \in [0, t+\delta]} |f(z)| \sup_{z \in [0, t+\delta]} |g^n(z)|\delta^{\beta(h)}. \end{aligned}$$

Yet, since  $g^n$  is convergent, we have  $\sup_{n \geq 1} \sup_{z \in [0, t+\delta]} |g^n(z)| < +\infty$  (see [21, Proposition VI.2.4.] for instance) which implies that for all  $\theta \geq 0$ ,

$$\sup_{n \geq 1} \sup_{t \in [0, \theta]} |c^n(t+\delta) - c^n(t)| \rightarrow 0 \quad \text{as } \delta \rightarrow 0.$$

Hence, the sequence  $(c^n)_{n \geq 1}$  is uniformly continuous. Moreover, for all  $n \geq 1$ ,  $c^n(0) = 0$  and the uniform continuity gives the uniform boundedness

$$\sup_{n \geq 1} \sup_{t \in [0, \theta]} |c^n(t)| < +\infty.$$

Then, Ascoli-Arzelà theorem implies that the sequence  $(c^n)_{n \geq 1}$  is relatively compact. It only remains to identify the limit for all  $t \geq 0$ . Yet, as a consequence of the dominated convergence and the fact that for almost every  $z$ ,  $g^n(z) \rightarrow g(z)$ , we have  $\int_0^t h(t-z)f(z)g^n(z)dz \rightarrow \int_0^t h(t-z)f(z)g(z)dz$ .

Now, if  $h(0) \neq 0$ , one can use the following decomposition,

$$c^n(t) = \int_0^t (h(t-z) - h(0))f(z)g^n(z)dz + h(0) \int_0^t f(z)g^n(z)dz.$$

The first term is convergent thanks to what we have done in the case  $h(0) = 0$  whereas the convergence of the second one is simpler and left to the reader.  $\square$

**Acknowledgement** This research was partly supported by the french Agence Nationale de la Recherche (ANR 2011 BS01 010 01 projet Calibration), by the interdisciplinary axis MTC-NSC of the University of Nice Sophia-Antipolis and by the Labex MME-DII (ANR11-LBX-0023-01). The author would like to thank François Delarue for helpful discussions which improved this paper.

## References

- [1] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003.
- [2] E. Bacry, K. Dayri, and J.-F. Muzy. Non-parametric kernel estimation for symmetric Hawkes processes. Application to high frequency financial data. *The European Physical Journal B-Condensed Matter and Complex Systems*, 85(5):1–12, 2012.
- [3] E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy. Scaling limits for Hawkes processes and application to financial statistics, Feb. 2012.
- [4] P. Bao, H.-W. Shen, X. Jin, and X.-Q. Cheng. Modeling and Predicting Popularity Dynamics of Microblogs using Self-Excited Hawkes Processes. *arXiv preprint arXiv:1503.02754*, 2015.
- [5] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [6] P. Brémaud. *Point processes and queues*. Springer-Verlag, New York, 1981. Martingale dynamics, Springer Series in Statistics.
- [7] N. Brunel and V. Hakim. Fast global oscillations in networks of integrate-and-fire neurons with low firing rates. *Neural computation*, 11(7):1621–1671, 1999.
- [8] M. A. Buice and C. C. Chow. Dynamic finite size effects in spiking neural networks. *PLoS Comput Biol*, 9(1):e1002872, 2013.
- [9] J. Chevallier. Mean-field limit of generalized Hawkes processes. *arXiv preprint arXiv:1510.05620*, 2015.
- [10] J. Chevallier. *Modelling large neural networks via Hawkes processes*. PhD thesis, Université Nice Sophia Antipolis, 2016. Manuscript available [here](#).
- [11] J. Chevallier, M. J. Cáceres, M. Doumic, and P. Reynaud-Bouret. Microscopic approach of a time elapsed neural model. *Mathematical Models and Methods in Applied Sciences*, 25(14):2669–2719, 2015.
- [12] R. Crane and D. Sornette. Robust dynamic classes revealed by measuring the response function of a social system. *Proceedings of the National Academy of Sciences*, 105(41):15649–15653, 2008.
- [13] G. Dumont et al. Private communication about ongoing work.
- [14] G. Dumont, J. Henry, and C. O. Tarniceriu. Theoretical connections between mathematical neuronal models corresponding to different expressions of noise. *arXiv preprint arXiv:1602.03764*, 2016.
- [15] O. Faugeras and J. Maclaurin. Asymptotic description of stochastic neural networks. I. Existence of a large deviation principle. *Comptes Rendus Mathématique*, 352(10):841–846, 2014.
- [16] B. Fernandez and S. Méléard. A Hilbertian approach for fluctuations on the McKean-Vlasov model. *Stochastic Process. Appl.*, 71(1):33–53, 1997.
- [17] W. Gerstner and W. M. Kistler. *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge university press, 2002.
- [18] R. D. Gill, N. Keiding, and P. K. Andersen. *Statistical models based on counting processes*. Springer, 1997.
- [19] G. Gusto and S. Schbath. FADO: A Statistical Method to Detect Favored or Avoided Distances between Occurrences of Motifs using the Hawkes’ Model. *Statistical Applications in Genetics and Molecular Biology*, 4(1), 2005.
- [20] A. G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.
- [21] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 2003.
- [22] A. Joffe and M. Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Advances in Applied Probability*, pages 20–65, 1986.
- [23] B. Jourdain and S. Méléard. Propagation of chaos and fluctuations for a moderate model with smooth initial data. In *Annales de l’IHP Probabilités et statistiques*, volume 34, pages 727–766, 1998.

- [24] Y. Y. Kagan. Statistical distributions of earthquake numbers: consequence of branching process. *Geophysical Journal International*, 180(3):1313–1328, 2010.
- [25] T. J. Liniger. *Multivariate hawkes processes*. PhD thesis, Diss., Eidgenössische Technische Hochschule ETH Zürich, Nr. 18403, 2009, 2009.
- [26] E. Luçon and W. Stannat. Transition from Gaussian to non-Gaussian fluctuations for mean-field diffusions in spatial interaction. *arXiv preprint arXiv:1502.00532*, 2015.
- [27] M. Mattia and P. Del Giudice. Finite-size dynamics of inhibitory and excitatory interacting spiking neurons. *Physical Review E*, 70(5):052903, 2004.
- [28] S. Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 42–95. Springer, Berlin, 1996.
- [29] F. Merlevède, M. Peligrad, et al. Rosenthal-type inequalities for the maximum of partial sums of stationary processes and examples. *The Annals of Probability*, 41(2):914–960, 2013.
- [30] G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita. Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493), 2011.
- [31] Y. Ogata. Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2):379–402, 1998.
- [32] K. Pakdaman, B. Perthame, and D. Salort. Dynamics of a structured neuron population. *Nonlinearity*, 23(1):55, 2010.
- [33] K. Pakdaman, B. Perthame, and D. Salort. Relaxation and self-sustained oscillations in the time elapsed neuron network model. *SIAM Journal on Applied Mathematics*, 73(3):1260–1279, 2013.
- [34] R. Rebolledo. Central limit theorems for local martingales. *Z. Wahrsch. Verw. Gebiete*, 51(3):269–286, 1980.
- [35] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion (Grundlehren der mathematischen Wissenschaften)*. Springer-Verlag, 3rd edition, 1999.
- [36] P. Reynaud-Bouret, S. Schbath, et al. Adaptive estimation for Hawkes processes; application to genome analysis. *The Annals of Statistics*, 38(5):2781–2822, 2010.
- [37] M. G. Riedler, M. Thieullen, and G. Wainrib. Limit theorems for infinite-dimensional piecewise deterministic Markov processes. Applications to stochastic excitable membrane models. *Electron. J. probab.*, 17(55):1–48, 2012.
- [38] G. R. Shorack and J. A. Wellner. *Empirical processes with applications to statistics*, volume 59. Siam, 2009.
- [39] A.-S. Sznitman. Topics in propagation of chaos. In *École d’Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Math.*, pages 165–251. Springer, Berlin, 1991.
- [40] C. Tuleau-Malot, A. Rouis, F. Grammont, and P. Reynaud-Bouret. Multiple Tests Based on a Gaussian Approximation of the Unitary Events Method with delayed coincidence count. appearing in *Neural Computation*, 26:7, 2014.
- [41] G. Wainrib. *Randomness in neurons: a multiscale probabilistic analysis*. PhD thesis, PhD thesis, 2010.
- [42] K. Yosida. *Functional analysis*, volume 123 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin-New York, sixth edition, 1980.
- [43] L. Zhu. *Nonlinear Hawkes Processes*. PhD thesis, New York University, 2013.