



**HAL**  
open science

## An expansion formula for Hawkes processes and application to cyber-insurance derivatives \*

Caroline Hillairet, Anthony Réveillac, Mathieu Rosenbaum

### ► To cite this version:

Caroline Hillairet, Anthony Réveillac, Mathieu Rosenbaum. An expansion formula for Hawkes processes and application to cyber-insurance derivatives \*. *Stochastic Processes and their Applications*, 2023, 160, pp.89-119. 10.1016/j.spa.2023.02.012 . hal-03189601

**HAL Id: hal-03189601**

**<https://insa-toulouse.hal.science/hal-03189601v1>**

Submitted on 4 Apr 2021

**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.

# An expansion formula for Hawkes processes and application to cyber-insurance derivatives \*

Caroline Hillairet<sup>†</sup>     Anthony Réveillac<sup>‡</sup>     Mathieu Rosenbaum<sup>§</sup>

April 4, 2021

## Abstract

In this paper we provide an expansion formula for Hawkes processes which involves the addition of jumps at deterministic times to the Hawkes process in the spirit of the well-known integration by parts formula (or more precisely the Mecke formula) for Poisson functional. Our approach allows us to provide an expansion of the premium of a class of cyber insurance derivatives (such as reinsurance contracts including generalized Stop-Loss contracts) or risk management instruments (like Expected Shortfall) in terms of so-called shifted Hawkes processes. From the actuarial point of view, these processes can be seen as "stressed" scenarios. Our expansion formula for Hawkes processes enables us to provide lower and upper bounds on the premium (or the risk evaluation) of such cyber contracts and to quantify the surplus of premium compared to the standard modeling with a homogenous Poisson process.

**Keywords:** Hawkes process; Malliavin calculus; pricing formulae; cyber insurance derivatives.

## 1 Introduction

In actuarial science, the classical Cramer-Lundberg model used to describe the surplus process of an insurance portfolio relies on the assumptions of the claims arrival being modeled by a Poisson process, and of independence among claim sizes and between claim sizes and claim inter-occurrence times. However, in practice those assumptions are often too restrictive and there is a need for more general models. A first generalization in the modeling of claims arrivals consists in using Cox processes (also known as doubly stochastic Poisson processes), in the context of ruin theory such as in Albrecher and Asmussen (2006) [2], or for pricing stop-loss catastrophe insurance contract and catastrophe insurance derivatives such as in Dassios and Jang (2003) [13] and (2013) [14], or Hillairet et al. (2018) [11].

---

\*This research is supported by a grant of the French National Research Agency (ANR), "Investissements d'Avenir" (LabEx Ecodec/ANR-11-LABX-0047) and the Joint Research Initiative "Cyber Risk Insurance: actuarial modeling" with the partnership of AXA Research Fund.

<sup>†</sup>ENSAE Paris, CREST UMR 9194, 5 avenue Henry Le Chatelier 91120 Palaiseau, France. Email: [caroline.hillairet@ensae.fr](mailto:caroline.hillairet@ensae.fr)

<sup>‡</sup>INSA de Toulouse, IMT UMR CNRS 5219, Université de Toulouse, 135 avenue de Rangueil 31077 Toulouse Cedex 4 France. Email: [anthony.reveillac@insa-toulouse.fr](mailto:anthony.reveillac@insa-toulouse.fr)

<sup>§</sup>Ecole Polytechnique, CMAP UMR 7641, Route de Saclay, 91120 Palaiseau, France. Email: [mathieu.rosenbaum@polytechnique.edu](mailto:mathieu.rosenbaum@polytechnique.edu)

Besides, self-exciting effects have been highlighted in cyber risk, in favor of modeling the claims arrivals by a Hawkes process, that is adapted to model aftershocks of cyber attacks. Such processes have been recently used in the cyber security field, for instance by Peng et al. (2016) [?] who focused on extreme cyber attacks rates. Baldwin et al. (2017) also studied in [?] the threats to 10 important IP services, using industry standard SANS data, and they claim that Hawkes processes provides the adequate modeling of cyber attacks into information systems because they capture both shocks and persistence after shocks that may form attack contagion. In cyber insurance, the statistical analysis of Bessy et al. (2020) [?] on the public Privacy Rights Clearinghouse database highlights the ability of Hawkes models to capture self-excitation and interactions of data-breaches. Although the application in this paper focuses on cyber risk, this methodology can be applied to other risks presenting self-exciting properties, such as credit risk.

Hawkes processes, which have wide applications in many fields (such as seismology, finance, neuroscience or social networks), are also beginning to get studied in actuarial sciences. Magnusson Thesis (2015) [13] is dedicated to Hawkes processes with exponential decay and their application to insurance. Dassios and Zhao (2012) [5] consider the ruin problem in a model where the claims arrivals follow a Hawkes process with decreasing exponential kernel. Stabile and Torrisi (2010) [18] study the asymptotic behavior of infinite and finite horizon ruin probabilities, assuming non-stationary Hawkes claims arrivals and under light-tailed conditions on the claims. Gao and Zhu (2018) [8] establish large deviations results for Hawkes processes with an exponential kernel and develop approximations for finite-horizon ruin probabilities. Swishchuk (2018) [?] applies limit theorems for risk model based on general compound Hawkes process to compute premium principles and ruin times.

Rather than giving explicit computations for probabilities of ruin, our paper proposes to give pricing formulae of insurance contracts, such as Stop-Loss contracts. Stop-Loss is a non-proportional type of reinsurance and works similarly to excess-of-loss reinsurance. While excess-of-loss is related to single loss amounts, either per risk or per event, stop-loss covers are related to the total amount of claims in a year. The reinsurer pays the part of the total loss that exceeds a certain amount  $K$ . The reinsurer's liability is often limited to a given threshold. Stop-loss reinsurance offers protection against an increase in either or both severity and frequency of a company's loss experience. Various approximations of stop-loss reinsurance premiums are described in literature, some of them assuming certain dependence structure, such as Gerber (1982) [9], Albers (1999) [1], De Lourdes Centeno (2005) [6] or Reijnen et al. (2005) [17].

Stop-loss contracts are the paradigm of reinsurance contracts, but we aim at dealing with more general payoffs (of maturity  $T$ ), whose valuation involves the computation of the quantity of the form

$$\mathbb{E}[K_T h(L_T)] \text{ where } \begin{cases} K_T \text{ is the effective loss covered by the reinsurance company,} \\ L_T \text{ is the loss quantity that activates the contract.} \end{cases}$$

For example, for stop loss contracts  $h(L_T) = \mathbf{1}_{\{L_T \geq K\}}$ . Similarly, this methodology can be applied to valuation of credit derivatives, such as Credit Default Obligation tranches. It also goes beyond the analysis of pricing and finds application in the computation of the expected shortfall of contingent claims : the expected shortfall is a useful risk measure, that takes into account the size of the expected loss above the value at risk. We refer to [11] for more details on those analogies.

Our paper considers cyber insurance contracts, with underlying a cumulative loss indexed by a Hawkes process. We propose to compute an explicit closed form pricing formula. Although Monte Carlo procedures are certainly the most efficient to compute numerically the premium of such general contracts, the closed form expansion formula we develop allows to compute lower and upper bounds for the premium, and to quantify the surplus of premium compared to the standard modeling with a homogenous Poisson process. A correct estimate of this surplus of premium is a crucial challenge for cyber insurance, to avoid an underestimation of the risk induced by a Poisson process model. Such formula could also be efficient for sensitivity analysis. The formula relies on the so-called shifted Hawkes processes, which, from the actuarial point of view, can be seen as "stressed" scenarios.

From the probabilistic point of view, the quantity  $\mathbb{E}[K_T h(L_T)]$  can be expressed (conditioning with respect to the claims) as  $\mathbb{E}\left[\int_{(0,T]} Z_t dH_t F\right]$  where  $Z$  is a predictable process and  $F := h(L_T)$  is a functional of the Hawkes process. In the case where the counting process is a Poisson process (or a Cox process), Malliavin calculus enables one to transform this quantity. More precisely, to simplify the discussion, assume  $H$  is an homogeneous Poisson process with intensity  $\mu > 0$  (in other words the self-exciting kernel  $\Phi$  is put to 0), the Malliavin integration by parts formula allows us to derive that<sup>1</sup> :

$$\mathbb{E}\left[\int_{(0,T]} Z_t dH_t F\right] = \mu \int_0^T \mathbb{E}[Z_t F \circ \varepsilon_t^+] dt, \quad (1.1)$$

where the notation  $F \circ \varepsilon_t^+$  denotes the functional on the Poisson space where a deterministic jump is added to the paths of  $H$  at time  $t$ . This expression turns out to be particularly interesting from an actuarial point of view since adding a jump at some time  $t$  corresponds to realising a stress test by adding artificially a claim at time  $t$ . This approach has been followed in [11] for Cox processes (that is doubly stochastic Poisson processes with stochastic but independent intensity). Naturally, in case of a Poisson process, the additional jump at some time  $t$  only impacts the payoff of the contract by adding a new claim in the contract but it does not impact the dynamic of the counting process  $H$ .

The goal of this paper is two-fold:

1. First we provide in Theorem 3.13 a generalization of Equation (1.1) in case  $H$  is a Hawkes process. The main ingredient consists in using a representation of a Hawkes process known as the "Poisson embedding" (related to the "Thinning Algorithm") in terms of a Poisson process  $N$  on  $[0, T] \times \mathbb{R}_+$  to which the Malliavin integration by parts formula can be applied. As the adjunction of a jump at a given time impacts the dynamic of the Hawkes process, we refer to the obtained expression more to an "expansion" rather than an "integration by parts formula" for the Hawkes process, as it involves what we name "shifted Hawkes processes" for which jumps at deterministic times are added to the process accordingly to the self-exciting kernel  $\Phi$ . We refer to Theorem 3.13 and to Remark 3.14 for a discussion on this expansion and its

---

<sup>1</sup>Note that strictly speaking this formula is not the Malliavin integration by parts formula on the Poisson space as the stochastic integral  $\int_{(0,T]} Z_t dH_t$  is not exactly the divergence of  $Z$ , which explains why the dual operator on the right-hand side is not exactly the Malliavin difference operator. However, in case of predictable integrators  $Z$ , the classical integration by parts formula can be reduced to this form which is sufficient for our purpose.

link to the one obtained for homogeneous Poisson process.

2. Then, we apply our main result to the specific quantity  $\mathbb{E}[K_T h(L_T)]$  which is at the core for determining the premium of a large class of insurance derivatives or risk management instruments. Our main result on that regard is given in Theorem 4.3. As pointed out in the discussion at the beginning of Section 4.3, the shifted processes  $H^{v_n, \dots, v_1}$  (see Definition 3.6 for a precise statement) appearing in the form of the premium are of the same complexity than the original Hawkes process  $H$ . However, they exhibit deterministic jumps at some times  $v_1, \dots, v_n$  which are weighted by correlation factors of the form  $\Phi(v_i - v_{i-1})$ . In other words, this formula make appears  $n$ -jumps of the Hawkes process at some deterministic times. This provides an additional input compared to classical estimates for Hawkes processes for obtaining lower or upper bounds of their CDF in terms of the one of a Poisson process for instance. We benefit from this formulation to derive in Proposition 4.5 and Proposition 4.8 a lower and an upper bound respectively for the quantity  $\mathbb{E}[K_T h(L_T)]$ .

We proceed as follows. In the next section, we provide general notations and elements of Malliavin calculus on the (classical) Poisson space. In particular, the shift operators (which will play a central role in our analysis) on the Poisson space are introduced. We also explain the representation of a Hawkes process using the Poisson embedding. Section 3 provides the derivation of the expansion formula in Theorem 3.13. Note that it requires the introduction and analysis of what we named the shifted Hawkes processes resulting from the shifts on the Poisson space of the original Hawkes process (this material is presented in Section 3.1). Insurance contracts of interest are presented in Section 4 together with the main result for the representation of the premium of such contracts in Theorem 4.3. Lower and upper bounds for this premium are presented in Propostion 4.5 and Proposition 4.8, and in Corollaries 4.7 and 4.9. Finally, we postponed some technical material in Section 5.

## 2 Elements of stochastic analysis on the Poisson space, Hawkes process and thinning

This short section provides some generalities and elements of stochastic analysis on the Poisson space, and in particular the integration by parts formula for the Poisson process. Hawkes processes are also defined and their representation through the thinning procedure is presented. Throughout this paper  $T > 0$  denotes a fixed positive real number. For  $X$  a topological space, we set  $\mathcal{B}(X)$  the  $\sigma$ -algebra of Borelian sets.

### 2.1 Elements of stochastic analysis on the Poisson space

Let the space of configurations

$$\Omega^N := \left\{ \omega^N = \sum_{i=1}^n \delta_{t_i, \theta_i}, i = 1, \dots, n, 0 = t_0 < t_1 < \dots < t_n \leq T, \theta_i \in \mathbb{R}_+, n \in \mathbb{N} \cap \{+\infty\} \right\}.$$

Each path of a counting process is represented as an element  $\omega^N$  in  $\Omega^N$  which is a  $\mathbb{N}$ -valued measure on  $[0, T] \times \mathbb{R}_+$ . Let  $\mathcal{F}_T^N$  be the  $\sigma$ -field associated to the vague topology on  $\Omega^N$ , and  $\mathbb{P}^N$  the Poisson measure on  $\Omega^N$  under which the counting process  $N$  defined as [the canonical](#)

process on  $\Omega^N$  as

$$(N(\omega))([0, t] \times [0, b])(\omega) := \omega([0, t] \times [0, b]), \quad t \in [0, T], b \in \mathbb{R}_+,$$

is an homogeneous Poisson process with intensity one (so that  $N([0, t] \times [0, b])$  is a Poisson random variable with intensity  $bt$  for any  $(t, b) \in [0, T] \times \mathbb{R}_+$ ). We set  $\mathbb{F}^N := (\mathcal{F}_t^N)_{t \in [0, T]}$  the natural filtration of  $N$ , that is  $\mathcal{F}_t^N := \sigma(N(\mathcal{T} \times B))$ ,  $\mathcal{T} \subset \mathcal{B}([0, t])$ ,  $B \in \mathcal{B}(\mathbb{R}_+)$ . The expectation with respect to  $\mathbb{P}^N$  is denoted by  $\mathbb{E}[\cdot]$ .

One of the main ingredient in our approach will be the integration by parts formula for the Poisson process  $N$  and the shift operators defined below.

**Definition 2.1** (Shift operator). We define for  $(t, \theta)$  in  $[0, T] \times \mathbb{R}_+$  the measurable map

$$\begin{aligned} \varepsilon_{(t, \theta)}^+ : \Omega^N &\rightarrow \Omega^N \\ \omega &\mapsto \varepsilon_{(t, \theta)}^+(\omega), \end{aligned}$$

with  $(\varepsilon_{(t, \theta)}^+(\omega))(A) := \omega(A \setminus (t, \theta)) + \mathbf{1}_A(t, \theta)$ ,  $A \in \mathcal{B}([0, T] \times \mathbb{R}_+)$  and where

$$\mathbf{1}_A(t, \theta) := \begin{cases} 1, & \text{if } (t, \theta) \in A, \\ 0, & \text{else.} \end{cases}$$

**Remark 2.2.** Let  $(t_0, \theta_0)$  in  $(0, T) \times \mathbb{R}_+$ ,  $t_0 < s < t$ ,  $\mathcal{T} \in \{(s, t), (s, t], [s, t), [s, t]\}$  and  $B$  in  $\mathcal{B}(\mathbb{R}_+)$ . We have that

$$N \circ \varepsilon_{(t_0, \theta_0)}^+(\mathcal{T} \times B) = \varepsilon_{(t_0, \theta_0)}^+(N(\mathcal{T} \times B)) = N(\mathcal{T} \times B), \quad \mathbb{P} - a.s..$$

**Lemma 2.3.** Let  $t$  in  $[0, T]$  and  $F$  be an  $\mathcal{F}_t^N$ -measurable random variable. Let  $v > t$  and  $\theta_0 \geq 0$ . It holds that

$$F \circ \varepsilon_{(v, \theta_0)}^+ = F, \quad \mathbb{P} - a.s..$$

*Proof.* The proof consists in noticing that a  $\mathcal{F}_t^N$ -measurable random variable is a functional of  $N_{\cdot \wedge t}$ .  $\square$

Similarly, for any  $\omega$  in  $\Omega^N$  and  $(t, \theta)$  in  $[0, T] \times \mathbb{R}_+$ , we set the measure  $\varepsilon_{(t, \theta)}^-(\omega)$  defined as

$$(\varepsilon_{(t, \theta)}^-(\omega))(A) := \omega(A \setminus \{(t, \theta)\}), \quad A \in \mathcal{B}([0, T] \times \mathbb{R}_+).$$

We conclude this section with the integration by parts formula (ore more specifically the Mecke formula) on the Poisson space (see [16, Corollaire 5] or [14]).

**Proposition 2.4** (Mecke's Formula). Let  $F$  be in  $L^1(\Omega, \mathcal{F}_T^N, \mathbb{P})$  and  $\mathcal{Z} = (\mathcal{Z}(t, \theta))_{t \in [0, T], \theta \in \mathbb{R}_+}$  be a  $\mathbb{F}^N$ -adapted process<sup>2</sup> with  $\mathbb{E} \left[ \int_0^T |\mathcal{Z}(t, \theta)| dt \right] < +\infty$  and such that

$$\mathcal{Z}(t, \theta) \circ \varepsilon_{(t, \theta)}^- = \mathcal{Z}(t, \theta), \quad \mathbb{P} \otimes dt \otimes d\theta, a.e.. \quad (2.1)$$

We have that

$$\mathbb{E} \left[ F \int_{[0, T] \times \mathbb{R}_+} \mathcal{Z}(t, \theta) N(dt, d\theta) \right] = \mathbb{E} \left[ \int_0^T \int_{\mathbb{R}_+} \mathcal{Z}(t, \theta) (F \circ \varepsilon_{(t, \theta)}^+) dt d\theta \right].$$

For the expansion formula in Theorem 3.13, the Mecke's formula will be applied for the process  $\mathcal{Z}(t, \theta) = Z_t \mathbf{1}_{\{\theta \leq \Lambda_t\}}$  where  $\Lambda$  will denote the intensity of the Hawkes process (we refer to the proof of Theorem 3.13 for a more precise relation between this formula and our result).

<sup>2</sup>Note that only measurability with respect to  $\mathcal{F}_T^N$  is necessary here.

## 2.2 Representation of Hawkes processes

We first recall the definition of a Hawkes process.

**Definition 2.5** (Standard Hawkes process, [10]). Let  $(\Omega, \mathcal{F}_T, \mathbb{P}, \mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]})$  be a filtered probability space,  $\mu > 0$  and  $\Phi : [0, T] \rightarrow \mathbb{R}_+$  be a bounded non-negative map with  $\|\Phi\|_1 < 1$ . A standard Hawkes process  $H := (H_t)_{t \in [0, T]}$  with parameters  $\mu$  and  $\Phi$  is a counting process such that

- (i)  $H_0 = 0$ ,  $\mathbb{P}$ -a.s.,
- (ii) its ( $\mathbb{F}$ -predictable) intensity process is given by

$$\Lambda_t := \mu + \int_{(0, t)} \Phi(t-s) dH_s, \quad t \in [0, T],$$

that is for any  $0 \leq s \leq t \leq T$  and  $A \in \mathcal{F}_s$ ,

$$\mathbb{E}[\mathbf{1}_A(H_t - H_s)] = \mathbb{E}\left[\int_{(s, t]} \mathbf{1}_A \Lambda_r dr\right].$$

This definition can be generalized as follows, by considering a starting date  $v > 0$  and allowing starting points (for the Hawkes process itself and its intensity) that are  $\mathcal{F}_v$ -measurable (that is that are known at time  $v$ ).

**Definition 2.6** (Generalized Hawkes process). Let  $(\Omega, \mathcal{F}_T, \mathbb{P}, \mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]})$  be a filtered probability space. Let  $v$  in  $[0, T]$ ,  $h^v$  be a  $\mathcal{F}_v$ -measurable random variable with valued in  $\mathbb{N}$ ,  $\mu^v := (\mu^v(t))_{t \in [v, T]}$  a positive map such that  $\mu^v(t)$  is  $\mathcal{F}_v$ -measurable for any  $t \geq v$ , and  $\Phi : [0, T] \rightarrow \mathbb{R}_+$  be a bounded non-negative map with  $\|\Phi\|_1 < 1$ . A Hawkes process on  $[v, T]$  with parameters  $\mu^v$ ,  $h^v$  and  $\Phi : [0, T] \rightarrow \mathbb{R}_+$  is a ( $\mathbb{F}$ -adapted) counting process  $H := (H_t)_{t \in [v, T]}$  such that

- (i)  $H_v = h^v$ ,  $\mathbb{P}$ -a.s.,
- (ii) its ( $\mathbb{F}$ -predictable) intensity process is given by

$$\Lambda_t := \mu^v(t) + \int_{(v, t)} \Phi(t-s) dH_s, \quad t \in [v, T],$$

that is for any  $v \leq s \leq t \leq T$  and  $A \in \mathcal{F}_s$ ,

$$\mathbb{E}[\mathbf{1}_A(H_t - H_s) | \mathcal{F}_v] = \mathbb{E}\left[\int_{(s, t]} \mathbf{1}_A \Lambda_r dr \middle| \mathcal{F}_v\right].$$

Our main result relies on the following representation of a Hawkes process known as the "Poisson embedding" and related to the "Thinning Algorithm" (see *e.g.* [2, 3, 4, 15] and references therein). To this end we consider the filtered probability space  $(\Omega, \mathcal{F}_T, \mathbb{P}, \mathbb{F})$  as follows

$$\Omega := \Omega^N, \quad \mathcal{F}_T := \mathcal{F}_T^N, \quad \mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}, \quad \mathcal{F}_t := \mathcal{F}_t^N, \quad t \in [0, T], \quad \mathbb{P} := \mathbb{P}^N.$$

**Theorem 2.7.** Let  $v$  in  $[0, T]$  and  $(\mu^v(t))_{t \in [v, T]}$  be a non-negative stochastic process such that for any  $t \geq v$ ,  $\mu^v(t)$  is a  $\mathcal{F}_v^N$ -measurable random variable. Let in addition  $h^v$  be a  $\mathcal{F}_v^N$ -measurable random with values in  $\mathbb{N}$ . On the probability space  $(\Omega, \mathcal{F}_T, \mathbb{P})$ , the SDE (2.3) with

$$\begin{cases} \hat{H}_t^v = h^v + \int_{(v, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \hat{\Lambda}_s^v\}} N(ds, d\theta), & t \in [v, T] \\ \hat{\Lambda}_t^v = \mu^v(t) + \int_{(v, t)} \Phi(t-u) d\hat{H}_u^v \end{cases} \quad (2.2)$$

admits a unique  $\mathbb{F}^N$ -adapted solution  $\hat{H}$ . Uniqueness is understood in the strong sense, that is, if  $\hat{H}^1, \hat{H}^2$  denote two solutions then

$$\mathbb{P} \left[ \sup_{t \in [0, T]} |\hat{H}_t^1 - \hat{H}_t^2| \neq 0 \right] = 0.$$

This result can be deduced from the general construction of point process of Jacod et al [12]. Nevertheless for seek of completeness, a direct proof for Hawkes process is postponed to Section 5.1.

**Corollary 2.8.** Let  $\mu \in \mathbb{R}_+$ . We consider  $(H, \Lambda)$  the unique solution to SDE

$$\begin{cases} H_t = \int_{(0, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_s\}} N(ds, d\theta), & t \in [0, T] \\ \Lambda_t = \mu + \int_{(0, t)} \Phi(t-u) dH_u. \end{cases} \quad (2.3)$$

We set  $\mathbb{F}^H := (\mathcal{F}_t^H)_{t \in [0, T]}$  the natural filtration of  $H$  (obviously  $\mathcal{F}_t^H \subset \mathcal{F}_t^N$ ). Then  $H$  is a standard Hawkes process.

*Proof.* Let  $0 \leq s \leq t \leq T$  and  $A$  in  $\mathcal{F}_s^H$ . Using Lemma 2.3,  $\mathcal{Z}(u, \theta) = \mathbf{1}_{\{\theta \leq \Lambda_u\}}$  satisfies Relation (2.1) (as  $\Lambda$  is predictable), one can thus apply Proposition 2.4 (Mecke's Formula) for  $\mathcal{Z}(u, \theta) = \mathbf{1}_{\{\theta \leq \Lambda_u\}}$ . Then, we have (by conditioning with respect to  $\mathcal{F}_s^H$ )

$$\begin{aligned} & \mathbb{E}[1_{A_s}(H_t - H_s)] \\ &= \mathbb{E} \left[ 1_{A_s} \int_{(s, t]} dH_u \right] \\ &= \mathbb{E} \left[ 1_{A_s} \int_{(s, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_u\}} N(du, d\theta) \right] \\ &= \mathbb{E} \left[ 1_{A_s} \int_{(s, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_u\}} dud\theta \right] \\ &= \mathbb{E} \left[ 1_{A_s} \int_{(s, t]} \Lambda_u du \right], \end{aligned}$$

where we have used again Lemma 2.3 to prove that  $1_{A_s} \circ \varepsilon_{(u, \theta)}^+ = 1_{A_s}$  for  $u > s$ .  $\square$



### 3 An expansion formula for Hawkes processes

Thanks to the previous representation of a Hawkes process using the Poisson embedding, we derive in this section an expansion formula. The main result is stated in Theorem 3.13. It requires an accurate definition and analysis of what we named the shifted Hawkes processes resulting from the shifts on the Poisson space of the original Hawkes process.

#### 3.1 The shifted Hawkes processes

We now introduce shifted Hawkes processes that is the effect of the shift operators  $\varepsilon^+$  on the Hawkes process. As we will see the resulting Hawkes process can also be described in terms of a Poisson SDE.

**Definition 3.1** (One shift Hawkes process). Let  $H$  be a standard Hawkes process with initial intensity  $\mu > 0$  and bounded excitation function  $\Phi : [0, T] \rightarrow \mathbb{R}_+$  such that  $\|\Phi\|_1 < 1$ . Let  $v$  in  $(0, T)$ . We set

$$\begin{cases} H_t^v = \mathbf{1}_{[0,v]}(t)H_t + \mathbf{1}_{[v,T]}(t) \left( H_{v-}^v + 1 + \int_{(v,t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_s^v\}} N(ds, d\theta) \right), \\ \Lambda_t^v = \mathbf{1}_{(0,v]}(t)\Lambda_t + \mathbf{1}_{(v,T]}(t) \left( \mu^{v,1}(t) + \int_{(v,t)} \Phi(t-u) dH_u^v \right), \end{cases} \quad (3.1)$$

with  $\mu^{v,1}(t) := \mu + \int_{(0,v]} \Phi(t-u) dH_u^v = \mu + \int_{(0,v)} \Phi(t-u) dH_u + \Phi(t-v)$ .

**Remark 3.2.**  $(H^v, \Lambda^v)$  is called a shifted (at time  $v$ ) Hawkes process but it is not a Hawkes process as it possesses a jump at the deterministic time  $v$ . However, it is a Hawkes process on  $(0, v)$  (and coincides with the original Hawkes process  $(H, \Lambda)$ ) and is a Hawkes process in the sense of Definition 2.6 (generalized Hawkes process) on  $[v, T]$ .

**Definition 3.3.** Let  $v$  in  $[0, T]$ , we set

$$\begin{aligned} \varepsilon_{(v, \Lambda_v)}^+ : \Omega^N &\rightarrow \Omega^N \\ \omega &\mapsto (\varepsilon_{(v, \theta)}^+(\omega))_{\theta = \Lambda_v(\omega)}. \end{aligned}$$

**Remark 3.4.** Let  $v$  in  $[0, T]$  and  $\theta_0 \geq 0$ . As  $\mathbb{P}[N(\{v\} \times \mathbb{R}_+) > 0] = 0$ , a direct computation gives that

$$\begin{aligned} H_v \circ \varepsilon_{(v, \theta_0)}^+ &= \left( \int_{[0,v]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_s\}} N(ds, d\theta) \right) \circ \varepsilon_{(v, \theta_0)}^+ \\ &= H_{v-} + \mathbf{1}_{\{\theta_0 \leq \Lambda_v\}}, \quad \mathbb{P} - a.s.. \end{aligned}$$

Hence

$$\mathbf{1}_{\{\theta_0 \leq \Lambda_v\}} \left( H_v \circ \varepsilon_{(v, \theta_0)}^+ \right) = \mathbf{1}_{\{\theta_0 \leq \Lambda_v\}} \left( H_v \circ \varepsilon_{(v, \Lambda_v)}^+ \right), \quad \mathbb{P} - a.s.. \quad (3.2)$$

As Equation (2.2) is completely determined by the values  $h^v$  and  $(\mu^v(t))_{t \in [0, T]}$  (recalling that  $\Phi$  does not depend on  $v$ ), we deduce that for any  $\theta_0 \geq 0$ , on the set  $\{\theta_0 \leq \Lambda_v\}$ ,

$$(H_t \circ \varepsilon_{(v, \theta_0)}^+, \Lambda_t \circ \varepsilon_{(v, \theta_0)}^+)_{t \in [v, T]} = (H_t \circ \varepsilon_{(v, \Lambda_v)}^+, \Lambda_t \circ \varepsilon_{(v, \Lambda_v)}^+)_{t \in [v, T]}.$$

This remark leads us to the following lemma.

**Lemma 3.5.** *Let  $v$  in  $[0, T]$ . We have that*

$$(H \circ \varepsilon_{(v, \Lambda_v)}^+, \Lambda \circ \varepsilon_{(v, \Lambda_v)}^+) = (H^v, \Lambda^v),$$

where  $(H^v, \Lambda^v)$  is defined in (3.1).

*Proof.* Recall that

$$\begin{cases} H_t = H_v + \int_{(v, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda(s)\}} N(ds, d\theta), & t \in [0, T] \\ \Lambda(t) = \mu + \int_{(0, v]} \Phi(t - u) dH_u + \int_{(v, t]} \Phi(t - u) dH_u. \end{cases}$$

Thus, by Lemma 2.3,  $H_t \circ \varepsilon_{(v, \Lambda_v)}^+ = H_t$  for  $t < v$  and  $\Lambda_t \circ \varepsilon_{(v, \Lambda_v)}^+ = \Lambda_t$  for  $t \leq v$ . Let  $t \geq v$ , we have that (recall that the jump times of  $N$  shifted by  $\Lambda_v$  after time  $v$  coincide with those of  $N$ )

$$\begin{cases} H_t \circ \varepsilon_{(v, \Lambda_v)}^+ = H_{v-} + 1 + \int_{(v, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda \circ \varepsilon_{(v, \Lambda_v)}^+(s)\}} N(ds, d\theta), & t \in [0, T] \\ \Lambda_t \circ \varepsilon_{(v, \Lambda_v)}^+ = \mu + \int_{(0, v)} \Phi(t - u) dH_u + \Phi(t - v) + \int_{(v, t]} \Phi(t - u) dH_u \circ \varepsilon_{(v, \Lambda_v)}^+. \end{cases}$$

The result follows by uniqueness of the solution to this SDE.  $\square$

We now proceed iteratively to construct a Multi-shifted Hawkes process: the shifts  $(v_1, \dots, v_{n-1})$  being chosen, the  $n^{\text{th}}$  shift is taken in the interval  $]0, v_n[$ .

**Definition 3.6** (Multi-shifted Hawkes process).

Let  $n \in \mathbb{N}^*$ , and  $0 < v_n < v_{n-1} < \dots < v_1 < T$ . We set

$$\begin{cases} H_t^{v_n, \dots, v_1} = \mathbf{1}_{[0, v_n)}(t) H_t + \sum_{i=1}^n \mathbf{1}_{[v_i, v_{i-1})}(t) \left( H_{v_i-}^{v_n, \dots, v_1} + 1 + \int_{(v_i, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_s^{v_n, \dots, v_1}\}} N(ds, d\theta) \right), \\ \Lambda_t^{v_n, \dots, v_1} = \mathbf{1}_{(0, v_n]}(t) \Lambda_t + \sum_{i=1}^n \mathbf{1}_{(v_i, v_{i-1}]}(t) \left( \mu^{v_i, n}(t) + \int_{(v_i, t]} \Phi(t - u) dH_u^{v_n, \dots, v_1} \right), \end{cases} \quad (3.3)$$

with  $\mu^{v_i, n}(t) := \mu + \int_{(0, v_i]} \Phi(t - u) dH_u^{v_n, \dots, v_1} = \mu + \int_{(0, v_i)} \Phi(t - u) dH_u^{v_n, \dots, v_1} + \Phi(t - v_i)$ .

**Remark 3.7.** Note that the process  $H^{v_n, \dots, v_1}$  is not a Hawkes process as it has deterministic jumps at times  $v_n, \dots, v_1$  but it is a generalized Hawkes process on each interval  $(v_{i-1}, v_i)$ .

**Proposition 3.8.** *Let  $n \in \mathbb{N}^*$ , and  $0 < v_n < v_{n-1} < \dots < v_1 < T$ . We have that*

$$(H, \Lambda) \circ \varepsilon_{(v_1, \Lambda_{v_1})}^+ \circ \dots \circ \varepsilon_{(v_n, \Lambda_{v_n})}^+ = (H^{v_n, \dots, v_1}, \Lambda^{v_n, \dots, v_1}).$$

The proof is postponed to Section 5.2. We conclude this section on shifted Hawkes processes with the following remark.

**Remark 3.9.** Let  $v < v_n$  and  $\theta_0 \geq 0$ . Following the lines of Remark 3.4, we get that

$$\mathbf{1}_{\{\theta_0 \leq \Lambda_v\}} (H_t^{v_n, \dots, v_1} \circ \varepsilon_{(v, \theta_0)}^+, \Lambda_t^{v_n, \dots, v_1} \circ \varepsilon_{(v, \theta_0)}^+)_{t \in [v, T]} = \mathbf{1}_{\{\theta_0 \leq \Lambda_v\}} (H_t^{v_n, \dots, v_1} \circ \varepsilon_{(v, \Lambda_v)}^+, \Lambda_t^{v_n, \dots, v_1} \circ \varepsilon_{(v, \Lambda_v)}^+)_{t \in [v, T]}.$$

### 3.2 Expansion formula for the Hawkes process

Let  $H = (H_t)_{t \in [0, T]}$  be a standard Hawkes process with parameters  $\mu > 0$  and bounded non-negative kernel  $\Phi$  with  $\|\Phi\|_1 < 1$ , solution to the SDE (2.3) and build on the Poisson space  $(\Omega^N, \mathcal{F}_T^N, \mathbb{P}^N)$ , according to Section 2. In line of the results obtained in the previous section, we can derive the lemma below.

**Lemma 3.10.** *Let  $F$  be a  $\mathcal{F}_T^N$ -measurable random variable and  $v > 0$ . We have that*

$$\int_{\mathbb{R}_+} (F \circ \varepsilon_{(v, \theta)}^+) \mathbf{1}_{\{\theta \leq \Lambda_v\}} d\theta = \Lambda_v (F \circ \varepsilon_{(v, \Lambda_v)}^+), \quad \mathbb{P} - a.s..$$

*Proof.* Let  $\theta \geq 0$ . Remarks 3.5 and 3.9 entail

$$\mathbf{1}_{\{\theta \leq \Lambda_v\}} (F \circ \varepsilon_{(v, \theta)}^+) = \mathbf{1}_{\{\theta \leq \Lambda_v\}} (F \circ \varepsilon_{(v, \Lambda_v)}^+),$$

which concludes the proof.  $\square$

**Definition 3.11.** Let  $F$  be a  $\mathcal{F}_T^N$ -measurable random variable and  $Z$  a  $\mathbb{F}^N$ -predictable process. We set

(i) for  $v$  in  $(0, T)$ ,

$$F^v := F \circ \varepsilon_{(v, \Lambda_v)}^+, \quad Z^v := Z \circ \varepsilon_{(v, \Lambda_v)}^+;$$

(ii) for  $n \geq 2$  and  $0 < v_n < v_{n-1} < \dots < v_1 < T$ ,

$$F^{v_n, \dots, v_1} := F^{v_{n-1}, \dots, v_1} \circ \varepsilon_{(v_n, \Lambda_{v_n})}^+, \quad Z^{v_n, \dots, v_1} := Z^{v_{n-1}, \dots, v_1} \circ \varepsilon_{(v_n, \Lambda_{v_n})}^+. \quad (3.4)$$

**Notation 3.12.** For  $n = 1$ , we set  $m_\Phi(\Delta^1) := 1$  and for  $n \geq 2$ ,

$$m_\Phi(\Delta^n) := \int_0^T \int_0^{v_1} \dots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) dv_n \dots dv_1,$$

To ensure the convergence of the forthcoming expansion formula, one will need the following assumption

$$\lim_{n \rightarrow +\infty} m_\Phi(\Delta^n) = 0. \quad (3.5)$$

Remark that if  $\|\Phi\|_\infty < 1$ ,  $m_\Phi(\Delta^n) \leq \frac{T^n}{n!}$  and Relation (3.5) is satisfied.

We have now introduced all the ingredients to state the expansion formula for the Hawkes process.

**Theorem 3.13** (Expansion formula for the Hawkes process).

*Let  $F$  be a bounded  $\mathcal{F}_T^N$ -measurable random variable and  $Z = (Z_t)_{t \in [0, T]}$  be a bounded  $\mathbb{F}^H$ -predictable process. Then for any  $M \in \mathbb{N}^*$ ,*

$$\begin{aligned} & \mathbb{E} \left[ F \int_{[0, T]} Z_t dH_t \right] \\ &= \mu \int_0^T \mathbb{E} [Z_v F^v] dv \end{aligned}$$

$$\begin{aligned}
& + \mu \sum_{n=2}^M \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \mathbb{E} [Z_{v_1}^{v_n, \dots, v_2} F^{v_n, \dots, v_1}] dv_n \cdots dv_1 \\
& + \int_0^T \int_0^{v_1} \cdots \int_0^{v_M} \prod_{i=2}^{M+1} \Phi(v_{i-1} - v_i) \mathbb{E} [Z_{v_1}^{v_{M+1}, \dots, v_2} F^{v_{M+1}, \dots, v_1} \Lambda_{v_{M+1}}] dv_{M+1} \cdots dv_1. \quad (3.6)
\end{aligned}$$

In addition if Relation (3.5) is satisfied ( $\lim_{n \rightarrow +\infty} m_\Phi(\Delta^n) = 0$ ), we have that

$$\begin{aligned}
& \mathbb{E} \left[ F \int_{[0, T]} Z_t dH_t \right] \\
& = \mu \int_0^T \mathbb{E} [Z_v F^v] dv \\
& + \mu \sum_{n=2}^{+\infty} \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \mathbb{E} [Z_{v_1}^{v_n, \dots, v_2} F^{v_n, \dots, v_1}] dv_n \cdots dv_1. \quad (3.7)
\end{aligned}$$

**Remark 3.14.** Remark that the first term  $\mu \int_0^T \mathbb{E} [Z_v F^v] dv$  corresponds to the formula for a Poisson process (setting the self-exciting kernel  $\Phi$  at zero). Therefore the sum in the second term can be interpreted as a correcting term due to the self-exciting property of the counting process  $H$ . Besides, this formula can be used to provide lower and upper bounds on the premium of insurance contracts.

*Proof.* Set  $\mathcal{Z}(t, \theta) = Z_t \mathbf{1}_{\{\theta \leq \Lambda_t\}}$ . As  $Z$  is  $\mathbb{F}^H$ -predictable, it is  $\mathbb{F}^N$ -predictable and thus Relation (2.1) is satisfied. Thus Mecke's formula for Poisson functionals (see Proposition 2.4) gives that

$$\begin{aligned}
& \mathbb{E} \left[ F \int_{[0, T]} Z_t dH_t \right] \\
& = \mathbb{E} \left[ F \int_{[0, T]} \int_{\mathbb{R}_+} Z_t \mathbf{1}_{\{\theta \leq \Lambda_t\}} N(dt, d\theta) \right] \\
& = \mathbb{E} \left[ \int_{[0, T]} \int_{\mathbb{R}_+} Z_t (F \circ \varepsilon_{(t, \theta)}^+) \mathbf{1}_{\{\theta \leq \Lambda_t\}} dt d\theta \right] \\
& = \int_{[0, T]} \mathbb{E} \left[ Z_t \int_{\mathbb{R}_+} (F \circ \varepsilon_{(t, \theta)}^+) \mathbf{1}_{\{\theta \leq \Lambda_t\}} d\theta \right] dt \\
& = \int_{[0, T]} \mathbb{E} \left[ Z_t (F \circ \varepsilon_{(t, \Lambda_t)}^+) \Lambda_t \right] dt \\
& = \mu m_1 + I_1,
\end{aligned}$$

with

$$\begin{aligned}
m_1 & := \int_{[0, T]} \mathbb{E} \left[ Z_t (F \circ \varepsilon_{(t, \Lambda_t)}^+) \right] dt, \\
I_1 & := \int_{[0, T]} \mathbb{E} \left[ Z_v (F \circ \varepsilon_{(v, \Lambda_v)}^+) \int_{(0, v)} \Phi(v - u) dH_u \right] dv.
\end{aligned}$$

Setting  $F^v := F \circ \varepsilon_{(v, \Lambda_v)}^+$ , once again by Proposition 2.4 we have that

$$I_1 = \int_{[0, T]} \mathbb{E} \left[ Z_{v_1} F^{v_1} \int_{(0, v_1)} \Phi(v_1 - v_2) dH_{v_2} \right] dv_1$$

$$\begin{aligned}
&= \int_{[0,T]} \mathbb{E} \left[ Z_{v_1} F^{v_1} \int_{(0,v_1)} \Phi(v_1 - v_2) \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_{v_2}\}} N(dv_2, d\theta) \right] dv_1 \\
&= \int_{[0,T]} \int_0^{v_1} \mathbb{E} \left[ \Phi(v_1 - v_2) \int_{\mathbb{R}_+} Z_{v_1} \circ \varepsilon_{(v_2, \theta)}^+ F^{v_1} \circ \varepsilon_{(v_2, \theta)}^+ \mathbf{1}_{\{\theta \leq \Lambda_{v_2}\}} d\theta \right] dv_2 dv_1 \\
&= \int_{[0,T]} \int_0^{v_1} \Phi(v_1 - v_2) \mathbb{E} \left[ Z_{v_1} \circ \varepsilon_{(v_2, \Lambda_{v_2})}^+ F^{v_1} \circ \varepsilon_{(v_2, \Lambda_{v_2})}^+ \Lambda_{v_2} \right] dv_2 dv_1 \\
&= \mu m_2 + I_2,
\end{aligned}$$

with

$$m_2 := \int_{[0,T]} \int_0^{v_1} \Phi(v_1 - v_2) \mathbb{E} \left[ Z_{v_1} \circ \varepsilon_{(v_2, \Lambda_{v_2})}^+ F^{v_1} \circ \varepsilon_{(v_2, \Lambda_{v_2})}^+ \right] dv_2 dv_1,$$

and

$$I_2 := \int_{[0,T]} \int_0^{v_1} \Phi(v_1 - v_2) \mathbb{E} \left[ Z_{v_1} \circ \varepsilon_{(v_2, \Lambda_{v_2})}^+ F^{v_1} \circ \varepsilon_{(v_2, \Lambda_{v_2})}^+ \int_{(0,v_2]} \Phi(v_3 - v_2) dH_{v_3} \right] dv_2 dv_1.$$

For  $n \geq 2$  we set

$$m_n := \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \mathbb{E} \left[ Z_{v_1}^{v_n, \dots, v_2} F^{v_n, \dots, v_1} \right] dv_n \cdots dv_1,$$

and

$$I_n := \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \mathbb{E} \left[ Z_{v_1}^{v_n, \dots, v_2} F^{v_n, \dots, v_1} \int_{(0, v_n]} \Phi(v_{n+1} - v_n) dH_{v_{n+1}} \right] dv_n \cdots dv_1.$$

We have that

$$\begin{aligned}
I_n &= \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \mathbb{E} \left[ Z_{v_1}^{v_n, \dots, v_2} F^{v_n, \dots, v_1} \int_{(0, v_n]} \Phi(v_{n+1} - v_n) \mathbf{1}_{\{\theta \leq \Lambda_{v_{n+1}}\}} N(dv_{n+1}, d\theta) \right] dv_n \cdots dv_1 \\
&= \int_0^T \int_0^{v_1} \cdots \int_0^{v_n} \prod_{i=2}^{n+1} \Phi(v_{i-1} - v_i) \mathbb{E} \left[ \int_{\mathbb{R}_+} (Z_{v_1}^{v_n, \dots, v_2} F^{v_n, \dots, v_1}) \circ \varepsilon_{(v_{n+1}, \theta)}^+ \mathbf{1}_{\{\theta \leq \Lambda_{v_{n+1}}\}} d\theta \right] dv_{n+1} \cdots dv_1 \\
&= \int_0^T \int_0^{v_1} \cdots \int_0^{v_n} \prod_{i=2}^{n+1} \Phi(v_{i-1} - v_i) \mathbb{E} \left[ (Z_{v_1}^{v_n, \dots, v_2} F^{v_n, \dots, v_1}) \circ \varepsilon_{(v_{n+1}, \Lambda_{v_{n+1}})}^+ \Lambda_{v_{n+1}} \right] dv_{n+1} \cdots dv_1 \\
&= \int_0^T \int_0^{v_1} \cdots \int_0^{v_n} \prod_{i=2}^{n+1} \Phi(v_{i-1} - v_i) \mathbb{E} \left[ Z_{v_1}^{v_{n+1}, \dots, v_2} F^{v_{n+1}, \dots, v_1} \Lambda_{v_{n+1}} \right] dv_{n+1} \cdots dv_1 \\
&= \mu m_{n+1} + I_{n+1}.
\end{aligned}$$

Hence, by induction, we obtain the first part of Theorem 3.13 (Equation (3.6)).

Classical estimates (using the expression of  $\Lambda$  in Definition 2.5 and [?, Lemma 3]) yield that  $\mathbb{E}[\Lambda_t] \leq \mu \left( 1 + \frac{\|\Phi\|_1}{1 - \|\Phi\|_1} \right)$ , for any  $t$ . Therefore, as  $Z$  and  $F$  are assumed to be bounded, there exists  $C > 0$  (depending on  $\|Z\|_\infty$ ,  $\|F\|_\infty$ ,  $\mu$  and  $\|\Phi\|_1$ ) such that

$$\int_0^T \int_0^{v_1} \cdots \int_0^{v_M} \prod_{i=2}^{M+1} \Phi(v_{i-1} - v_i) \mathbb{E} \left[ Z_{v_1}^{v_{M+1}, \dots, v_2} F^{v_{M+1}, \dots, v_1} \Lambda_{v_{M+1}} \right] dv_{M+1} \cdots dv_1$$

$$\leq C \int_0^T \int_0^{v_1} \cdots \int_0^{v_M} \prod_{i=2}^{M+1} \Phi(v_{i-1} - v_i) dv_{M+1} \cdots dv_1 = C m_\Phi(\Delta^{M+1})$$

which converges to zero as  $M \rightarrow +\infty$  due to Relation (3.5).  $\square$

An alternative form can be given for the representation (3.7). Let  $n \geq 1$  and  $\Delta^n$  the  $n$ -dimensional simplex of  $[0, T]$  :

$$\Delta^n := \{0 < v_n < \cdots < v_1 < T\}.$$

Let  $\mathcal{U}^n$  be a flat Dirichlet distribution on  $\Delta^n$  that is a continuous random variable uniformly distributed on  $\Delta^n$ , then for any Borelian integrable map  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\mathbb{E}[\varphi(\mathcal{U}^n)] = \frac{n!}{T^n} \int_0^T \cdots \int_0^{v_{n-1}} \varphi(v_n, \dots, v_1) dv_n \cdots dv_1.$$

**Remark 3.15.** Let  $(\mathcal{U}^n)_{n \geq 2}$  be a sequence of independent random variables such that  $\mathcal{U}^i := (\mathcal{U}_1^i, \dots, \mathcal{U}_i^i)$  is a Dirichlet distribution on  $\Delta^i$ . Then

$$\begin{aligned} \mathbb{E} \left[ F \int_{[0, T]} Z_t dH_t \right] &= \mu \int_0^T \mathbb{E} [Z_v F^v] dv \\ &+ \mu \sum_{n=2}^{+\infty} \frac{T^n}{n!} \mathbb{E} \left[ \prod_{i=2}^n \Phi(\mathcal{U}_{i-1}^n - \mathcal{U}_i^n) Z_{\mathcal{U}_1^n}^{\mathcal{U}_n^n, \dots, \mathcal{U}_2^n} F^{\mathcal{U}_n^n, \dots, \mathcal{U}_1^n} \right]. \end{aligned} \quad (3.8)$$

## 4 Insurance derivatives for cyber risk

The expansion formula can be applied to compute closed formula for the premium of a class of insurance and financial derivatives (such as reinsurance contracts including generalized Stop-Loss contracts, or CDO tranches) or risk management instruments (like Expected Shortfall), to provide generalizations of results that have been proved in [11] in a Cox model setting.

In this section, we choose to focus on cyber reinsurance contracts. Indeed, one important feature of cyber risk is the presence of accumulation phenomena and contagion, than can not be accurately modeled by Poisson processes (see the statistical analysis of Bessy et al. [?], based on the public Privacy Rights Clearinghouse database). This means that assuming a Poisson process to model the frequency of cyber risk may induce an underestimation of the risk, due to a misspecification of the dependency and self-exciting components of the risk. Developing a valuation formula for cyber contracts, taking into account this self-exciting feature is thus a crucial challenge for cyber insurance.

### 4.1 The cumulative loss processes and derivatives payoffs

The so-called cumulative loss process is a key process for risk analysis in insurance and reinsurance. It corresponds to the cumulative sum of claims amounts, the sum being indexed by a counting process modeling the arrival times of the claims. In standard models, the counting process is assumed to be a Poisson process, that is the claims inter-arrivals are assumed iid with exponential distribution. Nevertheless, for cyber risk, a Hawkes process modeling is more appropriate, due to the auto-excitation feature of cyber risk (see [?], [?] or [?]). In the following, we propose closed-form formula for insurance derivatives, such as stop-loss contracts, for

which the cumulative loss process is indexed by a Hawkes process  $L_t := \sum_{i=1}^{H_t} X_i$  where  $(H_t)$  is a Hawkes process and  $X_i$  models the  $i^{\text{th}}$ -claim amount, that arrives at the random time  $\tau_i^H$ .

More precisely, the mathematical framework is the following. Let  $(\Omega^N, \mathcal{F}^N, \mathbb{P}^N)$  the probability space as defined previously. We consider  $(\Omega^C, \mathcal{F}^C, \mathbb{P}^C)$  a probability space on which we define  $(\eta_1, \vartheta_1)$  a  $\mathbb{R}_+^2$ -valued random variable (we denote by  $\mu$  its distribution) and  $(\eta_i, \vartheta_i)_{i \geq 2}$  and  $(\bar{\eta}_i, \bar{\vartheta}_i)_{i \geq 1}$  that are independent copies of  $(\eta_1, \vartheta_1)$ . We set

$$\Omega := \Omega^N \times \Omega^C, \quad \mathbb{F} := (\mathcal{F}_t)_{t \in [0, T]}, \quad \mathcal{F}_t := \mathcal{F}_t^H \otimes \mathcal{F}_t^C, \quad \mathbb{P} := \mathbb{P}^N \otimes \mathbb{P}^C.$$

Note that variables (or processes) defined only on  $\Omega^N$  (respectively  $\Omega^C$ ) naturally extend to  $\Omega$ . In addition,  $H$  and the underlying Poisson process  $N$  are independent of the variables  $\eta_i, \bar{\eta}_i, \vartheta_i, \bar{\vartheta}_i$  ( $i \geq 1$ ).

We now define the cumulative loss processes and the insurance derivatives we consider. One typical example of insurance contract is Stop-Loss contract. A Stop-Loss contract provides to its buyer protection against losses which are larger than a given threshold  $\underline{K}$ . If the  $i^{\text{th}}$ -claim size is  $f(\eta_i)$ , then the cumulative loss process is given by  $L_t := \sum_{i=1}^{H_t} e^{-\kappa(t-\tau_i^H)} f(\eta_i)$ , where  $e^{-\kappa(t-\tau_i^H)}$  is a discount factor. The process  $(L_t)$  is the loss that activates the contract. Sometimes the compensation amount are not exactly the ones that are computed to activate the reinsurance contract and may depend on other losses  $\vartheta_i$ . The "generalized" loss process is then given by  $K_t := \sum_{i=1}^{H_t} e^{-\kappa(t-\tau_i^H)} g(\eta_i, \vartheta_i)$ . For example, for a stop loss contract, the reinsurance company pays the loss amount above a threshold  $\underline{K}$ , up to a given amount  $(\bar{K} - \underline{K})$  fixed by the contract. The payoff of a generalized stop loss contract is then given by

$$\begin{cases} 0, & \text{if } L_T \leq \underline{K} \\ K_T - \underline{K}, & \text{if } \underline{K} \leq L_T \leq \bar{K} \\ \bar{K} - \underline{K}, & \text{if } L_T \geq \bar{K}. \end{cases} \quad (4.1)$$

Then, the premium<sup>3</sup> of such a contract is

$$\mathbb{E} [K_T \mathbf{1}_{\{L_T > \underline{K}\}}] - \underline{K} \mathbb{P} [L_T \in [\underline{K}, \bar{K}]] + (\bar{K} - \underline{K}) \mathbb{P} [L_T \geq \bar{K}]. \quad (4.2)$$

Therefore one has to compute quantities of the form  $\mathbb{E} [K_T h(L_T)]$ . The definitions are gathered below.

**Definition 4.1** (Insurance derivative contract).<sup>4</sup>

We denote by  $(\tau_i^H)_{i \in \mathbb{N}}$  the jump times of the counting process  $H$ .

- (i) Given  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  a bounded deterministic function, the loss process is

$$L_t := \sum_{i=1}^{H_t} e^{-\kappa(t-\tau_i^H)} f(\eta_i), \quad t \in [0, T]. \quad (4.3)$$

<sup>3</sup>Remark that when the risk of the contract is neither hedgeable nor even related to a financial market, the premium of the contract relies on the computation of the expectation of the payoff, under the physical probability measure  $\mathbb{P}$ .

<sup>4</sup>We put strong a condition of boundedness on the functions  $f$ ,  $g$  and  $h$  since this condition is in force for actuarial derivatives, but it can be relaxed to a weaker integrability condition

(ii) Given  $g : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$  a bounded deterministic function, the generalized loss process is

$$K_t := \sum_{i=1}^{H_t} e^{-\kappa(t-\tau_i^H)} g(\eta_i, \vartheta_i), \quad t \in [0, T]. \quad (4.4)$$

(iii) Let  $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  be a bounded deterministic function. We aim at computing the expectation of the derivatives payoff, that is the quantity

$$\mathbb{E} [K_T h(L_T)]. \quad (4.5)$$

## 4.2 General pricing formula

For the computation of  $\mathbb{E} [K_T h(L_T)]$ , we will rely on the formula in Theorem 3.13, by writing the cumulative generalized loss process  $K_t$  as an integral with respect to the Hawkes process. Namely,

$$K_T = \int_{(0, T]} Z_s dH_s, \quad t \in [0, T]$$

with

$$Z_s := \sum_{i=1}^{+\infty} g(\eta_i, \vartheta_i) e^{-\kappa(T-s)} \mathbf{1}_{(\tau_{i-1}^H, \tau_i^H]}(s), \quad s \in [0, T].$$

We first introduce the following definition.

**Definition 4.2.** Let  $n \in \mathbb{N}^*$  and  $0 < v_n < \dots < v_1 < T$ .

We set  $\tau_i^{v_n, \dots, v_1}$  the  $i^{\text{th}}$ -jump time of the shifted Hawkes process  $H_T^{v_n, \dots, v_1}$  and

$$L_T^{v_n, \dots, v_1} := \sum_{i=1}^{H_T^{v_n, \dots, v_1}} e^{-\kappa(T-\tau_i^{v_n, \dots, v_1})} f(\eta_i).$$

**Theorem 4.3.** Under the previous assumptions and Relation (3.5), it holds that

$$\begin{aligned} & \mathbb{E}[K_T h(L_T)] \\ &= \mu \int_0^T e^{-\kappa(T-v_1)} \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( e^{-\kappa(T-v_1)} f(\bar{\eta}_1) + \sum_{\substack{i=1 \\ \tau_i^{v_1} \neq v_1}}^{(H_T^{v_1})-1} e^{-\kappa(T-\tau_i^{v_1})} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right] dv_1 \\ &+ \mu \sum_{n=2}^{+\infty} \int_0^T \int_0^{v_1} \dots \int_0^{v_{n-1}} e^{-\kappa(T-v_1)} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \\ & \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{\substack{i=1 \\ \tau_i^{v_n, \dots, v_1} \neq v_n, \dots, v_1}}^{(H_T^{v_n, \dots, v_1})-n}} e^{-\kappa(T-\tau_i^{v_n, \dots, v_1})} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right] dv_n \dots dv_1. \end{aligned} \quad (4.6)$$

where  $(\bar{\eta}_i, \bar{\vartheta}_i)_{i \geq 1}$  are independent copies of  $(\eta_1, \vartheta_1)$ , independent of all other variables.



*Proof.* For all  $t \in [0, T]$ ,  $K_t = \int_{(0,t]} Z_s dH_s$  with  $Z_s := \sum_{i=1}^{+\infty} e^{-\kappa(T-s)} g(\eta_i, \vartheta_i) \mathbf{1}_{(\tau_{i-1}^H, \tau_i^H]}(s)$ . Note that  $Z_s = e^{-\kappa(T-s)} g(\eta_{1+H_{s-}}, \vartheta_{1+H_{s-}})$  for any  $s$  in  $(0, T]$ . Recall that  $L_T = \sum_{i=1}^{H_T} e^{-\kappa(T-\tau_i)} f(\eta_i)$ . By Theorem 3.13, we have that

$$\begin{aligned} & \mathbb{E}[K_T h(L_T)] \\ &= \mathbb{E} \left[ h(L_T) \int_{(0,T]} Z_t dH_t \right] \\ &= \mathbb{E} \left[ h(L_T) \int_{(0,T]} e^{-\kappa(T-t)} g(\eta_{1+H_{t-}}, \vartheta_{1+H_{t-}}) dH_t \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ h(L_T) \int_{(0,T]} Z_t dH_t \middle| \mathcal{F}^C \right] \right] \\ &= \mu \int_0^T \mathbb{E} \left[ e^{-\kappa(T-v_1)} g(\eta_{1+H_{v_1-}}, \vartheta_{1+H_{v_1-}}) h(L_T^{v_1}) \right] dv_1 + \mu \sum_{n=2}^{+\infty} I_n \end{aligned}$$

with for  $n \geq 2$

$$I_n := \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \mathbb{E} \left[ e^{-\kappa(T-v_1)} g(\eta_{1+H_{v_1-}^{v_n, \dots, v_2}}, \vartheta_{1+H_{v_1-}^{v_n, \dots, v_2}}) h(L_T^{v_n, \dots, v_1}) \right] dv_n \cdots dv_1$$

and for  $n = 1$

$$I_1 := \int_0^T \mathbb{E} \left[ e^{-\kappa(T-v_1)} g(\eta_{1+H_{v_1-}}, \vartheta_{1+H_{v_1-}}) h(L_T^{v_1}) \right] dv_1.$$

By definition,  $1 + H_{v_1-}^{v_n, \dots, v_2} = H_{v_1-}^{v_n, \dots, v_1}$ ,  $\mathbb{P}$ -a.s.. Using the fact that the  $(\eta_i)$  are iid, we can separate the claim at time  $v_1$  (that we will represent using the random variables  $(\bar{\eta}_1, \bar{\vartheta}_1)$ ) from the other claims. Thus

$$\begin{aligned} & I_n \\ &= \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \\ & \mathbb{E} \left[ e^{-\kappa(T-v_1)} g(\bar{\eta}_1, \bar{\vartheta}_1) h \left( e^{-\kappa(T-v_1)} f(\bar{\eta}_1) + \sum_{\substack{i=1 \\ \tau_i^{v_n, \dots, v_1} \neq v_1}}^{H_T^{v_n, \dots, v_1} - 1}} e^{-\kappa(T-\tau_i^{v_n, \dots, v_1})} f(\eta_i) \right) \right] dv_n \cdots dv_1 \\ &= \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \\ & \mathbb{E} \left[ e^{-\kappa(T-v_1)} g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( e^{-\kappa(T-v_1)} f(\bar{\eta}_1) + \sum_{\substack{i=1 \\ \tau_i^{v_n, \dots, v_1} \neq v_1}}^{H_T^{v_n, \dots, v_1} - 1}} e^{-\kappa(T-\tau_i^{v_n, \dots, v_1})} f(\eta_i) \right) \middle| (\bar{\eta}_1, \bar{\vartheta}_1) \right] \right] dv_n \cdots dv_1 \\ &= \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \end{aligned}$$

$$\mathbb{E} \left[ e^{-\kappa(T-v_1)} g(\bar{\eta}_1, \bar{v}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{\substack{i=1 \\ \tau_i^{v_n, \dots, v_1} \neq v_n, \dots, v_1}}^{(H_T^{v_n, \dots, v_1})-n} e^{-\kappa(T-\tau_i^{v_n, \dots, v_1})} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right] dv_n \cdots dv_1.$$

The treatment for  $I_1$  is similar, which concludes the proof.  $\square$

The expansion formula (4.6) is written in terms of the shifted Hawkes process  $H^{v_n, \dots, v_1}$ . The shifted Hawkes process has three "types" of jumps

- the spontaneous jumps induced by an homogeneous Poisson with intensity  $\mu$
- the deterministic enforced jumps at time  $v_n < \dots < v_1$
- the auto-excited jumps that are induced by previous jumps of the process

Controlling this different types of jumps allows us to perform bounds on the premium.

### 4.3 Lower and upper bounds

We now use Relation (4.6) to perform lower and upper bounds on the premium.

Recall that the premium takes the form  $\mathbb{E}[K_T h(L_T)]$  with  $K_T = \sum_{i=1}^{H_T} e^{-\kappa(T-\tau_i^H)} X_i$ .

In the more simple case without discounting ( $\kappa = 0$ ), computing lower or upper bounds on the premium relies on two types of estimates :

- 1) estimates on the CDF of a given sum of claims  $\sum_{i=1}^n X_i$  (where  $n$  is prescribed)
- 2) estimates on the CDF of the counting process  $H$ .

Usually, for insurance derivatives a specific model of claims is considered so that the former quantity  $\sum_{i=1}^n X_i$  is accessible. The main issue of course lies in the estimates for the value of  $H_T$ . For instance, one can make the following type of estimates on the Hawkes process :

- (i) Obtain an upper bounds on  $\mathbb{P}[H_T \geq C]$  for some constant  $C$  (by Markov's inequality for instance), but this is only an upper bound.
- (ii) Get a lower bound on the premium by noting that  $H \geq \tilde{N}$  where  $\tilde{N}$  is an homogeneous Poisson process with intensity  $\mu$  (as  $\Lambda_t \geq \mu$  for any  $t \geq 0$ ).

In both approaches one makes rough estimates on the Hawkes process. Our approach allows for an intermediary situation as the processes  $H^{v_n, \dots, v_1}$  have deterministic jumps at times  $v_n, \dots, v_1$  which are weighted by the kernels  $\Phi(v_{i-1} - v_i)$  over the simplex. So we can make less stringent estimates by at least knowing  $n$  jumps of the (shifted) Hawkes process. Obviously, on each intervals  $[v_{i-1}, v_i]$  we have to consider a Hawkes process for which estimates of the form (i)-(ii) above are the best tools available.

#### 4.3.1 Lower bound

Using Theorem 4.3, one can already obtain a first lower bound by just considering the  $n$  enforced jumps of the shifted Hawkes process and ignoring the other jumps (which is clearly a very rough estimate).

**Proposition 4.4.** *Assume  $h$  is non-decreasing. We have that*

$$\begin{aligned} \mathbb{E}[K_T h(L_T)] &\geq \mu \int_0^T e^{-\kappa(T-v_1)} \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) h \left( e^{-\kappa(T-v_1)} f(\bar{\eta}_1) \right) \right] dv_1 \\ &\quad + \mu \sum_{n=2}^{+\infty} \int_0^T \int_0^{v_1} \cdots \int_0^{v_{n-1}} e^{-\kappa(T-v_1)} \\ &\quad \prod_{i=2}^n \Phi(v_{i-1} - v_i) \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) \right) \middle| \bar{\eta}_1 \right] \right] dv_n \cdots dv_1. \end{aligned}$$

If  $\kappa = 0$ , the lower bound simplifies as

$$\mathbb{E}[K_T h(L_T)] \geq \mu \sum_{n=1}^{+\infty} \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n f(\bar{\eta}_k) \right) \middle| \bar{\eta}_1 \right] \right] m_\Phi(\Delta^n). \quad (4.7)$$

*Proof.* We apply Relation (4.6) and note that for any  $n \geq 1$ , we have simply used the fact that  $H_T^{v_n, \dots, v_1} \geq n$ .  $\square$

The formula above takes only into account the deterministic jumps added to the process. The proposition below is more accurate by including the jumps of a homogeneous Poisson process with constant intensity  $\mu$ .

**Proposition 4.5.** *Assume  $h$  is non-decreasing. We consider a family  $(\mathcal{U}^p)_{p \geq 1}$  of independent random variables (which are constructed from  $N$  only), where for each  $p$ ,  $\mathcal{U}^p$  is a flat Dirichlet distributions on  $\Delta^p$ . Set for  $n \geq 1$ ,  $0 < v_n < \cdots < v_1 < T$  :*

$$\begin{aligned} &\alpha_n(v_n, \dots, v_1) \\ &:= \sum_{p=0}^{+\infty} e^{-(T\mu)} \frac{(T\mu)^p}{p!} \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{i=1}^p e^{-\kappa(T-\mathcal{U}_i^p)} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right]. \end{aligned}$$

It holds that

$$\begin{aligned} \mathbb{E}[K_T h(L_T)] &\geq \mu \int_0^T e^{-\kappa(T-v_1)} \alpha_1(v_1) dv_1 \\ &\quad + \mu \sum_{n=2}^{+\infty} \int_0^T \cdots \int_0^{v_{n-1}} e^{-\kappa(T-v_1)} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \alpha_n(v_n, \dots, v_1) dv_n \cdots dv_1. \quad (4.8) \end{aligned}$$

*Proof.* Let  $n \geq 1$  and  $0 < v_n < \cdots < v_1 < T$ . By Proposition 3.8,  $H_T^{v_n, \dots, v_1} \geq n + H_T$ ,  $\mathbb{P}$ -a.s.. In addition by Lemma 5.1 (see Section 5.3), we have that  $H_T \geq \tilde{N}_T$ , with  $\tilde{N}_t := N([0, t] \times [0, \mu])$  which is an homogeneous Poisson process with intensity  $\mu$  and by construction, any jump  $\tau_i^{\tilde{N}}$  of  $\tilde{N}$  different of  $v_n, \dots, v_1$  is a jump of  $H_T^{v_n, \dots, v_1}$  (different of  $v_n, \dots, v_1$ ). Hence, using Relation (4.6), we have that

$$\begin{aligned} &\mathbb{E}[K_T h(L_T)] \\ &\geq \mu \int_0^T e^{-\kappa(T-v_1)} \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( e^{-\kappa(T-v_1)} f(\bar{\eta}_1) + \sum_{\substack{i=1 \\ \tau_i^{\tilde{N}} \neq v_1}}^{\tilde{N}_T} e^{-\kappa(T-\tau_i^{\tilde{N}})} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right] dv_1 \end{aligned}$$

$$\begin{aligned}
& + \mu \sum_{n=2}^{+\infty} \int_0^T \cdots \int_0^{v_{n-1}} e^{-\kappa(T-v_1)} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \\
& \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\nu}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{\substack{\tilde{N}_T \\ \tau_i^{\tilde{N}} \neq v_n, \dots, v_1}} e^{-\kappa(T-\tau_i^{\tilde{N}})} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right] dv_n \cdots dv_1.
\end{aligned}$$

For  $n \geq 1$  and  $0 < v_n < \cdots < v_1 < T$ , it holds that :

$$\begin{aligned}
& \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{\substack{\tilde{N}_T \\ \tau_i^{\tilde{N}} \neq v_n, \dots, v_1}} e^{-\kappa(T-\tau_i^{\tilde{N}})} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \\
& = \sum_{p=0}^{+\infty} e^{-(T\mu)} \frac{(T\mu)^p}{p!} \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{\substack{p \\ \tau_i^{\tilde{N}} \neq v_n, \dots, v_1}} e^{-\kappa(T-\tau_i^{\tilde{N}})} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \\
& = \sum_{p=1}^{+\infty} e^{-(T\mu)} \frac{(T\mu)^p}{p!} \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{i=1}^p e^{-\kappa(T-\mathcal{U}_i^p)} f(\eta_i) \right) \middle| \bar{\eta}_1 \right],
\end{aligned}$$

which concludes the proof.  $\square$

**Remark 4.6.** A direct approach for exhibiting  $p$  jumps of the Hawkes process would require conditioning on  $H_T \geq p$  and then to obtain a lower bound for  $\mathbb{P}[H_T \geq p]$  by for example  $\mathbb{P}[\tilde{N}_T \geq p]$  (where  $\tilde{N}$  is the homogeneous Poisson process obtained from  $N$  with intensity  $\mu$ ) which is what we do. But the advantage of our approach lies in the fact that on top of these jumps of  $\tilde{N}$  we can go further and consider  $n$  jumps of the Hawkes process which somehow are really produced by the self-excitation phenomenon as these jumps are weighted by the kernel  $\Phi$  along the jump times  $v_n, \dots, v_1$ .

As a corollary to Proposition 4.5, we can deduce the following lower bound in case the discounting factor  $\kappa$  is equal to 0.

**Corollary 4.7.** *Assume  $h$  is non-decreasing and  $\kappa = 0$ . It holds that*

$$\mathbb{E}[K_T h(L_T)] \geq \mu \sum_{n=1}^{+\infty} m_{\Phi}(\Delta^n) \sum_{p=0}^{+\infty} e^{-(T\mu)} \frac{(T\mu)^p}{p!} \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\nu}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n f(\bar{\eta}_k) + \sum_{i=1}^p f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right], \quad (4.9)$$

where we recall that  $m_{\Phi}(\Delta^n)$  is defined in Notation 3.12.

As pointed out in Remark 3.14, the first term in the sum indexed by  $n$  (that is for  $n = 1$ ) corresponds to the formula for a Poisson process with intensity  $\mu$ . Therefore the sum of the remaining terms (for  $n \geq 2$ ) corresponds to a lower bound for the correcting term due to the self-exciting property of the counting process  $H$ . This quantity should be at least added to a computation of the premium based on a standard Poisson process model.

### Application for Stop Loss contract with deductible.

We consider the following setting of Stop Loss contract ( $h(x) = \mathbf{1}_{x \geq \underline{K}}$ ) with no discounting ( $\kappa = 0$ ), and with a deductible on the reporting of the claims such that only claims whose amount exceeds a threshold  $\underline{f}$  are reported. A lower bound for the surplus of premium due to the self-exciting property of the counting process  $H$  is

$$\mu \mathbb{E} [g(\bar{\eta}_1, \bar{\vartheta}_1)] \sum_{n=2}^{+\infty} m_{\Phi}(\Delta^n) \sum_{p+n \geq \lfloor \underline{K}/\underline{f} \rfloor + 1}^{+\infty} e^{-(T\mu)} \frac{(T\mu)^p}{p!}$$

where  $\mu$  is the intensity of the "spontaneous" jumps (induced an homogeneous Poisson with intensity  $\mu$ ) and  $\mathbb{E} [g(\bar{\eta}_1, \bar{\vartheta}_1)]$  is the mean cost of one claim. If we assume furthermore a decreasing excitation kernel  $\Phi$ , then  $m_{\Phi}(\Delta^n) \geq \Phi(T)^{n-1} \frac{T^n}{n!}$  and the lower bound becomes

$$\mu \mathbb{E} [g(\bar{\eta}_1, \bar{\vartheta}_1)] \sum_{n=2}^{+\infty} \Phi(T)^{n-1} \frac{T^n}{n!} \sum_{p+n \geq \lfloor \underline{K}/\underline{f} \rfloor + 1}^{+\infty} e^{-(T\mu)} \frac{(T\mu)^p}{p!}.$$

#### 4.3.2 Upper bound

We now turn to the upper bound. We introduce the following quantities, as in [7].

**Proposition 4.8.** *Assume  $h$  is non-decreasing and  $\Phi$  is non-increasing. Let  $C_2 := \left( \int_0^T \Psi_1(t) dt \right)^2$  with  $\Psi_1$  solution to*

$$\Psi_1(t) = 1 + \int_0^t \Phi(t-s) \Psi_1(s) ds, \quad t \in [0, T]$$

and  $C_1 := \int_0^T \Psi_2(t) dt$ , with  $\Psi_2$  solution to

$$\Psi_2(t) = (\Psi_1(t))^2 + \int_0^t \Phi(s) \Psi_2(t-s) ds, \quad t \in [0, T].$$

For  $n \geq 1$ , let also

$$c_n := (\mu + n\Phi(0))C_1 + (\mu + n\Phi(0))^2 C_2.$$

Finally set for  $n \geq 1$  and  $0 < v_n < \dots < v_1 < T$

$$\begin{aligned} & \beta_n(v_n, \dots, v_1) \\ & := e^{-T(\mu+n\Phi(0))} \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) \right) \middle| \bar{\eta}_1 \right] \right] \\ & + \sum_{p=1}^{+\infty} \left( \frac{c_n}{p^2} \wedge 1 \right) \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{i=1}^p f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right] \end{aligned}$$

We have that

$$\begin{aligned} \mathbb{E}[K_T h(L_T)] & \leq \mu \int_0^T e^{-\kappa(T-v_1)} \beta_1(v_1) dv_1 \\ & + \mu \sum_{n=2}^{+\infty} \int_0^T \dots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) e^{-\kappa(T-v_1)} \beta_n(v_n, \dots, v_1) dv_n \dots dv_1. \end{aligned} \quad (4.10)$$

*Proof.* Fix  $n \geq 1$ . By construction (see Lemma 5.2),  $H_T^{v_n, \dots, v_1} \leq H_T^{\mu+n\Phi(0)} + n$ , where  $H^{\mu+n\Phi(0)}$  denotes the Hawkes process (constructed from  $N$ ) with initial intensity  $\mu$  replaced by  $\mu+n\Phi(0)$  (in other words,  $H^{\mu+n\Phi(0)}$  is solution to Equation (2.3) with initial intensity  $\mu+n\Phi(0)$ ). Denote by  $(\tau_i^n)_i$  the jump times of  $H_T^{\mu+n\Phi(0)}$ ; by the thinning procedure, the jump times  $(\tau_i^H)_i$  of  $H$  are necessarily included in the jump times  $(\tau_i^n)_i$  of  $H^{\mu+n\Phi(0)}$ . By Lemma 5.2, we can make the estimate below. We have that

$$\begin{aligned}
& \mathbb{E}[K_T h(L_T)] \\
& \leq \mu \int_0^T \mathbb{E} \left[ e^{-\kappa(T-v_1)} g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( e^{-\kappa(T-v_1)} f(\bar{\eta}_1) + \sum_{i=1}^{H_T^{\mu+n\Phi(0)}} e^{-\kappa(T-\tau_i^n)} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right] dv_1 \\
& + \mu \sum_{n=2}^{+\infty} \int_0^T \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \\
& \mathbb{E} \left[ e^{-\kappa(T-v_1)} g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{i=1}^{H_T^{\mu+n\Phi(0)}} e^{-\kappa(T-\tau_i^n)} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right] dv_n \cdots dv_1 \\
& \leq \mu \int_0^T \mathbb{E} \left[ e^{-\kappa(T-v_1)} g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( e^{-\kappa(T-v_1)} f(\bar{\eta}_1) + \sum_{i=1}^{H_T^{\mu+n\Phi(0)}} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right] dv_1 \\
& + \mu \sum_{n=2}^{+\infty} \int_0^T \cdots \int_0^{v_{n-1}} \prod_{i=2}^n \Phi(v_{i-1} - v_i) \\
& \mathbb{E} \left[ e^{-\kappa(T-v_1)} g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{i=1}^{H_T^{\mu+n\Phi(0)}} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right] dv_n \cdots dv_1
\end{aligned}$$

where we used the upper bound  $e^{-\kappa(T-\tau_i^n)} \leq 1$  to get rid off the unknown jump times  $\tau_i^n$ . Let  $n \geq 1$ ,  $0 < v_n < \cdots < v_1 < T$ . We have that

$$\begin{aligned}
& \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{i=1}^{H_T^{\mu+n\Phi(0)}} f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \\
& = \sum_{p=0}^{+\infty} \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{i=1}^p f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \mathbb{P}[H_T^{\mu+n\Phi(0)} = p] \\
& \leq \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) \right) \middle| \bar{\eta}_1 \right] \mathbb{P}[H_T^{\mu+n\Phi(0)} = 0] \\
& + \sum_{p=1}^{+\infty} \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{i=1}^p f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \mathbb{P}[H_T^{\mu+n\Phi(0)} \geq p] \\
& \leq \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) \right) \middle| \bar{\eta}_1 \right] \mathbb{P}[H_T^{\mu+n\Phi(0)} = 0]
\end{aligned}$$

$$+ \sum_{p=1}^{+\infty} \frac{1}{p^2} \mathbb{E} \left[ h \left( \sum_{k=1}^n e^{-\kappa(T-v_k)} f(\bar{\eta}_k) + \sum_{i=1}^p f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \mathbb{E} \left[ \left( H_T^{\mu+n\Phi(0)} \right)^2 \right].$$

Then by [7, Proposition 5],  $\mathbb{E} \left[ \left( H_T^{\mu+n\Phi(0)} \right)^2 \right] = (\mu + n\Phi(0))C_1 + (\mu + n\Phi(0))^2 C_2$  and by [7, Proposition 7],  $\mathbb{P}[H_T^{\mu+n\Phi(0)} = 0] = e^{-(\mu+n\Phi(0))T}$ . The result follows by injecting these estimates in the previous one.  $\square$

Once again, we consider as a corollary the case where  $\kappa = 0$ .

**Corollary 4.9.** *Assume  $h$  is non-decreasing,  $\kappa = 0$  and  $\Phi$  non-increasing. We consider  $C_1$ ,  $C_2$  and  $c_n$  as defined in Proposition 4.8. For  $n \geq 1$  set*

$$\begin{aligned} \beta_n := & e^{-T(\mu+n\Phi(0))} \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n f(\bar{\eta}_k) \right) \middle| \bar{\eta}_1 \right] \right] \\ & + \sum_{p=1}^{+\infty} \frac{c_n}{p^2} \mathbb{E} \left[ g(\bar{\eta}_1, \bar{\vartheta}_1) \mathbb{E} \left[ h \left( \sum_{k=1}^n f(\bar{\eta}_k) + \sum_{i=1}^p f(\eta_i) \right) \middle| \bar{\eta}_1 \right] \right]. \end{aligned}$$

We have that

$$\mathbb{E}[K_T h(L_T)] \leq \mu \sum_{n=1}^{+\infty} m_\Phi(\Delta^n) \beta_n, \quad (4.11)$$

where we recall  $m_\Phi(\Delta^n)$  is defined in Notation 3.12.

**Remark 4.10.** Note that in case of a Poisson process, the sum in the right-hand-side of (4.11) resumes to the term  $n = 1$  (since  $\Phi \equiv 0$ ) which is exactly equal to  $\mathbb{E}[K_T h(L_T)]$ .

**Remark 4.11.** One can relax the monotony assumption on  $\Phi$  above by allowing a general bounded  $\Phi$  map. In that case, the shifts by  $n\Phi(0)$  are replaced with  $n\Phi^*$  with  $\Phi^* := \sup_{x \in \mathbb{R}_+} \Phi(x)$ .

## 5 Technical material

### 5.1 Proof of Theorem 2.7

We prove the existence and uniqueness of (2.2)

$$\begin{cases} \hat{H}_t^v = h^v + \int_{(v,t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \hat{\Lambda}_s^v\}} N(ds, d\theta), & t \in [v, T] \\ \hat{\Lambda}_t^v = \mu^v(t) + \int_{(v,t]} \Phi(t-u) d\hat{H}_u^v \end{cases}$$

for  $v$  in  $[0, T]$ , where  $(\mu^v(t))_{t \in [v, T]}$  a non-negative stochastic process such that for any  $t \geq v$ ,  $\mu^v(t)$  is a  $\mathcal{F}_v^N$ -measurable random variable and  $h^v$  a  $\mathcal{F}_v^N$ -measurable random with values in  $\mathbb{N}$ . The proof is composed of two parts : existence and uniqueness.

#### Existence

We start with the existence part. We set

$$\Lambda_t^{(1)} := \mu^v(t), \quad t \in [v, T].$$

We make use the following notation : if  $\tau_i^N$  denotes the  $i$ th jump time of  $N$ , then there exists a unique element  $\theta_i \geq 0$  such that  $N(\{(\tau_i^N) \times \{\theta_i\}\}) = 1$  (in other words we denote by  $\theta_i$  the mark associated to the  $i^{\text{th}}$ -jump time of  $N$ ). Let

$$\tau_1^{\hat{H}} := \inf \left\{ \tau_i^N \geq v, \theta_1 \leq \Lambda_{\tau_i^N}^{(1)}, \quad i \geq 1 \right\} \wedge T.$$

In addition,  $\tau_1^{\hat{H}}$  is a  $\mathbb{F}^N$ -stopping time, indeed for any  $t \in [0, T]$ ,

$$\{\tau_1^{\hat{H}} \leq t\} = \bigcup_{i \geq 1} \left[ \{v \leq \tau_i^N \leq t\} \cap \{\theta_i \leq \Lambda_{\tau_i^N}^{(1)}\} \right] \in \mathcal{F}_t^N.$$

For  $i \geq 1$ , we set

$$\begin{aligned} \Lambda_t^{(i+1)} &:= \Lambda_t^{(i)} \mathbf{1}_{v \leq t \leq \tau_i^{\hat{H}}} + \Phi(t - \tau_i^{\hat{H}}) \mathbf{1}_{\tau_i^{\hat{H}} < t \leq T}, \quad t \in [0, T], \\ \tau_{i+1}^{\hat{H}} &:= \inf \left\{ \tau_k^N, \tau_k^N > \tau_i^{\hat{H}}, \theta_k \leq \Lambda_{\tau_k^N}^{(i+1)}, \quad k \geq 1 \right\} \wedge T. \end{aligned} \quad (5.1)$$

By induction, one proves that for any  $i \geq 1$ ,  $\tau_i^{\hat{H}}$  is a  $\mathbb{F}^N$ -stopping time and  $\Lambda^{(i)}$  is a  $\mathbb{F}^N$ -predictable stochastic process as a càglàd,  $\mathbb{F}^N$ -adapted process. In addition, by construction,  $\Lambda^{(i+1)}$  and  $\Lambda^{(i)}$  coincide on  $[0, \tau_i^{\hat{H}}]$ . Furthermore,  $\lim_{i \rightarrow +\infty} \tau_i^{\hat{H}} = T$ ,  $\mathbb{P}$ -*p.s.* as  $\tau_i^{\hat{H}} \geq \tau_i^N$ ,  $\mathbb{P}$ -a.s. for any  $i \geq 1$ . We set :

$$\Lambda_t := \lim_{i \rightarrow +\infty} \Lambda_t^{(i)}, \quad t \in [0, T], \quad (5.2)$$

which is a  $\mathbb{F}$ -predictable process. We then set :

$$\hat{H}_t := h^v + \sum_{i=1}^{+\infty} \mathbf{1}_{\{t \geq \tau_i^{\hat{H}}\}}, \quad t \in [v, T].$$

### Uniqueness

We now turn to the uniqueness of the solution. To do so we give some immediate properties of any solution to (2.3). Consider,  $\hat{H}$  be a solution to (2.3). Then, by definition, we have that

$$\{\tau_k^{\hat{H}}, k \in \mathbb{N}\} \subset \{\tau_k^N, k \in \mathbb{N}\}.$$

In addition, by definition, the jump times of  $\hat{H}$  are given by

$$\tau_i^{\hat{H}} := \inf \left\{ \tau_k^N, \tau_k^N > \tau_{i-1}^{\hat{H}}, \theta_k \leq \Lambda_{\tau_k^N}, \quad k \geq 1 \right\} \wedge T,$$

where

$$\Lambda_t = \mu^v(t) + \int_{(v,t)} \Phi(t-s) d\hat{H}_s, \quad t \in [v, T]. \quad (5.3)$$

Consider,  $\hat{H}^1, \hat{H}^2$ , two solution processes to (2.3). Denote by  $\tau_k^{\hat{H}^1}$  (respectively  $\tau_k^{\hat{H}^2}$ ) the jump times of  $\hat{H}^1, \hat{H}^2$  the intensity function of  $\hat{H}^1$  (respectively  $\Lambda^{\hat{H}^2}$  the one of  $\hat{H}^2$ ), with

$$\Lambda_t^{\hat{H}^i} = \mu^v(t) + \int_{(v,t)} \Phi(t-s) d\hat{H}_s^i, \quad t \in [v, T], \quad i \in \{1, 2\}.$$



By the previous remark,  $\Lambda^{\hat{H}^1} = \Lambda^{\hat{H}^2}$  on  $[0, \tau_1^{\hat{H}^1} \wedge \tau_1^{\hat{H}^2})$  and so  $\tau_1^{\hat{H}^1} = \tau_1^{\hat{H}^2}$ ,  $\mathbb{P}$ -a.s.. Let  $\tau_1 := \tau_1^{\hat{H}^1} (= \tau_1^{\hat{H}^2})$ . We have thus that  $\Lambda^{\hat{H}^1}(\tau_1) = \Lambda^{\hat{H}^2}(\tau_1)$ . Let  $i \geq 1$ . Assume that

$$\tau_i^{\hat{H}^1} = \tau_i^{\hat{H}^2} (=:\tau_i), \quad \hat{H}^1 = \hat{H}^2 \text{ on } [0, \tau_i], \quad \text{and } \Lambda^{\hat{H}^1} = \Lambda^{\hat{H}^2} \text{ on } [0, \tau_i], \quad \mathbb{P} - a.s..$$

Then, as the intensity functions agree up to  $\tau_i$  and as the jump times of any solution is characterized by (5.2), we deduce that  $\tau_{i+1}^{\hat{H}^1} = \tau_{i+1}^{\hat{H}^2} (=:\tau_{i+1})$ ,  $\mathbb{P}$ -a.s.. Hence,

$$\hat{H}^1 = \hat{H}^2 \text{ on } [0, \tau_i + 1], \quad \text{and } \Lambda^{\hat{H}^1} = \Lambda^{\hat{H}^2} \text{ on } [0, \tau_{i+1}], \quad \mathbb{P} - a.s..$$

Thus,  $\hat{H}_t^1 = \hat{H}_t^2$  for any  $t$  in  $[0, T]$ ,  $\mathbb{P}$ -a.s. (as  $\lim_{i \rightarrow +\infty} \tau_i = T$ ,  $\mathbb{P} - a.s.$ ).

## 5.2 Proof of Proposition 3.8

Let  $n \in \mathbb{N}^*$ , and  $0 < v_n < v_{n-1} < \dots < v_1 < T$ . We prove that

$$(H, \Lambda) \circ \varepsilon_{(v_1, \Lambda_{v_1})}^+ \circ \dots \circ \varepsilon_{(v_n, \Lambda_{v_n})}^+ = (H^{v_n, \dots, v_1}, \Lambda^{v_n, \dots, v_1}).$$

*Proof.* We set  $(H^{v_0}, \Lambda^{v_0}) := (H, \Lambda)$ . By Lemma 3.5,  $(H^{v_1}, \Lambda^{v_1}) = (H^0 \circ \varepsilon_{(v_1, \Lambda_{v_1})}^+, \Lambda^0 \circ \varepsilon_{(v_1, \Lambda_{v_1})}^+)$ . Let  $n \geq 2$  and assume that

$$(H^{v_{n-2}, \dots, v_1} \circ \varepsilon_{(v_{n-1}, \Lambda_{v_{n-1}})}^+, \Lambda^{v_{n-2}, \dots, v_1} \circ \varepsilon_{(v_{n-1}, \Lambda_{v_{n-1}})}^+) = (H^{v_{n-1}, \dots, v_1}, \Lambda^{v_{n-1}, \dots, v_1}).$$

We prove that

$$(H^{v_{n-1}, \dots, v_1} \circ \varepsilon_{(v_n, \Lambda_{v_n})}^+, \Lambda^{v_{n-1}, \dots, v_1} \circ \varepsilon_{(v_n, \Lambda_{v_n})}^+) = (H^{v_n, \dots, v_1}, \Lambda^{v_n, \dots, v_1}).$$

For simplicity of notations, we set  $:(H^{n-1}, \Lambda^{n-1}) := (H^{v_{n-1}, \dots, v_1}, \Lambda^{v_{n-1}, \dots, v_1})$  and  $(H^{n-1,+}, \Lambda^{n-1,+}) := (H^{v_{n-1}, \dots, v_1} \circ \varepsilon_{(v_n, \Lambda_{v_n})}^+, \Lambda^{v_{n-1}, \dots, v_1} \circ \varepsilon_{(v_n, \Lambda_{v_n})}^+)$ . Since

$$\left\{ \begin{array}{l} H_t^{n-1} = \mathbf{1}_{[0, v_n)}(t)H_t + \mathbf{1}_{[v_n, v_{n-1})}(t)H_t \\ \quad + \sum_{i=1}^{n-1} \mathbf{1}_{[v_i, v_{i-1})}(t) \left( H_{v_i^-}^{n-1} + 1 + \int_{(v_i, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_s^{n-1}\}} N(ds, d\theta) \right), \\ \Lambda_t^{n-1} = \mathbf{1}_{(0, v_n]}(t)\Lambda_t + \mathbf{1}_{(v_n, v_{n-1}]}(t)\Lambda_t \\ \quad + \sum_{i=1}^{n-1} \mathbf{1}_{(v_i, v_{i-1}]}(t) \left( \mu + \int_{(0, v_i]} \Phi(t-u) dH_u^{n-1} + \int_{(v_i, t)} \Phi(t-u) dH_u^{n-1} \right), \end{array} \right.$$

$$\begin{aligned} H_t^{n-1,+} &= \mathbf{1}_{[0, v_n)}(t)H_t + \mathbf{1}_{[v_n, v_{n-1})}(t) \left( H_{v_n^-} + 1 + \int_{(v_n, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_s^{n-1,+}\}} N(ds, d\theta) \right) \\ &\quad + \sum_{i=n-1}^1 \mathbf{1}_{[v_i, v_{i-1})}(t) \left( H_{v_i^-}^{n-1,+} + 1 + \int_{(v_i, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_s^{n-1,+}\}} N(ds, d\theta) \right) \\ &= \mathbf{1}_{[0, v_n)}(t)H_t + \sum_{i=n}^1 \mathbf{1}_{[v_i, v_{i-1})}(t) \left( H_{v_i^-}^{n-1,+} + 1 + \int_{(v_i, t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \Lambda_s^{n-1,+}\}} N(ds, d\theta) \right). \end{aligned}$$

In addition,

$$\begin{aligned}
\Lambda_t^{n-1,+} &= \mathbf{1}_{(0,v_n]}(t)\Lambda_t \\
&+ \mathbf{1}_{(v_n,v_{n-1}]}(t) \left( \mu + \int_{(0,v_n)} \Phi(t-u)dH_u + \Phi(t-v_n) + \int_{(v_n,t)} \Phi(t-u)dH_u \circ \varepsilon_{((v_n,\Lambda(v_n)))}^+ \right) \\
&+ \sum_{i=1}^{n-1} \mathbf{1}_{(v_i,v_{i-1}]}(t) \left( \mu + \int_{(0,v_i)} \Phi(t-u)dH_u^{n-1,+} + \int_{(v_i,t)} \Phi(t-u)dH_u^{n-1,+} \right) \\
&= \mathbf{1}_{(0,v_n]}(t)\Lambda_t + \mathbf{1}_{(v_n,v_{n-1}]}(t) \left( \mu + \int_{(0,v_n)} \Phi(t-u)dH_u^{n-1,+} + \int_{(v_n,t)} \Phi(t-u)dH_u^{n-1,+} \right) \\
&+ \sum_{i=1}^{n-1} \mathbf{1}_{(v_i,v_{i-1}]}(t) \left( \mu + \int_{(0,v_i)} \Phi(t-u)dH_u^{n-1,+} + \int_{(v_i,t)} \Phi(t-u)dH_u^{n-1,+} \right) \\
&= \mathbf{1}_{(0,v_n]}(t)\Lambda_t + \sum_{i=1}^n \mathbf{1}_{(v_i,v_{i-1}]}(t) \left( \mu + \int_{(0,v_i)} \Phi(t-u)dH_u^{n-1,+} + \int_{(v_i,t)} \Phi(t-u)dH_u^{n-1,+} \right).
\end{aligned}$$

Hence,  $(H^{n-1,+}, \Lambda^{n-1,+})$  solves the same equation than  $(H^{v_n,\dots,v_1}, \Lambda^{v_n,\dots,v_1})$ , which concludes the proof.  $\square$

### 5.3 Two comparison lemma

We provide two comparison lemma based on the thinning algorithm.

**Lemma 5.1.** *Let  $\mu > 0$  and  $(\hat{H}, \hat{\Lambda})$  the unique solution to*

$$\begin{cases} \hat{H}_t = \int_{(0,t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \hat{\Lambda}_s\}} N(ds, d\theta), & t \in [0, T] \\ \hat{\Lambda}_t = \mu + \int_{(0,t)} \Phi(t-u)d\hat{H}_u. \end{cases}$$

*Consider the homogeneous Poisson process  $\tilde{H}$  defined as*

$$\tilde{H}_t = \int_{(0,t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\theta \leq \mu\}} N(ds, d\theta), \quad t \in [0, T].$$

*It holds that*

$$\hat{H}_t \geq \tilde{H}_t, \quad \forall t \in [0, T], \mathbb{P} - a.s..$$

*Proof.* Let  $t$  in  $[0, T]$ . As  $\hat{\Lambda}_t \geq \mu$  we have that

$$\begin{aligned}
\hat{H}_t - \tilde{H}_t &= \int_{(0,t]} \int_{\mathbb{R}_+} \left( \mathbf{1}_{\{\theta \leq \hat{\Lambda}_s\}} - \mathbf{1}_{\{\theta \leq \mu\}} \right) N(ds, d\theta) \\
&= \int_{(0,t]} \int_{\mathbb{R}_+} \mathbf{1}_{\{\mu < \theta \leq \hat{\Lambda}_s\}} N(ds, d\theta) \\
&\geq 0.
\end{aligned}$$

$\square$

**Lemma 5.2.** Assume  $\Phi$  is non-increasing. Let  $n \geq 1$ ,  $0 < v_n < \dots < v_1 < T$ . Let  $(H^{\mu+n\Phi(0)}, \Lambda^{\mu+n\Phi(0)})$  the solution to SDE (2.3) with initial intensity  $\mu + n\Phi(0)$  instead of  $\mu$ . Consider as well  $(H^{v_n, \dots, v_1}, \Lambda^{v_n, \dots, v_1})$  the process defined in Definition 3.6. It holds that :

$$H_t^{v_n, \dots, v_1} \leq n + H_t^{\mu+n\Phi(0)}, \quad \forall t \in [0, T], \mathbb{P} - a.s..$$

In addition, if  $\tau_i^{\mu+n\Phi(0)}$  denotes a jump time of  $H^{\mu+n\Phi(0)}$ , then

$$\mathbb{P}[\exists k \in \{1, \dots, n\}, \tau_i^{\mu+n\Phi(0)} = v_k] = 0$$

and  $\{\tau_i^{H^{v_n, \dots, v_1}}, i \geq 1\} \setminus \{v_1, \dots, v_n\} \subset \{\tau_i^{H^{\mu+n\Phi(0)}}, i \geq 1\}$  almost surely.

*Proof.* On  $[0, v_n)$ , we have that

$$H_t^{v_n, \dots, v_1} = H_t \leq H_t^{\mu+n\Phi(0)}$$

as

$$\Lambda_t^{v_n, \dots, v_1} = \Lambda_t \leq \Lambda_t^{\mu+n\Phi(0)}, \quad t \in [0, v_n].$$

At  $t = v_n$ , we have that  $H_t^{v_n, \dots, v_1} = H_t + 1 \leq H_t^{\mu+n\Phi(0)} + n$ . In addition,

$$\Lambda_{v_n+}^{v_n, \dots, v_1} = \Lambda_{v_n} + \Phi(0) \leq \Lambda_{v_n} + n\Phi(0) \leq \Lambda_{v_n+}^{\mu+n\Phi(0)}.$$

It is important to note that the jump at time  $v_n$  for  $H^{v_n, \dots, v_1}$  will impact the self-exciting part of the intensity while for  $H^{\mu+n\Phi(0)}$  we have only shifted the baseline intensity. Hence, as  $\Phi$  is decreasing, the previous inequality propagates beyond  $v_n$ . The previous inequalities transfer to any interval  $[v_i, v_{i-1})$  until time  $T$ . As  $\Lambda_t^{v_n, \dots, v_1} \leq \Lambda_t^{\mu+n\Phi(0)}$ , using the thinning algorithm, any non-deterministic jump of  $H_t^{v_n, \dots, v_1}$  is a jump  $H^{\mu+n\Phi(0)}$  (the deterministic jumps being  $v_n, \dots, v_1$ ).  $\square$

## References

- [1] Willem Albers. Stop-loss premiums under dependence. *Insurance: Mathematics and Economics*, 24(3):173–185, 1999.
- [2] Pierre Brémaud and Laurent Massoulié. Stability of nonlinear Hawkes processes. *Annals of Probability*, 24(3):1563–1588, 1996.
- [3] Manon Costa, Carl Graham, Laurence Marsalle, and Tran Viet Chi. Renewal in Hawkes processes and self-excitation and inhibition. Preprint : arXiv:1801.04645v2, 2020.
- [4] Daryl J. Daley and David Vere-Jones. *An Introduction to the Theory of Point Processes*. Springer. New York, 1988.
- [5] A. Dassios and H. Zhao. Ruin by dynamic contagion claims. *Insurance: Mathematics and Economics*, 51(1):93–106, 2012.
- [6] Maria de Lourdes Centeno. Dependent risks and excess of loss reinsurance. *Insurance: Mathematics and Economics*, 37(2):229–238, 2005.

- [7] Xuefeng Gao, Xiang Zhou, and Lingjiong Zhu. Transform analysis for Hawkes processes with applications in dark pool trading. *Quantitative Finance*, 18(2):265–282, 2018.
- [8] Xuefeng Gao, Lingjiong Zhu, et al. Large deviations and applications for markovian Hawkes processes with a large initial intensity. *Bernoulli*, 24(4A):2875–2905, 2018.
- [9] Hans U Gerber. On the numerical evaluation of the distribution of aggregate claims and its stop-loss premiums. *Insurance: Mathematics and Economics*, 1(1):13–18, 1982.
- [10] Alan G Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.
- [11] C. Hillairet, Y. Jiao, and A. Réveillac. Pricing formulae for derivatives in insurance using the malliavin calculus. *Probability, Uncertainty and Quantitative Risk*, 3(7):1–19, 2018.
- [12] Jean Jacod and Anatoli V Skorohod. Jumping filtrations and martingales with finite variation. In *Séminaire de Probabilités XXVIII*, pages 21–35. Springer, 1994.
- [13] Hanna Magnusson. A risk surplus model using Hawkes point processes, 2015.
- [14] D. Nualart and J. Vives. *Anticipative calculus for the poisson process based on the fock space.*, volume vol 1426 of *In: Azéma J., Yor M., Meyer P.A. (eds) Séminaire de Probabilités XXIV 1988/89. Lecture Notes in Mathematics*. Springer, Berlin, Heidelberg., 1990.
- [15] Yosihiko Ogata. On Lewis’ simulation method for point processes. *IEEE Transactions of Information Theory*, 27(23–31), 1981.
- [16] J. Picard. Formules de dualité sur l’espace de Poisson. *Ann. Inst. H. Poincaré Probab. Statist.*, 32(4):509–548, 1996.
- [17] Rajko Reijnen, Willem Albers, and Wilbert CM Kallenberg. Approximations for stop-loss reinsurance premiums. *Insurance: Mathematics and Economics*, 36(3):237–250, 2005.
- [18] Gabriele Stabile and Giovanni Luca Torrisi. Risk processes with non-stationary Hawkes claims arrivals. *Methodology and Computing in Applied Probability*, 12(3):415–429, 2010.