MOMENT GENERATING FUNCTION OF NON-MARKOV SELF-EXCITED CLAIMS PROCESSES

Donatien Hainaut

LIDAM Discussion Paper ISBA
2021 / 28

UCLouvain

LIDAM
ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

Moment generating function of non-Markov self-excited claims processes

Donatien Hainaut*
LIDAM, Université Catholique de Louvain, Belgium

July 2, 2021

Abstract

This article establishes the moment generating function (mgf) of self-excited claim processes with memory functions that admit a Fourier's transform representation. In this case, the claim and intensity processes may be reformulated as an infinite dimensional Markov processes in the complex plane. Approaching these processes by discretization and next considering the limit allows us to find their moment generating function. We illustrate the article by fitting non-Markov self-excited processes to the time-series of cyber-attacks targeting medical and other services, in the US from 2014 to 2018.

KEYWORDS : self-excited process, shot noise process, Hawkes process.
JEL CODE : C5, G22

1 Introduction

Self-excited processes offer a natural solution to introduce contagion between claims. In this approach, the occurrence of a claim depends on previous ones. This dynamic was initially introduced by Hawkes (1971a, b) and Hawkes and Oakes (1974). In the most common specification, the intensity of claim numbers, that is akin to the instantaneous probability of a claim, increases as soon as a claim is observed. The influence of this claim on the intensity next decays with time according to a memory function, also called kernel. Musmeci and Vere-Jones (1992) or Ogata (1998) propose such a point-process for earthquake occurrences. Porter and White (2012) use a similar self-excited process for modeling terrorist activity. Barsotti et al. (2016) model the contagion of lapses in life insurance with exponential self-excited processes. Hainaut (2017) introduces contagion between insurance and financial markets with a bivariate self-excited process. More recently Bessy-Roland et al. (2020) fit a multivariate extension of such processes to cyber-claims. The reader may refer to Reinhart (2018) for a detailed review of other applications and properties of Hawkes processes.

*Postal address: Voie du Roman Pays 20, 1348 Louvain-la-Neuve (Belgium). E-mail to: donatien.hainaut(at)uclouvain.be

In a standard version, the claim-intensity process is Markov. In this case, the moment generating function (mgf) admits an analytical solution, found by Itô's calculus. This approach is e.g. applied in Dassios and Jiwook (2003) to price catastrophe reinsurance, in Jiwook and Dassios (2013) for a bivariate model or in Jiwook et al. (2018) for the calculation of moments of shot-noise processes. In a general setting, we lose most of the analytical tractability offered by stochastic calculus mainly because the claim-intensity process is not Markov anymore. Apart from moments or asymptotic properties, very few results are available in the literature. For instance, Muzy et al. (2013) find the first moments of stationary processes and limit behaviour. Stabile and Torrisi (2010) study the asymptotic behavior of infinite and finite horizon ruin probabilities of non-stationary hawkes claim arrivals. This article fills a gap in the literature by providing a closed-form expression of the moment generating function (mgf) of non-Markov self-excited claim processes.

We show that when the memory kernel is the Fourier's transform of a measure of probability, the self-excited claim process can be reformulated as an infinite dimensional Markov process in the complex plane. We focus on four types of memory kernels associated to Laplace, Gaussian, logistic and Cauchy measures. Next, we approach the infinite dimensional Markov model by a finite dimensional process and obtained its mgf using standard stochastic calculus. The mgf of the infinite dimensional process is obtained by considering the limit of the finite dimensional approximation. As illustration, we exploit these results to compute the probability density function of cyber-attacks in the USA, targeting medical and other services.

The paper proceeds as follows. Section 2 introduces the multivariate self-excited claim processes and the four categories of memory kernels. Taking advantage of the Fourier representation of memory functions, we reformulate the model as an infinite dimensional Markov process in Section 3. The continuous probability measure defining the memory kernel is approached by an atomic measure in Section 4. We infer in this framework the mgf of claim processes. Section 5 presents the mgf of original processes. Section 6 focuses on memory kernels linked to a Cauchy measure. In this case and under some conditions on parameters, the claim-intensity process is Markov. This particular case has been extensively studied in the literature and Section 6 reconciliates the expression of the mgf proposed in this article with the standard formulation. The next section presents a numerical illustration. The different models are fitted to time series of cyber-attacks targeting medical and other services in the USA. In the last Section, we compute marginal and multivariate probability density functions of claim processes with a discrete Fourier transform.

2 The claim processes

We consider an insurance company with m business lines covering claims that are mutually and self-excited. The aggregate claim process of the k^{th} risk is denoted by $\left(L_t^k\right)_{t\geq 0}$ for

$k = 1, ..., m$. This is the sum of N_t^k claims noted $J_j^k \in \mathbb{R}^+$:

$$\left(L_t^k\right)_{t\geq 0} = \sum_{j=1}^{N_t^k} J_j^k \ , k = 1, ..., m. \tag{1}$$

The counting process $\left(N_t^k\right)_{t\geq 0}$ is a point process with an intensity denoted by $\left(\lambda_t^k\right)_{t\geq 0}$. All processes are defined on a probability space Ω with a measure $\mathbb{P}$ and a filtration $(\mathcal{F}_t)_{t\geq 0}$. The claims $\left(J_j^k\right)$ are an independent copies of a random variable J^k. The probability density function (pdf) of J^k is d $\zeta_k(z)$ for $z \in \mathbb{R}^+$. The intensities of counting processes $\left(N_t^k\right)_{k=1,\ldots,m}$ are akin to the instantaneous probabilities of observing a new claims of category k. To introduce memory and contagion in the dynamic of claims, the arrival intensities are assumed driven by the following equation

$$\lambda_t^k = \theta_k + \left(\lambda_0^k - \theta_k\right) g_{kk}(t) + \sum_{j=1}^{m} \eta_{kj} \int_0^t g_{kj}(t-s) dL_s^j \quad , \ k = 1, ..., m. \tag{2}$$

where $(g_{kj}(t))_{k,j=1,\ldots,m}$ are positive definite functions such that $g_{kj}(0) = 1$ and $\lim_{t\to\infty} g_{kj}(t) = 0$. The parameters θ_k and $\eta_{k,j}$ are positive. In this framework, the occurence of a claim J^j of type j immediately raises the probability of observing a new claim accross all categories, proportionally to $(\eta_{k,j} J^j)_{k=1,\ldots,m}$. Between two successive claims, the intensities revert to $(\theta_k)_{k=1,\ldots,m}$. The speed of reversion depends upon the decay rate of $g_{kk}(.)$. The functions $(g_{kj}(t))_{k,j=1,\ldots,m}$ are kernels functiong defining the "memory" of claim processes. If they are fast decreasing, the contagion effect between claims is limited in time. In the opposite case, the processes remember for a long period of time the history of claims.

Let us adopt the following vector and matrix notations: $\boldsymbol{\lambda}_t = (\lambda_t^1, ..., \lambda_t^m)^\top$, $\boldsymbol{\theta} = (\theta^1, ..., \theta^m)^\top$, $\boldsymbol{L}_t = (L_t^1, ..., L_t^m)^\top$,$\boldsymbol{N}_t = (N_t^1, ..., N_t^m)^\top$, $G(t) = (g_{kj}(t))_{k,j=1,\ldots,m}$ and $H = (\eta_{kj})_{k,j=1,\ldots,m}$. This allows us to reformulate the dynamic of intensity in a concise way

$$\boldsymbol{\lambda}_t = \boldsymbol{\theta} + (\boldsymbol{\lambda}_0 - \boldsymbol{\theta})^\top \operatorname{diag}(G(t)) + H \int_0^t G(t-s)\, d\boldsymbol{L}_s \,,$$

where $\operatorname{diag}(G(t))$ is the matrix of diagonal elements of $G(t)$. The vector of expected claim sizes is denoted by $\boldsymbol{\mu}_J = (\mathbb{E}(J^1), ..., \mathbb{E}(J^m))^\top$. Let

$$||\Xi|| = \left(\int_0^\infty |\eta_{kj} g_{kj}(s) \mathbb{E}(J^j)| ds\right)_{k,j=1,\ldots,m}$$

be the $m \times m$ matrix of L_1-norm of the elementwise product of H, $G(t)$ and expected claims. To ensure the stationarity of N_t, the spectral radius of the matrix $||\Xi||$ must be strictly lower than one. We assume that this assumption is satisfied in the remainder of this article. In this setting, the long-term expectation, noted $\bar{\boldsymbol{\lambda}} = \lim_{t\to\infty} \mathbb{E}\left(\lambda_t\right)$, is such that $\bar{\boldsymbol{\lambda}} = \boldsymbol{\theta} + ||\Xi|| \bar{\boldsymbol{\lambda}}$ and therefore is equal to $\bar{\boldsymbol{\lambda}} = \left(I - ||\Xi||\right)^{-1} \boldsymbol{\theta}$.

We consider in this article functions $g_{kj}(.)$ that admit an integral representation. Bochner

(1933) showed that positive definite functions may be defined by a spectral measure as stated in the next proposition.

Proposition (Bochner) A continuous function $g(h)$ from $\mathbb{R}^+$ to the complex plane $\mathbb{C}$ is positive definite if and only if it may be represented as the Fourier transform of a measure $\nu(.)$ on $\mathbb{R}$:

$$g(h) \;=\; \int_{\mathbb{R}} e^{ihu}\nu(du)\,, \tag{3}$$

where $\nu(.)$ is a bounded, real valued function such that $\int_A \nu(du) \geq 0$ for all $A \subset \mathbb{R}$.

The function $\nu(.)$ is called the spectral distribution function for g. In the sequel of this article, we furthermore assume that $\nu(.)$ is a measure of probability, i.e. $\int_{\mathbb{R}} \nu(du) = 1$. By definition, an alternative representation of the function $g(.)$ is

$$g(h) \;=\; \int_{\mathbb{R}} \left(\cos\left(hu\right) + i\sin\left(hu\right)\right)\nu(du)\,.$$

Therefore, if $\boldsymbol{\lambda}_t$ is $\mathbb{R}$-valued, $g(h)$ is also $\mathbb{R}$-valued and the imaginary part in this last equation must be null: $\int_{\mathbb{R}} \sin\left(hu\right)\nu(du) = 0$. This implies that the spectral distribution is symmetric around the origin. If $\nu(.)$ is continuous, we have then $\nu(du) = \nu(-du)$ and

$$\begin{aligned} g(h) \;&=\; \int_{\mathbb{R}} \cos\left(hu\right)\nu(du)\,, \\ &=\; 2\int_0^{\infty} \cos\left(|h|u\right)\nu(du)\,. \end{aligned}$$

We consider four spectral distributions for $\nu(.)$: the Laplace, Gaussian, logistic and Cauchy measures. The next paragraphs detail each one.

1) The first measure considered is the Laplace or double exponential measure. Let $\alpha \in \mathbb{R}^+$ be constant. A symmetric exponential measure is defined by the following relation:

$$\nu_L(du) \;=\; \frac{1}{2}\left(\alpha e^{-\alpha u}1_{\{u\geq 0\}} + \alpha e^{\alpha u}1_{\{u\leq 0\}}\right)du\,. \tag{4}$$

The variance of $\nu_L(.)$ is $\sigma_L^2(\alpha) = \frac{2}{\alpha^2}$. If we use this function as spectral distribution, we obtain the following memory kernel

$$\begin{aligned} g_L(h) \;&=\; \frac{1}{2}\left[\int_{\mathbb{R}^+} \alpha e^{ihu}e^{-\alpha u}du + \int_{\mathbb{R}^-} \alpha e^{ihu}e^{\alpha u}du\right] \\ &=\; \frac{\alpha^2}{\alpha^2+h^2}\,, \end{aligned} \tag{5}$$

which is a power decreasing function. If $m = 1$, as $\int_0^{\infty} \eta\,\mu_J\,g_L(s)\,ds = \frac{\pi\eta\mu_J\alpha}{2}$ we infer that $\alpha < \frac{2}{\pi\eta\mu_J}$ to ensure stability of λ_t at long-term.

2) The second category of functions $g(.)$ is generated with a Gaussian spectral measure,

$$\nu_G(du) = \sqrt{\frac{\alpha}{\pi}} e^{-\alpha u^2} du\,. \tag{6}$$

The variance of $\nu_G(.)$ is $\sigma_G^2(\alpha) = \frac{1}{2\alpha}$. Given that $\int_{\mathbb{R}^+} e^{-\alpha u^2} \cos(hu)\, du = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}} e^{-\frac{h^2}{4\alpha}}$, the kernel is in this case equal to

$$\begin{aligned} g_G(h) &= 2\sqrt{\frac{\alpha}{\pi}} \int_{\mathbb{R}^+} e^{-\alpha u^2} \cos(|h|u)\, du \\ &= e^{-\frac{h^2}{4\alpha}}\,. \end{aligned} \tag{7}$$

In the one-dimensional case ($m = 1$), $\int_0^\infty \eta\mu_J e^{-\frac{h^2}{4\alpha}} ds = \eta\mu_J\sqrt{\pi}\sqrt{\alpha}$ and we infer that $\sqrt{\alpha} < \frac{1}{\eta\mu_J\sqrt{\pi}}$ to ensure stability of λ_t.

3) The third considered distribution is the logistic one. For $\alpha > 0$, the spectral measure is in this case given by

$$\nu_{log}(du) = \frac{e^{-u/\alpha}}{\alpha\left(1 + e^{-u/\alpha}\right)^2} du, \tag{8}$$

The variance of $\nu_{log}(.)$ is $\sigma_{log}^2(\alpha) = \frac{\alpha^2\pi^2}{3}$ and the memory kernel is equal to

$$g_{log}(h) = \begin{cases} \frac{\pi\alpha h}{\sinh(\pi\alpha h)} & h > 0 \\ 1 & h = 0 \end{cases}.$$

If $m = 1$, as $\int_0^\infty \eta\, \mu_J\, g(s)\, ds = \frac{\eta\mu_J\pi}{4\alpha}$ we infer that $\frac{\eta\mu_J\pi}{4} < \alpha$ to ensure stability of λ_t.

4) The fourth type of spectral measure that we consider is the Cauchy distribution:

$$\nu(du) = \frac{1}{\pi} \frac{\alpha}{\alpha^2 + u^2} du\,, \tag{9}$$

where $\alpha \in \mathbb{R}^+$. By construction this measure is symmetric and the function $g(h)$ is an exponential decreasing function:

$$\begin{aligned} g(h) &= \frac{2\alpha}{\pi} \int_{\mathbb{R}^+} \frac{\cos(|h|u)}{\alpha^2 + u^2} du \\ &= e^{-\alpha|h|}\,. \end{aligned} \tag{10}$$

If $m = 1$, as $\int_0^\infty \eta\mu_J e^{-\alpha s} ds = \frac{\eta\mu_J}{\alpha}$, we infer that $\eta\mu_J < \alpha$ to ensure stability of λ_t.

This fourth spectral measure encompasses the classical Hawkes process (1971 a,b). This kind of processes has already been widely studied in the literature and will serves as benchmarck in this work. If the function $(g_{kj})_{k,j=1,\ldots,m}$ are defined by Equation (10) with parameters

$\alpha_{kj} = \alpha_k$ for $k = 1, ..., m$, the vector of densities is solution of the stochastic differential equations (SDE's):

$$d\boldsymbol{\lambda}_t = \text{diag}\left((\alpha_k)_{k=1,\ldots,m}\right)(\boldsymbol{\lambda}_0 - \boldsymbol{\theta})\,dt + H\,d\boldsymbol{L}_t\,,$$

and the multivariate process $(\boldsymbol{\lambda}_t, \boldsymbol{L}_t)_{t\geq 0}$ is Markov. It means that for a well behaved function $f(\boldsymbol{\lambda}_t, \boldsymbol{L}_t)$, the conditional expectation $\mathbb{E}\left(f(\boldsymbol{\lambda}_s, \boldsymbol{L}_s)|\mathcal{F}_t\right)$ for $s \geq t$, depends exclusively on $\boldsymbol{\lambda}_t$ and $\boldsymbol{L}_t$. Furthermore, we can rely on Itô calculus for determining e.g. the moment generating functions (mgf's) of claim processes. The mgf is of particular importance since it may be numerically inverted by a discrete Fourier's transform to retrieve the pdf of $\boldsymbol{L}_t$. Notice that when the condition $\alpha_{kj} = \alpha_k$ is not fulfilled, the process$(\boldsymbol{\lambda}_t, \boldsymbol{L}_t)_{t\geq 0}$ is not Markov anymore.

In order to compare memory kernels, we adjust parameters α in order to match variances of spectral densities. For a given $\sigma \in \mathbb{R}^+$, we set parameters of Laplace, Gaussian and logistic kernels to

$$\alpha_L = \sqrt{\frac{2}{\sigma^2}}\,,\ \alpha_G = \frac{1}{2\sigma^2}\,,\ \alpha_{log} = \sqrt{\frac{3\sigma^2}{\pi^2}}\,.$$

For the Cauchy distribution, the variance is not defined. Therefore, we simply find an α by minimizing the sum of least square between the Log an Cauchy memory kernels. The result of this procedure for $\sigma^2 = 16$ is illustrated in Figure 1. This graph emphasizes that the Laplace kernel has a slower decay than Gaussian and logistic ones while the Gaussian kernel quickly vanishes at mid-term.

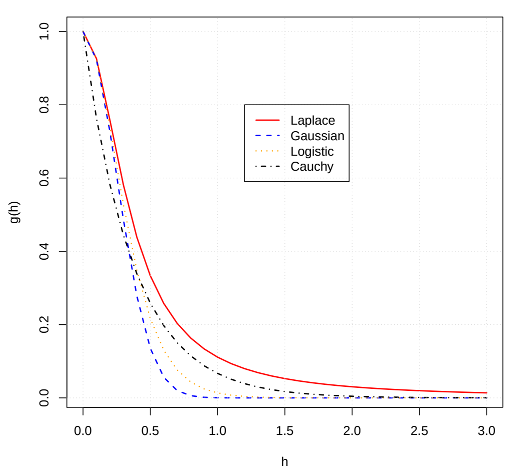

Figure 1: Comparison of Laplace, Gaussian, Logistic kernels for $\sigma_L^2(\alpha_L) = \sigma_G^2(\alpha_G) = \sigma_{log}^2(\alpha_{log}) = 4$. The Cauchy kernel is computed with $\alpha_C = 2.71$.

For most of memory kernels, the process $(\boldsymbol{\lambda}_s, \boldsymbol{L}_s)_{s\geq 0}$ is not Markov and we have very few information on its behavior conditionally to the filtration $\mathcal{F}_t$ for $t \leq s$. The next sections propose an alternative formulation that will allows us to find later the moment generating function of claim processes.

3 Infinite dimensional reformulation

Except for the Cauchy kernel with $\alpha_{kj} = \alpha_k$, the process $(\boldsymbol{L}_t, \boldsymbol{\lambda}_t)_{t\geq 0}$ is not Markov. Nevertheless, if functions $g_{kj}(t)$ admit a spectral representation with a measure $\nu_{kj}(z)$, we can reformulate the model as an infinite dimensional Markov process in the complex plane. Indeed, from the Bochner equation (3) and after changing the order of integration, the intensity process may be rewritten as

$$\begin{aligned}
\lambda_t^k &= \theta_k + \left(\lambda_0^k - \theta_k\right) g_{kk}(t) + \sum_{j=1}^{m} \eta_{kj} \int_0^t \int_{\mathbb{R}} e^{i(t-s)\xi} \nu_{kj}(d\xi) dL_s^j \\
&= \theta_k + \left(\lambda_0^k - \theta_k\right) g_{kk}(t) + \sum_{j=1}^{m} \eta_{kj} \int_{\mathbb{R}} \left(\int_0^t e^{i(t-s)\xi} dL_s^j \right) \nu_{kj}(d\xi)\,,
\end{aligned}$$

for $k = 1, ..., m$. Let us define a family of processes $Y_t^{(j,\xi)}$ indexed by $\xi \in \mathbb{R}$ and $j = 1, ..., m$ that is

$$Y_t^{(j,\xi)} = \int_0^t e^{i(t-s)\xi}\, dL_s^j\,.$$

These processes are defined on the complex plane $\mathbb{C}$ and their differential is provided by the SDE

$$dY_t^{(j,\xi)} \quad = \quad i\,\xi\, Y_t^{(j,\xi)} dt + dL_t^j\,. \tag{11}$$

The intensities can then be rewritten as a sum of integrals with respect to $Y_t^{(j,\xi)}$,

$$\lambda_t^k = \theta_k + \left(\lambda_0^k - \theta_k\right) g_{kk}(t) + \sum_{j=1}^{m} \eta_{kj} \int_{\mathbb{R}} Y_t^{(j,\xi)}\, \nu_{kj}(d\xi)\,.$$

As $\nu_{kj}(.)$ is a probability measure, we can check that $\lim_{\xi\to\infty} Y_t^{(j,\xi)}\, \nu_{kj}(d\xi) = 0$ and $\lim_{\xi\to-\infty} Y_t^{(j,\xi)}\, \nu_{kj}(d\xi) = 0$. The differential of intensities is given by

$$d\lambda_t^k = \left(\lambda_0^k - \theta_k\right) \frac{dg_{kk}(t)}{dt} dt + \sum_{j=1}^{m} \eta_{kj} \int_{\mathbb{R}} dY_t^{(j,\xi)}\, \nu_{kj}(d\xi)\,.$$

It is worth noting that λ_t in this reformulated model is well a real number despite that $Y_t^{(j,\xi)}$ is a complex process. To check this, let us denote by τ_l the time of the l^{th} claim of type j. We can then rewrite processes $Y_t^{(j,\xi)}$ and $Y_t^{(j,-\xi)}$ as the following sums

$$\begin{aligned}
Y_t^{(j,\xi)} &= \sum_{l=1}^{N_t^j} e^{i(t-\tau_l)\xi} J_l^j = \sum_{l=1}^{N_t^j} \left[\cos\left((t-\tau_l)\xi\right) + i \sin\left((t-\tau_l)\xi\right)\right] J_l^j\,, \\
Y_t^{(j,-\xi)} &= \sum_{l=1}^{N_t^j} e^{-i(t-\tau_l)\xi} J_l^j = \sum_{l=1}^{N_t^j} \left[\cos\left((t-\tau_l)\xi\right) - i \sin\left((t-\tau_l)\xi\right)\right] J_l^j\,.
\end{aligned}$$

Given that the measures $\nu_{kj}(.)$ are symmetric, the sum of $Y_t^{(j,\xi)}$ and $Y_t^{(j,-\xi)}$ weighted by $\nu_{kj}(d\xi)$ and $\nu_{kj}(-d\xi)$,

$$Y_t^{(j,\xi)}\,\nu_{kj}(d\xi)+Y_t^{(j,-\xi)}\nu_{kj}(-d\xi) \;=\; 2\sum_{l=1}^{N_t^j}\cos\left((t-\tau_l)\xi\right)J_l^j\,\nu_{kj}(d\xi)\,,$$

is real valued and so $\int_{\mathbb{R}} dY_t^{(j,\xi)}\,\nu_{kj}(d\xi)$. From the rewriting of $Y_t^{(j,\xi)}$ as $\sum_{l=1}^{N_t^j} e^{i(t-\tau_l)\xi}J_l^j$, we immediately infer that the quadratic variation of $Y_t^{(j,\xi)}$ is given by

$$[Y^{(j,\xi)},Y^{(j,\xi)}]_t=\sum_{l=1}^{N_t^j} e^{2i(t-\tau_l)\xi}\left(J_l^j\right)^2\,,$$

whereas the quadratic covariation of $Y^{(j_1,\xi_1)}$ and $Y^{(j_2,\xi_2)}$ is equal to

$$[Y^{(j_1,\xi_1)},Y^{(j_2,\xi_2)}]_t=\begin{cases}\sum_{l=1}^{N_t^{j_1}} e^{i(t-\tau_l)(\xi_1+\xi_2)}\left(J_l^{j_1}\right)^2 & j_1=j_2\,,\\ 0 & j_1\neq j_2\,.\end{cases}$$

We have reformulated the process $(\boldsymbol{\lambda}_t,\boldsymbol{L}_t)$ as an infinite dimensional Markov process $\left(\boldsymbol{\lambda}_t,\boldsymbol{L}_t,\left(Y_t^{(j,\xi)}\right)_{\xi\in\mathbb{R},j=1,\dots,m}\right)$. We propose a discrete approximation in the next section and infer the moment generating function of claim processes in this setting. By considering the limit, we retrieve in Section 5 the mgf of the original process.

4 Finite dimensional approximation

Instead of considering an infinity of processes $\left(Y_t^{(j,\xi)}\right)_{\xi\in\mathbb{R}}$, we approach them with a finite number of equivalent processes. This presents several advantages. Firstly, this makes possible implementing our model. Secondly, we can rely on the Itô's calculus to find the mgf's of $\left(\lambda_t^k\right)_{k=1,\dots,m}$. The key step consists to approach $\nu_{kj}(.)$ by a symmetric discrete measure with a finite numbers of atoms. For this purpose we consider a symmetric partition $\mathcal{E}^{(n)}:=\{-\infty<\xi_{-n}^{(n)}<\dots<\xi_0^{(n)}=0<\dots<\xi_n^{(n)}<\infty\}$, with $\xi_{-l}^{(n)}=-\xi_l^{(n)}$. The mid point of each interval $(\xi_l^{(n)},\xi_{l+1}^{(n)})$ is denoted by

$$\begin{aligned} b_{l+1} &= \frac{\xi_l^{(n)}+\xi_{l+1}^{(n)}}{2}\,,l\in\{0,\dots,n-1\} \\ b_l &= \frac{\xi_l^{(n)}+\xi_{l+1}^{(n)}}{2}\,,l\in\{-n,\dots,-1\}\end{aligned} \tag{12}$$

The mass of corresponding atoms for the measure $\nu_{kj}(.)$ is defined as the measure of intervals:

$$\begin{aligned} w_{l+1}^{(k,j)} &= \int_{\xi_l^{(n)}}^{\xi_{l+1}^{(n)}}\nu_{kj}(dz)\,,l\in\{0,\dots,n-1\} \\ w_l^{(k,j)} &= \int_{\xi_l^{(n)}}^{\xi_{l+1}^{(n)}}\nu_{kj}(dz)\,,l\in\{-n,\dots,-1\}\end{aligned} \tag{13}$$

for $l =\in \{-n,...,n\}\backslash\{0\}$ and $k,j=1,...,m$. Due to the symmetry of $\nu_{kj}(.)$, $w_{-l}^{(k,j)} = w_l^{(k,j)}$ and $b_{-l}=-b_l$. The discrete measure for a partition of size n is defined as follows

$$\tilde{\nu}_{kj}(z) = \sum_{l=-n}^{n} w_l^{(k,j)} \delta_{b_l}(z)\,, \tag{14}$$

where $\delta_{b_l^{(k,j)}}(z)$ is the Dirac measure located at point b_l. Notice that the notation $\sum_{l=-n}^{n}$ corresponds to $\sum_{l\in\{-n,...,n\}\backslash\{0\}}$, to lighten the writing. We consider that the following assumption holds for the partition $\mathcal{E}^{(n)}$. Firslty, $\xi_{-n}^{(n)} \to -\infty$ and $\xi_n^{(n)} \to \infty$ when $n\to\infty$. Secondly, $\max|\xi_{i+1}^{(n)} - \xi_i^{(n)}| \to 0$ when $n\to\infty$. Thirdly, $\mathcal{E}^{(n)} \subset \mathcal{E}^{(n+1)}$. In this case, for any function $f(.)$, integrable with respect to $\nu_{kj}(.)$, we have that $\lim_{n\to\infty}\int_{-\infty}^{\infty} f(z)\,\tilde{\nu}_{kj}(dz) = \int_{-\infty}^{\infty} f(z)\,\nu_{kj}(dz)$. To lighten further developments, we adopt the following notations $\tilde{Y}_t^{(j,l)} := Y_t^{(j,b_l)}$ for $j=1,...m$ and $l=-n,...,n$. Each of the $m\times(2n+1)$ processes, $\tilde{Y}_t^{(k,l)}$ is ruled by the SDE

$$d\tilde{Y}_t^{(j,l)} = i\,b_l \tilde{Y}_t^{(j,l)} dt + d\tilde{L}_t^j \quad ,$$

where $\tilde{L}_t^j = \sum_{l=1}^{\tilde{N}_t^j} J_l^j$ is the j^{th} claim process in the discretized model. Its intensity is denoted by $\tilde{\lambda}_t^k$ and is such that

$$\tilde{\lambda}_t^k = \theta_k + \left(\tilde{\lambda}_0^k - \theta_k\right) g_{kk}(t) + \sum_{j=1}^{m} \eta_{kj} \sum_{l=-n}^{n} w_l^{(k,j)} \tilde{Y}_t^{(j,l)}\,.$$

The symetry of $w_l^{(k,j)}$ (which follows from the symetry of the partition $\mathcal{E}^{(n)}$ and of ν_{kj}) ensures that λ_t^k is a real number. Its differential is given by

$$\begin{aligned} d\tilde{\lambda}_t^k &= \left(\tilde{\lambda}_0^k - \theta_k\right) \frac{dg_{kk}(t)}{dt} dt + \sum_{j=1}^{m} \eta_{kj} \sum_{l=-n}^{n} w_l^{(k,j)} d\tilde{Y}_t^{(j,l)} \\ &= \left(\tilde{\lambda}_0^k - \theta_k\right) \frac{dg_{kk}(t)}{dt} dt + \sum_{j=1}^{m}\sum_{l=-n}^{n} \eta_{kj} w_l^{(k,j)} i\,b_l \tilde{Y}_t^{(j,l)} dt \\ &\quad + \sum_{j=1}^{m}\sum_{l=-n}^{n} \eta_{kj} w_l^{(k,j)} d\tilde{L}_t^j\,. \end{aligned} \tag{15}$$

To lighten notations, we reformulate the dynamic of intensities in matrix form. For this purpose, we adopt the following notations: $\tilde{\boldsymbol{\lambda}}_t = \left(\tilde{\lambda}_t^1,...,\tilde{\lambda}_t^m\right)^\top$, $\tilde{\boldsymbol{L}}_t = \left(\tilde{L}_t^1,...,\tilde{L}_t^m\right)^\top$, $\tilde{\boldsymbol{Y}}_t^l = \left(\tilde{Y}_t^{(1,l)},...,\tilde{Y}_t^{(m,l)}\right)^\top$ and $W_l = (w_l^{(k,j)})_{k,j=1,...,m}$ for $l=-n,...,n$. Notice that by construction, we have the two limits,

$$\begin{aligned} \lim_{n\to\infty}\sum_{l=-n}^{n} w_l^{(k,j)} \tilde{Y}_t^{(j,l)} &= \int_{\mathbb{R}} Y_t^{(j,\xi)} \nu_{kj}(d\xi)\,, \\ \lim_{n\to\infty}\sum_{l=-n}^{n} W_l \tilde{\boldsymbol{Y}}_t^l &= \left(\sum_{j=1}^{m}\int_{\mathbb{R}} Y_t^{(j,\xi)} \nu_{kj}(d\xi)\right)_{k=1,...,m}\,. \end{aligned}$$

If $\odot$ is the Hadamard (element wise) product, we reformulate the dynamic of intensities, Equation (15), as follows

$$d\tilde{\boldsymbol{\lambda}}_t = \left[\left(\tilde{\boldsymbol{\lambda}}_0 - \boldsymbol{\theta}\right)^\top \operatorname{diag}\left(\frac{dG(t)}{dt}\right) + i\sum_{l=-n}^{n} b_l\,\left(H\odot W_l\right)\,\tilde{\boldsymbol{Y}}_t^l\right] dt + \sum_{l=-n}^{n}\left(H\odot W_l\right)\,d\tilde{\boldsymbol{L}}_t\,.$$

whereas $d\tilde{\boldsymbol{Y}}_t^l = i\,b_l\tilde{\boldsymbol{Y}}_t^l\,dt + d\tilde{\boldsymbol{L}}_t$. Therefore the differential of any real function $f\left(t,\tilde{\boldsymbol{\lambda}}_t,\left(\tilde{\boldsymbol{Y}}_t^l\right)_{l=-n,\dots,n},\tilde{\boldsymbol{L}}_t\right)$ differentiable once with respect to its arguments can be obtained with the Itô's formula,

$$\begin{aligned} df(.) &= \partial_t f(.)\,dt + i\sum_{l=-n}^{n} b_l\left(\partial_{\tilde{\boldsymbol{Y}}^l} f(.)\right)^\top \tilde{\boldsymbol{Y}}_t^l\,dt + \left(\partial_{\tilde{\boldsymbol{\lambda}}} f(.)\right)^\top \times \\ &\left[\left(\tilde{\boldsymbol{\lambda}}_0 - \boldsymbol{\theta}\right)^\top \operatorname{diag}\left(\frac{dG(t)}{dt}\right) + i\sum_{l=-n}^{n} b_l\,\left(H\odot W_l\right)\,\tilde{\boldsymbol{Y}}_t^l\right] dt + \\ &\int_{\mathbb{R}^{m,+}} f\left(t,\tilde{\boldsymbol{\lambda}}_t + \sum_{l=-n}^{n}\left(H\odot W_l\right)\,\boldsymbol{z},\left(\tilde{\boldsymbol{Y}}_t^l + \boldsymbol{z}\right)_{l=-n,\dots,n},\tilde{\boldsymbol{L}}_t + \boldsymbol{z}\right) \\ &-f(.)\,\tilde{\boldsymbol{L}}(dt,d\boldsymbol{z}) \end{aligned}$$

where $\tilde{\boldsymbol{L}}(.,.)$ is the random measure such that $\tilde{\boldsymbol{L}}_t = \int_0^t\int_{\mathbb{R}^{m,+}} \boldsymbol{z}\,d\tilde{\boldsymbol{L}}(du,d\boldsymbol{z})$. The notations $\partial_t f$, $\partial_{\tilde{\boldsymbol{Y}}^l} f$ and $\partial_{\tilde{\boldsymbol{\lambda}}} f$ are the partial derivatives of $f(.)$ with respect to time, $\tilde{\boldsymbol{Y}}_t^l$ and $\tilde{\boldsymbol{\lambda}}$.

As already mentioned, the mgf of the claim process is of particular importance since it may be numerically inverted by a discrete Fourier's transform to retrieve the pdf of $\boldsymbol{L}_t$. The next proposition provides an analytical expression of this function.

Proposition 4.1. *The moment generating function of claim processes $\mathbb{E}\left(e^{\boldsymbol{\omega}^\top\tilde{\boldsymbol{L}}_s}\,|\,\mathcal{F}_t\right)$ for $\boldsymbol{\omega} = (\omega_1,\dots,\omega_m)^\top \in \mathbb{C}^m$ is equal to*

$$\mathbb{E}\left(e^{\boldsymbol{\omega}^\top\tilde{\boldsymbol{L}}_s}\,|\,\mathcal{F}_t\right) = \exp\left(q_0(t,s) + \boldsymbol{q}_\lambda(t,s)^\top\tilde{\boldsymbol{\lambda}}_t + \sum_{l=-n}^{n}\boldsymbol{q}_l(t,s)^\top\left(H\odot W_l\right)\tilde{\boldsymbol{Y}}_t^l + \boldsymbol{\omega}^\top\tilde{\boldsymbol{L}}_t\right) \tag{16}$$

where $q_0(t,s) : \left(\mathbb{R}^+\right)^2 \to \mathbb{R}$, $\boldsymbol{q}_\lambda(t,s) : \left(\mathbb{R}^+\right)^2 \to \mathbb{R}^m$ and $\boldsymbol{q}_l(t,s) : \left(\mathbb{R}^+\right)^2 \to \mathbb{R}^m$. We denote by $\boldsymbol{e}_j = (0,\dots,1,\dots,0)$, the j^{th} vector of the orthonormed basis of $\mathbb{R}^n$. The element of vectors $\boldsymbol{q}_\lambda(t,s)$ and $\boldsymbol{q}_l(t,s)$ are noted $q_\lambda^k(t,s)$ and $q_l^k(t,s)$. These functions are solutions of the system of ordinary equations s (ODE's)

$$\begin{cases} \partial_t q_0(t,s) = -\boldsymbol{q}_\lambda(t,s)^\top diag\left(\frac{dG(t)}{dt}\right)\left(\tilde{\boldsymbol{\lambda}}_0 - \boldsymbol{\theta}\right) \\ \partial_t q_\lambda^j(t,s) = -\int_0^\infty e^{\left(\sum_{l=-n}^{n}\left(\boldsymbol{q}_\lambda(t,s)^\top + \boldsymbol{q}_l(t,s)^\top\right)(H\odot W_l) + \boldsymbol{\omega}^\top\right)\boldsymbol{e}_j\,z} - 1\,\zeta_j(dz) \\ \partial_t \boldsymbol{q}_l(t,s) = -i\,b_l\left(\,\boldsymbol{q}_l(t,s) + \boldsymbol{q}_\lambda(t,s)\right) \end{cases} \tag{17}$$

for $l=-n,...,n$ and $j=1,...,m$. The terminal conditions are $q_0(s,s)=0$, $q_\lambda^j(s,s)=0$ and $q_l^j(s,s)=0$.

Proof Let us denote by $f\left(t,\tilde{\boldsymbol{\lambda}}_t,\left(\tilde{\boldsymbol{Y}}_t^l\right)_{l=-n,\dots,n},\tilde{\boldsymbol{L}}_t\right)$, the moment generating of claim processes: $\mathbb{E}\left(e^{\boldsymbol{\omega}^\top\tilde{\boldsymbol{L}}_s}\,|\,\mathcal{F}_t\right)$. As $f(.)$ is a conditional expectation, it is also a martingale and therefore $f(.)$ is solution of the following differential equation:

$$
\begin{aligned}
0 &= \partial_t f(.) + i\sum_{l=-n}^{n} b_l\left(\partial_{\tilde{\boldsymbol{Y}}^l} f(.)\right)^\top \tilde{\boldsymbol{Y}}_t^l + \left(\partial_{\tilde{\boldsymbol{\lambda}}} f(.)\right)^\top \times \qquad (18)\\
&\left[\operatorname{diag}\left(\frac{dG(t)}{dt}\right)\left(\tilde{\boldsymbol{\lambda}}_0-\boldsymbol{\theta}\right)+i\sum_{l=-n}^{n} b_l\,(H\odot W_l)\,\tilde{\boldsymbol{Y}}_t^l\right]+\sum_{j=1}^{m}\tilde{\lambda}_t^j\times\\
&\int_0^\infty f\left(t,\tilde{\boldsymbol{\lambda}}_t+\sum_{l=-n}^{n}(H\odot W_l)\,\boldsymbol{e}_j\,z\,,\left(\tilde{\boldsymbol{Y}}_t^l+\boldsymbol{e}_j\,z\right)_{l=-n,\dots,n},\tilde{\boldsymbol{L}}_t+\boldsymbol{e}_j\,z\right)\\
&-f(.)\,\zeta_j(dz)\,.
\end{aligned}
$$

We do the ansatz that $f(.)$ is an exponential affine function of the form

$$
f(.)=\exp\left(q_0(t,s)+\boldsymbol{q}_\lambda(t,s)^\top\tilde{\boldsymbol{\lambda}}_t+\sum_{l=-n}^{n}\boldsymbol{q}_l(t,s)^\top(H\odot W_l)\,\tilde{\boldsymbol{Y}}_t^l+\boldsymbol{\omega}^\top\tilde{\boldsymbol{L}}_t\right).
$$

The mgf is real and therefore $\sum_{l=-n}^{n}\boldsymbol{q}_l(t,s)^\top(H\odot W_l)\,\tilde{\boldsymbol{Y}}_t^l$ must be real. Given that $\tilde{Y}_t^{(j,l)}$ are random and

$$
\sum_{l=-n}^{n}\boldsymbol{q}_l(t,s)^\top(H\odot W_l)\,\tilde{\boldsymbol{Y}}_t^l \;=\; \sum_{l=-n}^{n}\sum_{k=1}^{m}\sum_{j=1}^{m} q_l^k(t,s)\eta_{kj}w_l^{(k,j)}\tilde{Y}_t^{(j,l)}\,,
$$

the sum $\sum_{l=-n}^{n} q_l^k(t,s)w_l^{(k,j)}\tilde{Y}_t^{(j,l)}\in\mathbb{R}$. This necessary implies that $\mathrm{Re}(q_l^k(.))=\mathrm{Re}(q_{-l}^k(.))$ and $-\mathrm{Im}(q_l^k(.))=\mathrm{Im}(q_{-l}^k(.))$. This relation can be checked once that $q_l^k(.)$ is found. The partial derivatives of $f(.)$ with respect to state variables and times are given by

$$
\begin{aligned}
\partial_{\tilde{\boldsymbol{Y}}^l} f(.) &= f(.)\,(H\odot W_l)^\top\,\boldsymbol{q}_l(t,s)\\
\partial_{\tilde{\boldsymbol{\lambda}}} f(.) &= f(.)\boldsymbol{q}_\lambda(t,s)\,,
\end{aligned}
$$

and

$$
\partial_t f(.) \;=\; f(.)\left(\partial_t q_0(t,s)+\partial_t\boldsymbol{q}_\lambda(t,s)^\top\tilde{\boldsymbol{\lambda}}_t+\sum_{l=-n}^{n}\partial_t\boldsymbol{q}_l(t,s)^\top(H\odot W_l)\,\tilde{\boldsymbol{Y}}_t^l\right).
$$

Under the assumption that $f(.)$ is an exponential affine function, the jump term in Equation (18) becomes

$$
\begin{aligned}
&f\left(t,\tilde{\boldsymbol{\lambda}}_t+\sum_{l=-n}^{n}(H\odot W_l)\,\boldsymbol{e}_j\,z,\left(\tilde{\boldsymbol{Y}}_t^l+\boldsymbol{e}_j\,z\right)_{l=-n,\dots,n},\tilde{\boldsymbol{L}}_t+\boldsymbol{e}_j\,z\right)-f(.)=\\
&f(.)\left(e^{\boldsymbol{q}_\lambda(t,s)^\top\sum_{l=-n}^{n}(H\odot W_l)\,\boldsymbol{e}_j\,z+\sum_{l=-n}^{n}\boldsymbol{q}_l(t,s)^\top(H\odot W_l)\,\boldsymbol{e}_j\,z+\boldsymbol{\omega}^\top\boldsymbol{e}_j\,z}-1\right)
\end{aligned}
$$

Combining the previous equations allows us to infer that $q_0(t,s)$, $\boldsymbol{q}_\lambda(t,s)$ and $\boldsymbol{q}_l(t,s)$ satisfy the relation

$$
\begin{aligned}
0 = \partial_t q_0(t,s) + \partial_t \boldsymbol{q}_\lambda(t,s)^\top \tilde{\boldsymbol{\lambda}}_t + \sum_{l=-n}^{n} \partial_t \boldsymbol{q}_l(t,s)^\top \left(H \odot W_l\right) \tilde{\boldsymbol{Y}}_t^l \\
+ i \sum_{l=-n}^{n} b_l \left(\left(H \odot W_l\right)^\top \boldsymbol{q}_l(t,s) \right)^\top \tilde{\boldsymbol{Y}}_t^l + \boldsymbol{q}_\lambda(t,s)^\top \times \\
\left[\operatorname{diag}\left(\frac{dG(t)}{dt}\right) \left(\tilde{\boldsymbol{\lambda}}_0 - \boldsymbol{\theta}\right) + i \sum_{l=-n}^{n} b_l \, \left(H \odot W_l\right) \, \tilde{\boldsymbol{Y}}_t^l \right] + \sum_{j=1}^{m} \tilde{\lambda}_t^j \\
\int_0^\infty e^{\boldsymbol{q}_\lambda(t,s)^\top \sum_{l=-n}^{n} (H\odot W_l)\, \boldsymbol{e}_j \, z + \sum_{l=-n}^{n} \boldsymbol{q}_l(t,s)^\top (H\odot W_l)\, \boldsymbol{e}_j \, z + \boldsymbol{\omega}^\top \boldsymbol{e}_j \, z} \\
-1 \, \zeta_j(dz) \, .
\end{aligned}
$$

Regrouping terms leads allows us to infer the following ODE's

$$
\begin{aligned}
0 &= \partial_t q_0(t,s) + \boldsymbol{q}_\lambda(t,s)^\top \operatorname{diag}\left(\frac{dG(t)}{dt}\right) \left(\tilde{\boldsymbol{\lambda}}_0 - \boldsymbol{\theta}\right) \, , \\
0 &= \int_0^\infty \left[e^{\boldsymbol{q}_\lambda(t,s)^\top \sum_{l=-n}^{n} (H\odot W_l)\, \boldsymbol{e}_j \, z + \sum_{l=-n}^{n} \boldsymbol{q}_l(t,s)^\top (H\odot W_l)\, \boldsymbol{e}_j \, z + \boldsymbol{\omega}^\top \boldsymbol{e}_j \, z} - 1 \right] \zeta_j(dz) \\
&\quad + \partial_t q_{\lambda_j}(t,s) \, , \\
0 &= \sum_{l=-n}^{n} \partial_t \boldsymbol{q}_l(t,s)^\top \left(H \odot W_l\right) + i \sum_{l=-n}^{n} b_l \, \boldsymbol{q}_l(t,s)^\top \left(H \odot W_l\right) \\
&\quad + i \sum_{l=-n}^{n} b_l \, \boldsymbol{q}_\lambda(t,s)^\top \, \left(H \odot W_l\right) \, .
\end{aligned}
$$

The last equation is fulfilled if

$$
\partial_t \boldsymbol{q}_l(t,s) \quad = \quad -i \, b_l \, \left(\boldsymbol{q}_l(t,s) + \, \boldsymbol{q}_\lambda(t,s) \right)
$$

for $j = 1, ..., m$ from which we infer equations (17). Given that $b_l = -b_l$ and $q_\lambda \in \mathbb{R}$, we have well that $\operatorname{Re}(q_l^k(.)) = \operatorname{Re}(q_{-l}^k(.))$ and $-\operatorname{Im}(q_l^k(.)) = \operatorname{Im}(q_{-l}^k(.))$.

end

5 Moment generating function

We establish the moment generating function (mgf) of claim processes by considering the limit of the mgf of the discretized model detailed in the previous section. Firstly, we provide the analytical expression of $\boldsymbol{q}_l(.)$ in the next corollary:

Corollary 5.1. *The function $\boldsymbol{q}_l(t,s) : \left(\mathbb{R}^+\right)^2 \to \mathbb{R}^m$ is equal to*

$$
\boldsymbol{q}_l(t,s) = \int_t^s i \, b_l e^{i \, b_l \, (u-t)} \boldsymbol{q}_\lambda(u,s) \, du \, . \tag{19}
$$

This result is directly proven by checking that the proposed expression of $\boldsymbol{q}_l(.,.)$ is s solution of the ODE's (17),

$$\partial_t \boldsymbol{q}_l(t,s) \;=\; -i\, b_l \underbrace{\left(\int_t^s i\, b_l e^{i\, b_l\,(u-t)} \boldsymbol{q}_\lambda(u,s)\, du\right)}_{\boldsymbol{q}_l(t,s)} -i\, b_l \boldsymbol{q}_\lambda(t,s)\,.$$

As the process $\boldsymbol{L}_t$ is the limit of $\tilde{\boldsymbol{L}}_t$ when $n \to \infty$, we can find its moment generating function by considering the limit of Equation (16). The next proposition presents the main result of this article.

Proposition 5.2. *The moment generating function of the claim process,* $\mathbb{E}\left(e^{\boldsymbol{\omega}^\top \boldsymbol{L}_s} \,|\, \mathcal{F}_t\right)$ *for* $\boldsymbol{\omega} \in \mathbb{C}^{m-}$*, is*

$$\begin{aligned} \mathbb{E}\left(e^{\boldsymbol{\omega}^\top \boldsymbol{L}_s} \,|\, \mathcal{F}_t\right) \;&=\; \exp\left(q_0(t,s) + \boldsymbol{q}_\lambda(t,s)^\top \boldsymbol{\lambda}_t + \boldsymbol{\omega}^\top \boldsymbol{L}_t\right) \\ &\quad \times \exp\left(\int_0^t \int_t^s \boldsymbol{q}_\lambda(u,s)^\top \left(H \odot \frac{dG(u-v)}{du}\right) du\, d\boldsymbol{L}_v\right)\,, \end{aligned} \tag{20}$$

where $q_0(t,s) : (\mathbb{R}^+)^2 \to \mathbb{R}$ *and* $\boldsymbol{q}_\lambda(t,s) : (\mathbb{R}^+)^2 \to \mathbb{R}^m$ *are solutions of the system of ordinary equations s (ODE's)*

$$\begin{aligned} &\partial_t q_0(t,s) = -\boldsymbol{q}_\lambda(t,s)^\top diag\left(\frac{dG(t)}{dt}\right)(\boldsymbol{\lambda}_0 - \boldsymbol{\theta}) \\ &\partial_t q_\lambda^j(t,s) = -\int_0^\infty e^{\left(\boldsymbol{q}_\lambda(t,s)^\top H + \int_t^s \boldsymbol{q}_\lambda(u,s)^\top \left(H \odot \frac{dG(u-t)}{du}\right) du + \boldsymbol{\omega}^\top\right)\boldsymbol{e}_j\, z} - 1\, \zeta_j(dz) \end{aligned} \tag{21}$$

for $j = 1, ..., m$*. The terminal conditions are* $q_0(s,s) = 0$ *and* $q_\lambda^j(s,s) = 0$*.*

Proof By construction $\mathbb{E}\left(e^{\boldsymbol{\omega}^\top \boldsymbol{L}_s} \,|\, \mathcal{F}_t\right) = \lim_{n\to\infty} \mathbb{E}\left(e^{\boldsymbol{\omega}^\top \tilde{\boldsymbol{L}}_s} \,|\, \mathcal{F}_t\right)$. We can then develop the sum present in Equation (16) as

$$\lim_{n\to\infty} \sum_{l=-n}^{n} \boldsymbol{q}_l(t,s)^\top (H \odot W_l)\, \tilde{\boldsymbol{Y}}_t^l = \sum_{k=1}^{m}\sum_{j=1}^{m} \lim_{n\to\infty} \sum_{l=-n}^{n} q_l^k(t,s) \eta_{kj} w_l^{(k,j)} \tilde{Y}_t^{(j,l)}\,.$$

If we remember the definition of $w_l^{(k,j)}$ and $Y_t^{(j,\xi)}$, we obtain after a change of order of integration that

$$\begin{aligned} \lim_{n\to\infty} \sum_{l=-n}^{n} q_l^k(t,s) \eta_{kj} w_l^{(k,j)} \tilde{Y}_t^{(j,l)} \;&=\; \int_{\mathbb{R}} q_\xi^k(t,s) Y_t^{(j,\xi)} \eta_{kj} \nu_{kj}(d\xi) \\ &=\; \int_0^t \int_{\mathbb{R}} q_\xi^k(t,s) e^{i(t-v)\xi} \eta_{kj} \nu_{kj}(d\xi)\, dL_v^j\,. \end{aligned} \tag{22}$$

From the previous corollary, the function $q_l^k(t,s)$ is given by

$$q_l^k(t,s) = \int_t^s i\, b_l e^{i\, b_l\,(u-t)} q_\lambda^k(u,s)\, du$$

Therefore, the limit of this function when the size of the partition tends to infinity is

$$\lim_{n\to\infty} q_l^k(t,s) \;\; = q_\xi^k(t,s) = \;\; \int_t^s i\,\xi e^{i\,\xi\,(u-t)} q_\lambda^k(u,s)\,du\,,$$

for some $\xi \in \mathbb{R}$. Injecting this expression in the integrand of Equation (22) leads to

$$\begin{aligned}
&\int_{\mathbb{R}} q_\xi^k(t,s) e^{i(t-v)\xi} \eta_{kj} \nu_{kj}(d\xi) \\
&\quad = \int_{\mathbb{R}} \int_t^s i\,\xi e^{i\,\xi\,(u-t)} q_\lambda^k(u,s)\,du\, e^{i(t-v)\xi} \eta_{kj} \nu_{kj}(d\xi) \\
&\quad = \int_t^s q_\lambda^k(u,s) \eta_{kj} \int_{\mathbb{R}} i\,\xi e^{i\,\xi\,(u-v)} \nu_{kj}(d\xi)\,du\,.
\end{aligned} \tag{23}$$

As $g_{kj}(.)$ admits a spectral representation, its derivative is also equal to

$$\frac{dg_{kj}(u-v)}{du} = \int_{\mathbb{R}} i\,\xi\, e^{i\,\xi\,(u-v)} \nu_{kj}(d\xi)\,. \tag{24}$$

Combining Equations (23) and (24) gives us

$$\int_{\mathbb{R}} q_\xi^k(t,s) e^{i(t-v)\xi} \eta_{kj} \nu_{kj}(d\xi) = \int_t^s q_\lambda^k(u,s) \eta_{kj} \frac{dg_{kj}(u-v)}{du}\,du\,.$$

It is then possible to develop Equation (22) as follows

$$\lim_{n\to\infty} \sum_{l=-n}^{n} q_l^k(t,s) \eta_{kj} w_l^{(k,j)} \tilde{Y}_t^{(j,l)} \;\; = \;\; \int_0^t \int_t^s q_\lambda^k(u,s) \eta_{kj} \frac{dg_{kj}(u-v)}{du}\,du\,dL_v^j\,,$$

and the limit of $\sum_{l=-n}^{n} \boldsymbol{q}_l(t,s)^\top (H \odot W_l)\, \tilde{\boldsymbol{Y}}_t^l$ when $n \to \infty$ can then be rewritten under the matrix form:

$$\begin{aligned}
&\sum_{k=1}^{m} \sum_{j=1}^{m} \lim_{n\to\infty} \sum_{l=-n}^{n} q_l^k(t,s) \eta_{kj} w_l^{(k,j)} \tilde{Y}_t^{(j,l)} \\
&\quad = \int_0^t \int_t^s \sum_{k=1}^{m} q_\lambda^k(u,s) \sum_{j=1}^{m} \eta_{kj} \frac{dg_{kj}(u-v)}{du}\,du dL_v^j \\
&\quad = \int_0^t \int_t^s \boldsymbol{q}_\lambda(u,s)^\top \left(H \odot \frac{dG(u-v)}{du} \right) du\, d\boldsymbol{L}_v\,.
\end{aligned}$$

The expression (20) immediately follows from this last result. The derivative of q_λ^j with respect to time is

$$\partial_t q_\lambda^j(t,s) = -\lim_{n\to\infty} \int_0^\infty e^{\boldsymbol{q}_\lambda(t,s)^\top \sum_{l=-n}^{n} H \odot W_l\, \boldsymbol{e}_j\, z + \sum_{l=-n}^{n} \boldsymbol{q}_l(t,s)^\top W_l\, \boldsymbol{e}_j\, z + \boldsymbol{\omega}^\top \boldsymbol{e}_j\, z} - 1\, \zeta_j(dz)\,.$$

Using a similar reasoning as for previous developments, we infer that

$$\begin{aligned}
\lim_{n\to\infty}\sum_{l=-n}^{n}\boldsymbol{q}_l(t,s)^\top\left(H\odot W_l\right)\boldsymbol{e}_j &= \lim_{n\to\infty}\sum_{k=1}^{m}\sum_{l=-n}^{n} q_l^k(t,s)\eta_{kj}w_l^{(k,j)} \\
&= \sum_{k=1}^{m}\int_{\mathbb{R}}\int_t^s i\,\xi e^{i\,\xi\,(u-t)}q_\lambda^k(u,s)\,du\,\eta_{kj}\nu_{kj}(d\xi) \\
&= \sum_{k=1}^{m}\int_t^s q_\lambda^k(u,s)\eta_{kj}\int_{\mathbb{R}} i\,\xi e^{i\,\xi\,(u-t)}\,\nu_{kj}(d\xi)\,du \\
&= \sum_{k=1}^{m}\int_t^s q_\lambda^k(u,s)\eta_{kj}\frac{dg_{kj}(u-t)}{du}\,du \\
&= \int_t^s \boldsymbol{q}_\lambda(u,s)^\top\left(H\odot\frac{dG(u-t)}{du}\right)du\,\boldsymbol{e}_j\,.
\end{aligned}$$

As $\nu_{kj}(.)$ is a measure of probability, we have that $\lim_{n\to\infty}\sum_{l=-n}^{n} w_l^{(k,j)} = 1$ and

$$\begin{aligned}
\lim_{n\to\infty}\sum_{l=-n}^{n}\boldsymbol{q}_\lambda(t,s)^\top\left(H\odot W_l\right)\boldsymbol{e}_j &= \lim_{n\to\infty}\sum_{l=-n}^{n}\sum_{k=1}^{m} q_\lambda^k(t,s)\eta_{kj}\,w_l^{(k,j)} \\
&= \sum_{k=1}^{m} q_\lambda^k(t,s)\eta_{kj} \\
&= \boldsymbol{q}_\lambda(t,s)^\top H\boldsymbol{e}_j\,.
\end{aligned}$$

These two last expressions leads to Equation (21).

end

This result may be used to estimate the value of each and all business lines by indifference pricing. For instance, let us denote by $U(x) = -\frac{1}{\gamma}e^{-\gamma x}$, the insurer's utility function. The initial surplus is S_0 whereas $P(s)$ is the total premium for covering the m risks up to time s. The surplus at time s is defined as

$$S_s = S_0 + P(s) - \sum_{k=1}^{m} L_s^k$$

The premium is $P(s)$ estimated such that the company is indifferent to or not to ensure the risks, i.e.

$$U(S_0) = \mathbb{E}\left(U\left(S_0 + P(s) - \sum_{k=1}^{m} L_s^k\right)\mid\mathcal{F}_0\right).$$

If we denote by $\boldsymbol{e}$ the vector $(1,1,...,1)^\top$. The indifference premium for insuring the m risks is given

$$P(s) = \frac{1}{\gamma}\ln\mathbb{E}\left(e^{\gamma\,\boldsymbol{e}^\top\boldsymbol{L}_s}\mid\mathcal{F}_0\right),$$

where $\mathbb{E}\left(e^{\gamma \boldsymbol{e}^\top \boldsymbol{L}_s} \,|\, \mathcal{F}_0\right)$ can be calculated with Proposition 5.2 in which $\boldsymbol{\omega} = \gamma \boldsymbol{e}$. Another interesting use of the mgf is the computation of joint and marginal probability density functions of aggregated claims. This point is developed in the case study concluding the article. But before, we focus on a sub-category of Cauchy memory kernels for which $(\boldsymbol{\lambda}_t, \boldsymbol{L}_t)_{t\geq 0}$ is Markov.

6 Cauchy memory kernels

As already mentioned in Section 2, the process $(\boldsymbol{\lambda}_t, \boldsymbol{L}_t)_{t\geq 0}$ is Markov with Cauchy memory kernels such that $\alpha_{kj} = \alpha_k$ for $k = 1, ..., m$. This model is extensively studied in the literature. For such kernels, the moment generating function is obtained by taking advantages of the Markov properties of $(\boldsymbol{\lambda}_t, \boldsymbol{L}_t)_{t\geq 0}$ and Itô's calculus. Nevertheless, the expression of the mgf obtained by this way differs from the one presented in Proposition 5.2. The aim of this section is to reconciliate these results. We start by briefly reviewing the expression of the mgf inferred by Itô's calculus. If $g_{kj}(t) = e^{-\alpha_k t}$ with $\alpha_k \in \mathbb{R}^+$ for $k, j \in \{1, ..., m\}$, the multivariate process $(\boldsymbol{\lambda}_t, \boldsymbol{L}_t)_{t\geq 0}$ is Markov and

$$d\boldsymbol{\lambda}_t = \operatorname{diag}(\boldsymbol{\alpha})\,(\boldsymbol{\theta} - \boldsymbol{\lambda}_t)\,dt + H\,d\boldsymbol{L}_t\,,$$

where $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_m)^\top$. The matrix $G(t) = (e^{-\alpha_k t})_{k,j=1,...,m}$ has identical columns and

$$\frac{dG(t)}{dt} = \left(-\alpha_k e^{-\alpha_k t}\right)_{k,j=1,...,m} = -\operatorname{diag}(\boldsymbol{\alpha})\,G(t)$$

We write $\operatorname{diag}(\boldsymbol{\alpha} e^{-\boldsymbol{\alpha} t}) = \operatorname{diag}\left(\left(\alpha_k e^{-\alpha_k t}\right)_{k=1...m}\right)$ to lighten notations. According to Proposition 5.2, the mgf of cumulated claims process is given by

$$\begin{aligned}\mathbb{E}\left(e^{\boldsymbol{\omega}^\top \boldsymbol{L}_s} \,|\, \mathcal{F}_t\right) &= \exp\left(q_0(t,s) + \boldsymbol{q}_\lambda(t,s)^\top \boldsymbol{\lambda}_t + \omega^\top \boldsymbol{L}_t\right) \\ &\times \exp\left(-\int_0^t \int_t^s \boldsymbol{q}_\lambda(u,s)^\top \left(H \odot \left(\operatorname{diag}(\boldsymbol{\alpha})\, G(u-v)\right)\right) du\, d\boldsymbol{L}_v\right)\,,\end{aligned} \tag{25}$$

where $q_0(t,s) : (\mathbb{R}^+)^2 \to \mathbb{R}$, $\boldsymbol{q}_\lambda(t,s) : (\mathbb{R}^+)^2 \to \mathbb{R}^m$ are solutions of the system of ODE's

$$\begin{cases}\partial_t q_0(t,s) = \boldsymbol{q}_\lambda(t,s)^\top \operatorname{diag}(\boldsymbol{\alpha} e^{-\boldsymbol{\alpha} t})\,(\boldsymbol{\lambda}_0 - \boldsymbol{\theta}) \\ \partial_t q_\lambda^j(t,s) = -\int_0^\infty e^{\boldsymbol{q}_\lambda(t,s)^\top H \boldsymbol{e}_j\, z - \int_t^s \boldsymbol{q}_\lambda(u,s)^\top (H \odot (\operatorname{diag}(\boldsymbol{\alpha})\, G(u-t)))\, du\, \boldsymbol{e}_j z + \boldsymbol{\omega}^\top \boldsymbol{e}_j\, z} - 1\, \zeta_j(dz)\end{cases} \tag{26}$$

for $j = 1, ..., m$. The terminal conditions are $q_0(s,s) = 0$ and $q_\lambda^j(s,s) = 0$. This formulation of the mgf is not standard in the literature. The most common expression for the mgf is obtained with the Itô's lemma and recalled in the next proposition.

Proposition 6.1. *When the memory kernels are exponential decreasing with $\alpha_{kj} = \alpha_k$ for $k = 1, ..., m$, the moment generating function of cumulated claim processes for $\boldsymbol{\omega} \in \mathbb{C}^{m-}$ is:*

$$\mathbb{E}\left(e^{\boldsymbol{\omega}^\top \boldsymbol{L}_s} \,|\, \mathcal{F}_t\right) = \exp\left(p_0(t,s) + \boldsymbol{p}_\lambda(t,s)^\top \boldsymbol{\lambda}_t + \boldsymbol{\omega}^\top \boldsymbol{L}_t\right)\,, \tag{27}$$

where $p_0(t,s) : (\mathbb{R}^+)^2 \to \mathbb{R}$, $\boldsymbol{p}_\lambda(t,s) : (\mathbb{R}^+)^2 \to \mathbb{R}^m$ *are solutions of the system of ordinary equations s (ODE's)*

$$\begin{cases} \partial_t p_0(t,s) = -\boldsymbol{p}_\lambda(t,s)^\top \, diag\,(\boldsymbol{\alpha})\,\boldsymbol{\theta} \\ \partial_t p_\lambda^j(t,s) = \alpha_j p_\lambda^j(t,s) - \int_0^\infty e^{\boldsymbol{p}_\lambda(t,s)^\top H \boldsymbol{e}_j\, z + \boldsymbol{\omega}^\top \boldsymbol{e}_j\, z} - 1\, \zeta_j(dz) \end{cases} \tag{28}$$

for $j = 1, ..., m$*. The terminal conditions are* $p_0(s,s) = 0$ *and* $p_\lambda^j(s,s) = 0$*.*

Proof Let us momentaneously denote by $f(t, \boldsymbol{\lambda}_t, \boldsymbol{L}_t)$ the mgf $\mathbb{E}\left(e^{\boldsymbol{\omega}^\top \boldsymbol{L}_s} \,|\, \mathcal{F}_t\right)$. Since $f(.)$ is a conditional expectation, it is also a martingale. This is then solution of the following Itô equation:

$$\begin{aligned} 0 &= \partial_t f(.) + \partial f_{\boldsymbol{\lambda}}(.)^\top \mathrm{diag}(\boldsymbol{\alpha})(\boldsymbol{\theta} - \boldsymbol{\lambda}_t) \\ &\quad + \sum_{j=1}^m \lambda_t^j \int_0^\infty f(t, \boldsymbol{\lambda}_t + H\,\boldsymbol{e}_j\, z, \boldsymbol{L}_t + \boldsymbol{e}_j z) - f(.)\,\zeta_j(dz)\,. \end{aligned}$$

If we consider the ansatz $f\,(t, \boldsymbol{\lambda}_t, \boldsymbol{L}_t) = \exp(p_0(t,s) + \boldsymbol{p}_\lambda(t,s)^\top \boldsymbol{\lambda}_t + \boldsymbol{\omega}^\top \boldsymbol{L}_t)$ then

$$\begin{aligned} 0 &= \partial_t p_0(t,s) + \partial_t \boldsymbol{p}_\lambda(t,s)^\top \boldsymbol{\lambda}_t + \boldsymbol{p}_\lambda(t,s)^\top \mathrm{diag}(\boldsymbol{\alpha})(\boldsymbol{\theta} - \boldsymbol{\lambda}_t) \\ &\quad + \sum_{j=1}^m \lambda_t^j \int_0^\infty e^{\boldsymbol{p}_\lambda(t,s)^\top H\, \boldsymbol{e}_j\, z + \boldsymbol{\omega}^\top \boldsymbol{e}_j\, z} - 1\, \zeta_j(dz)\ . \end{aligned}$$

Regrouping terms leads to the result.

end

The next proposition reconciliates propositions 6.1 and 5.2.

Proposition 6.2. *When memory kernels are exponential decreasing with* $\alpha_{kj} = \alpha_k$ *for* $k = 1, ..., m$*, the functions* $p_0(t,s)$*,* $\boldsymbol{p}_\lambda(t,s)$ *defined by Equations (28) are related to functions* $q_0(t,s)$*,* $\boldsymbol{q}_\lambda(t,s)$ *of Equations (26) by the relations:*

$$\begin{cases} p_0(t,s) & = q_0(t,s) + \int_t^s \boldsymbol{q}_\lambda(u,s)^\top \, diag(e^{-\boldsymbol{\alpha}(u-t)})\,du \\ & \quad \times (diag(\boldsymbol{\alpha})\,\boldsymbol{\theta} + diag(\boldsymbol{\alpha} e^{-\boldsymbol{\alpha} t})\,(\boldsymbol{\lambda}_0 - \boldsymbol{\theta})) \\ \boldsymbol{p}_\lambda(t,s)^\top & = \boldsymbol{q}_\lambda(t,s)^\top - \int_t^s \boldsymbol{q}_\lambda(u,s)^\top\,(H \odot (diag(\boldsymbol{\alpha})\,G(u-t)))\,H^{-1}\,du \\ & = \boldsymbol{q}_\lambda(t,s)^\top - \int_t^s \boldsymbol{q}_\lambda(u,s)^\top diag(\boldsymbol{\alpha} e^{-\boldsymbol{\alpha}(u-t)})\,du \end{cases} \tag{29}$$

and

Proof Let us denote by $E = (1)_{k,j=1...m}$ the $m \times m$ matrix of ones. We start the proof by deriving the second equation of (29) with respect to time,

$$\begin{aligned} \partial_t \boldsymbol{p}_\lambda(t,s)^\top &= \partial_t \boldsymbol{q}_\lambda(t,s)^\top + \boldsymbol{q}_\lambda(t,s)^\top\,(H \odot (\mathrm{diag}(\boldsymbol{\alpha})E))\,H^{-1} \\ &\quad - \int_t^s \boldsymbol{q}_\lambda(u,s)^\top \left(H \odot \left(\mathrm{diag}(\boldsymbol{\alpha})^2\, G(u-t)\right)\right) H^{-1}\,du\,. \end{aligned} \tag{30}$$

Basic matrix algebra allows us to infer two interesting relations:

$$(H \odot (\mathrm{diag}(\boldsymbol{\alpha})E))\, H^{-1} = \mathrm{diag}(\boldsymbol{\alpha})\,,$$

and

$$\begin{aligned}\left(H \odot \left(\mathrm{diag}(\boldsymbol{\alpha})^2\, G(u-t)\right)\right) H^{-1} = \\ \mathrm{diag}(\boldsymbol{\alpha})\, (H \odot (\mathrm{diag}(\boldsymbol{\alpha})\, G(u-t)))\, H^{-1}\,.\end{aligned}$$

Equation (18) becomes then

$$\begin{aligned}\partial_t \boldsymbol{p}_\lambda(t,s)^\top &= \partial_t \boldsymbol{q}_\lambda(t,s)^\top + \boldsymbol{q}_\lambda(t,s)^\top \mathrm{diag}(\boldsymbol{\alpha}) \\ &\quad - \int_t^s \boldsymbol{q}_\lambda(u,s)^\top \mathrm{diag}(\boldsymbol{\alpha})\, (H \odot (\mathrm{diag}(\boldsymbol{\alpha})\, G(u-t)))\, H^{-1}\, du\end{aligned}$$

from which we infer that

$$\partial_t p_\lambda^j(t,s) \;=\; \partial_t q_\lambda^j(t,s) + \alpha_j \boldsymbol{p}_\lambda(t,s)^\top \boldsymbol{e}_j \tag{31}$$

By definition of $q_\lambda^j(t,s)$, we have

$$\partial_t q_\lambda^j(t,s) = -\int_0^\infty e^{\boldsymbol{p}_\lambda(t,s)^\top H \boldsymbol{e}_j\, z + \boldsymbol{\omega}^\top \boldsymbol{e}_j\, z} - 1\, \zeta_j(dz)$$

and we check that $\partial_t p_\lambda^j(t,s)$ satisfies the ODE's in Equation (28). Let us recall that $\boldsymbol{q}_\lambda(s,s) = 0$. If we denote by $\mathrm{diag}(\boldsymbol{\alpha}^2 e^{-\boldsymbol{\alpha} t})$ the $m \times m$ diagonal matrix $\mathrm{diag}\left((\alpha_k^2 e^{-\alpha_k t})_{k=1\ldots m}\right)$, the derivative of $p_0(t,s)$ with respect to t is given by

$$\begin{aligned}\partial_t p_0(t,s) = \partial_t q_0(t,s) - \int_t^s &\boldsymbol{q}_\lambda(u,s)^\top \mathrm{diag}(e^{-\boldsymbol{\alpha}(u-t)})\, du \\ &\times \left(\mathrm{diag}\left(\boldsymbol{\alpha}^2 e^{-\boldsymbol{\alpha} t}\right)(\boldsymbol{\lambda}_0 - \boldsymbol{\theta})\right) \\ - \boldsymbol{q}_\lambda(t,s)^\top &\left(\mathrm{diag}(\boldsymbol{\alpha})\, \boldsymbol{\theta} + \mathrm{diag}(\boldsymbol{\alpha} e^{-\boldsymbol{\alpha} t})\, (\boldsymbol{\lambda}_0 - \boldsymbol{\theta})\right) \\ + \int_t^s &\boldsymbol{q}_\lambda(u,s)^\top \mathrm{diag}(\boldsymbol{\alpha} e^{-\boldsymbol{\alpha}(u-t)})\, du \\ &\times \left(\mathrm{diag}(\boldsymbol{\alpha})\, \boldsymbol{\theta} + \mathrm{diag}(\boldsymbol{\alpha} e^{-\boldsymbol{\alpha} t})\, (\boldsymbol{\lambda}_0 - \boldsymbol{\theta})\right)\,.\end{aligned}$$

From the definition of $q_0(t,s)$, its derivative is equal to

$$\partial_t q_0(t,s) = \boldsymbol{q}_\lambda(t,s)^\top \mathrm{diag}\left(\boldsymbol{\alpha} e^{-\boldsymbol{\alpha} t}\right)(\boldsymbol{\lambda}_0 - \boldsymbol{\theta})$$

and therefore $\partial_t p_0(t,s)$ satisfies the ODE's in Equation (28):

$$\partial_t p_0(t,s) = -\underbrace{\left[\boldsymbol{q}_\lambda(t,s)^\top - \int_t^s \boldsymbol{q}_\lambda(u,s)^\top \mathrm{diag}(\boldsymbol{\alpha} e^{-\boldsymbol{\alpha}(u-t)})\, du\right]}_{\boldsymbol{p}_\lambda(t,s)^\top} \mathrm{diag}(\boldsymbol{\alpha})\, \boldsymbol{\theta}\,.$$

end

To conclude this section, we draw the attention of the reader on the fact that proposition 6.1 holds only if $\alpha_{kj} = \alpha_k$ for $k = 1, ..., m$. If this condition is not fulfilled the process is not Markov anymore and the mgf must be evaluated with propositions 4.1 or 5.2. Before explaining how we can retrieve the joint density function from the mgf, we briefly review the method for fitting parameters of a multivariate Hawkes process.

7 Estimation

Multivariate self-excited processes are estimated by log-likelihood maximization. Before presenting an illustration, we briefly recall the expression of this log-likelihood adapted to our framework. The time interval of observations is denoted by $[0, \mathcal{T}]$ whereas the jump times of L_t^k are denoted by τ_p^k for $k = 1, ..., m$. If the number of jumps of $L_{\mathcal{T}}^k$ is noted n_k and under the assumption that $\lambda_0^k = \theta_k$, the sample path of intensities is calculated as follows:

$$\lambda_t^k = \theta_k + \sum_{j=1}^{m} \sum_{l=1}^{N_t^k} \eta_{kj} g_{kj}(t - \tau_l^j) J_l^j \quad , \; k = 1, ..., m.$$

From e.g. Embrechts et al. (2001), we know that the log-likelihood is equal to

$$\ln \mathcal{L} \;\; = \;\; -\sum_{k=1}^{m} \int_0^{\mathcal{T}} \lambda_s^k ds + \sum_{k=1}^{m} \sum_{j=1}^{n_k} \log \left(\lambda_{\tau_j}^k \right) \, .$$

To benchmark our model, we also consider a Poisson process with constant intensities $(\lambda_k)_{k=1,...,m}$. In this case, the log-likelihood simply becomes:

$$\ln \mathcal{L} = -\sum_{k=1}^{m} \lambda_k \mathcal{T} + \sum_{k=1}^{m} n_k \log \left(\lambda_k \right) \, .$$

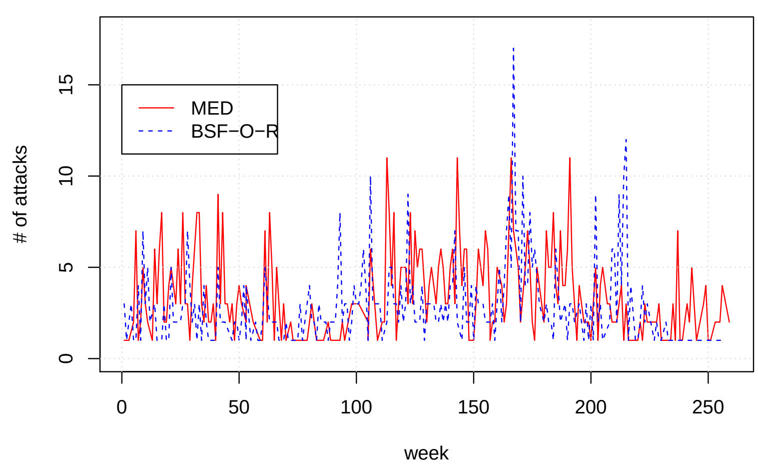

Figure 2: Weekly numbers of cyber-attacks in the MED and BSF-O-R sectors from 2014 to 2018.

We illustrate this procedure with the dataset from the Privacy Rights Clearinghouse[1] (PRC) that contains information about 8871 data breaches in USA since 2005. We focus on

[1] https://www.privacyrights.org/data-breaches

cyber-attacks by outside party and infections by malware from the 1/12014 to 31/12/2018. We consider two categories of sectors targeted by these attacks. The first one regroups healthcare, medical providers or medical Insurance services (MED). The second one gathers financial businesses (BSF), other businesses (BSO) and retail businesses (BSR). This dataset is analysed by Bessy-Roland et al. (2020) who provide evidence of mutual excitations between sectors and attacks. Over the considered period, 762 and 624 events were respectively reported for MED and BSF-O-R sectors. The evolution of weekly numbers of cyber-attacks between 2014 and 2018 is plotted in Figure 2. The dataset contains dates of attacks but do not provide information about the exact timing and about costs caused by them. As several events are reported on same days, we consider that claim sizes correspond to the number of attacks recorded for one day. Figure 3 shows the empirical distributions of the number of attacks recorded during the same day.

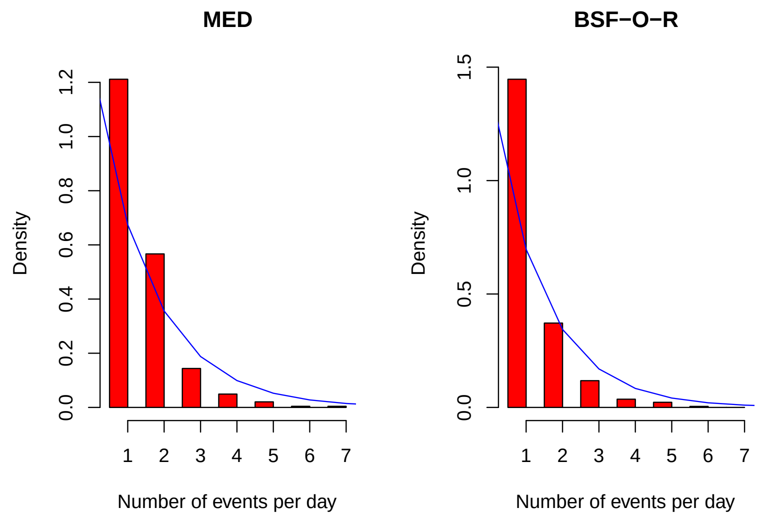

Figure 3: Empirical distributions of numbers of attacks per sector, reported on the same day.

Table 1 presents the results of the estimation procedure for the four memory kernels considered in this article. Except for the Gauss kernel, all other models achieve a similar goodness of fit in term of log-likelihoods. The MED and BSF-O-R sectors are respectively indexed by $k = 1$ and 2. The intensities reverts to values in the interval $[31.10\,, 44.23]$ depending upon the kernel. Whatever the model, self-excitation (measured by η_{11} and η_{22}) prevails on mutual contagion (η_{12} and η_{21}). The log-likelihood of a bivariate Poisson process is equal to 3276.43 and is clearly lower than the ones computed with our models. This emphasizes that considering self-excitation improves the quality of the fit. Notice that the corresponding Poisson intensities, $\lambda_1 = 94.60$ and $\lambda_2 = 91.00$ are twicer larger than mean reversion levels θ_1 and θ_2. Figure 4 compares the Logistic, Laplace and Cauchy memory functions $g_{kj}(.)$ computed with estimated parameters. In the next section, we use the parameter estimates obtained in this section to compute the joint and marginal distributions of the cumulated number of cyber-attacks.

	Kernel			
	Laplace	Gauss	Logistic	Cauchy
θ_1	36.2924	44.2328	40.9028	37.8796
θ_2	31.1032	39.8425	30.6998	39.1864
α_{11}	0.0456	0.0094	10.9251	13.4944
α_{12}	0.0816	1.4087	24.2666	21.0842
α_{21}	0.0103	0.4791	15.4625	23.3067
α_{22}	0.0520	0.0004	7.3451	13.6574
η_{11}	5.7489	3.1507	5.2305	5.3892
η_{12}	0.1209	0.0325	0.3872	0.0475
η_{21}	0.0764	0.6828	0.5021	0.1365
η_{22}	5.8786	8.5866	4.2646	5.8997
$\ln \mathcal{L}$	3321.72	3315.59	3321.12	3321.21

Table 1: Parameter estimates and log-likelihoods ($\ln \mathcal{L}$).

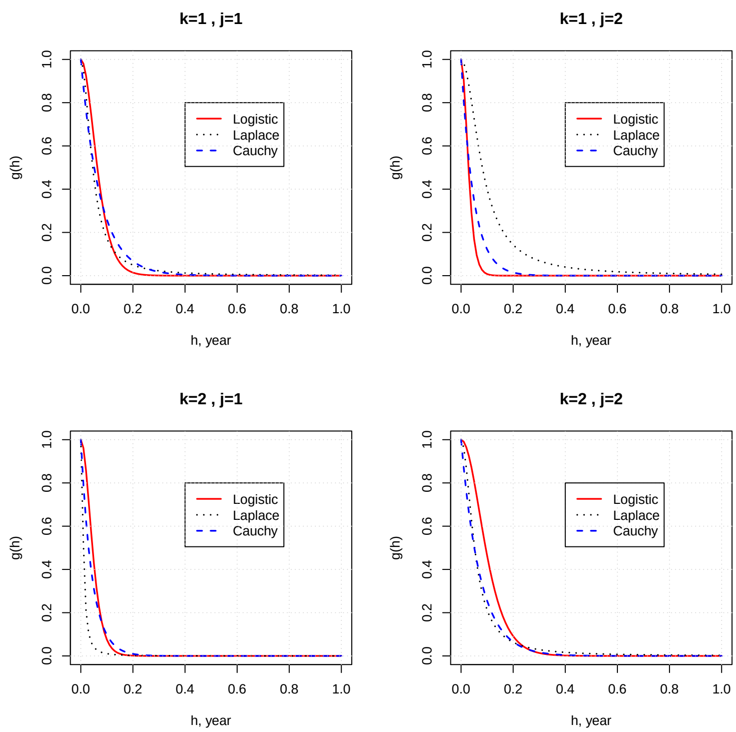

Figure 4: Comparison of estimated Laplace, logistic and Cauchy memory kernels.

8 Probability density functions by fast Fourier's transform

We have found the moment generating function of a multivariate claim process with mutual excitation. We exploit this result in this paragraph to compute the joint or the marginal probability density functions (pdf's) of this process. Our approach is based on a discrete fast Fourier's transform (DFFT). It roughly consists in inverting the characteristic function of the claim process, that is the mgf appraised on the imaginary axis.

For numerical reasons, we focus on the discrete version of the claim process, introduced in Section 4. Its characteristic function, denoted by $\Upsilon_{t,s}(i\boldsymbol{\omega}) = \mathbb{E}\left(e^{i\boldsymbol{\omega}^{\top}\tilde{\boldsymbol{L}}_s} \mid \mathcal{F}_t\right)$, is the inverse Fourier transform of the multivariate probability density function (pdf) $f_{t,s}(\boldsymbol{x})$ of $\tilde{\boldsymbol{L}}_s|\mathcal{F}_t$. Therefore, this density can be retrieved by computing the following multiple integrals (the Fourier transform of $\Upsilon_{t,s}(.)$):

$$\begin{aligned} f_{t,s}(\boldsymbol{x}) &= \frac{1}{(2\pi)^m}\mathcal{F}[\Upsilon_{t,s}(i\boldsymbol{\omega})](\boldsymbol{x}) \\ &= \frac{1}{(2\pi)^m}\int_{-\infty}^{+\infty}...\int_{-\infty}^{+\infty}\Upsilon_{t,s}(i\boldsymbol{\omega})e^{-i\boldsymbol{\omega}^{\top}\boldsymbol{x}}d\omega_1...d\omega_m\,. \end{aligned} \tag{32}$$

The calculation of this multiple integral is computationally intensive but it may be speed up if we discretize it in a compatible expression with the DFFT, as presented in the next proposition.

Proposition 8.1. *Let M be the number of steps used in the Discrete Fast Fourier Transform (DFFT) and $\Delta_x = \frac{x_{max}}{M}$, be the step of discretization. Let us denote $\Delta_\omega = \frac{2\pi}{M\,\Delta_x}$ and*

$$\boldsymbol{\omega}_{j_1,j_2,...,j_m} = \left(-\frac{M}{2}\Delta_\omega + (j_1-1)\Delta_\omega, -\frac{M}{2}\Delta_\omega + (j_2-1)\Delta_\omega, ..., -\frac{M}{2}\Delta_\omega + (j_m-1)\Delta_\omega\right)$$

for $j_1, ..., j_m \in \{1, ..., M+1\}$. The values of $f_{t,s}(.)$ at points

$$\boldsymbol{x}_{k_1,k_2,...,k_m} = ((k_1-1)\Delta_x\,,\,(k_2-1)\Delta_x\,,\,...\,,(k_m-1)\Delta_x)$$

for $k_1, ..., k_m \in \{1, ..., M\}$ are approached by

$$\begin{aligned} f_{t,s}(\boldsymbol{x}_{k_1,...,k_m}) &= (\Delta_\omega)^m\sum_{j_1=1}^{M+1}\sum_{j_2=1}^{M+1}\cdots\sum_{j_m=1}^{M+1}\delta_{j_1,j_2,...,j_m}\Upsilon_{t,s}\left(i(\boldsymbol{\omega}_{j_1,j_2,...,j_m})\right) \\ &\quad\times\exp\left(i\sum_{l=1}^{m}((k_l-1)\pi)\right)\times\exp\left(-i\sum_{l=1}^{m}(k_l-1)(j_l-1)\frac{2\pi}{M}\right) \end{aligned} \tag{33}$$

where$\Upsilon_{t,s}(i\omega) = \mathbb{E}\left(e^{i\,\omega\tilde{L}_s} \mid \mathcal{F}_t\right)$ *is defined by equation (16) and*

$$\delta_{j_1,j_2,...,j_m} = \left(\frac{1}{2}\right)^{(1_{\{j_1=1\}}+...+1_{\{j_m=1\}})}(1-1_{\{j_1\neq1,...,j_m\neq1\}}) + 1_{\{j_1\neq1,...,j_m\neq1\}}\,.$$

Proof. Using a trapezoidal approximation of integral (32), leads to the following discrete approximation

$$\int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \Upsilon_{t,s}(i\boldsymbol{\omega})e^{-i\boldsymbol{\omega}^\top \boldsymbol{x}} d\omega_1 ... d\omega_N =$$
$$\sum_{j_1=1}^{M+1}\sum_{j_2=1}^{M+1}\dots\sum_{j_m=1}^{M+1} \delta_{j_1,j_2,\dots,j_m} \Upsilon_{t,s}\left(i(\boldsymbol{\omega}_{j_1,j_2,\dots,j_m})\right) e^{-i\,\boldsymbol{\omega}^\top_{j_1,j_2,\dots,j_m}\boldsymbol{x}_{k_1,k_2,\dots,k_m}} \left(\Delta_\omega\right)^m .$$

and given that $\Delta_x\Delta_\omega = \frac{2\pi}{M}$, the scalar product $\boldsymbol{\omega}^\top_{j_1,j_2,\dots,j_m}\boldsymbol{x}_{k_1,k_2,\dots,k_m}$ is rewritten as follows:

$$\begin{aligned}
&\boldsymbol{\omega}^\top_{j_1,j_2,\dots,j_m}\boldsymbol{x}_{k_1,k_2,\dots,k_m} \\
&= \sum_{l=1}^{m}\left(-\frac{M}{2}(k_l-1)\Delta_x\Delta_\omega + (k_l-1)(j_l-1)\Delta_x\Delta_\omega\right) \\
&= \sum_{l=1}^{m}\left(-(k_l-1)\pi + (k_l-1)(j_l-1)\frac{2\pi}{M}\right) .
\end{aligned}$$

and we immediately retrieve Equation (33). □

The expression (33) emphasizes that $f_{t,s}(\boldsymbol{x}_{k_1,\dots,k_m})$ is computable with a standard DFFT algorithm applied to $(\Delta_\omega)^m\,\delta_{j_1,j_2,\dots,j_m}\Upsilon_{t,s}\left(i(\boldsymbol{\omega}_{j_1,j_2,\dots,j_m})\right)\exp\left(i\sum_{l=1}^{m}\left((k_l-1)\pi\right)\right)$ valued on the meshgrid $j_1,\dots,j_m \in \{1,\dots,M+1\}$ (eventually adjusted by $\frac{1}{(2\pi)^m}$, depending of the DFFT implementation).

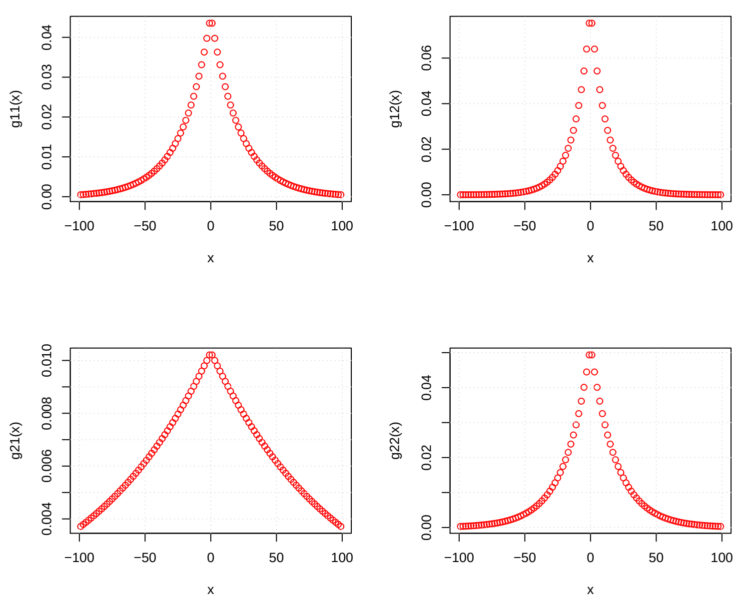

Figure 5: Plots of atomic measures approaching the Laplace distribution defining memory kernels g_{11}, g_{12}, g_{21} and g_{22}.

This procedure is illustrated by the left upper plot of Figure 6 that shows the joint pdf of accumulated numbers of cyber-attacks targeting MED and BSF-O-R sectors, over a one year period. The memory kernels are of Laplace type with parameters of Table 1. The corresponding probability measures are discretized on a grid of $n = 50$ atoms covering the largest 95% confidence interval of the four measures. Figure 5 presents these atomic measures.

The discrete distributions of the number of claims reported on the same day is approached by a continuous exponential random variable, as illustrated in Figure 3. The rates of these random variables are respectively 0.6391 and 0.7067 for the MED and BSF-O-R sectors. The system of ODE's (17) is numerically solved by backward iterations[2] with a time step of $1/200$. The bivariate DFFT is run with $M = 2^6$. We also set $\boldsymbol{\lambda}_0 = \boldsymbol{\theta}$.

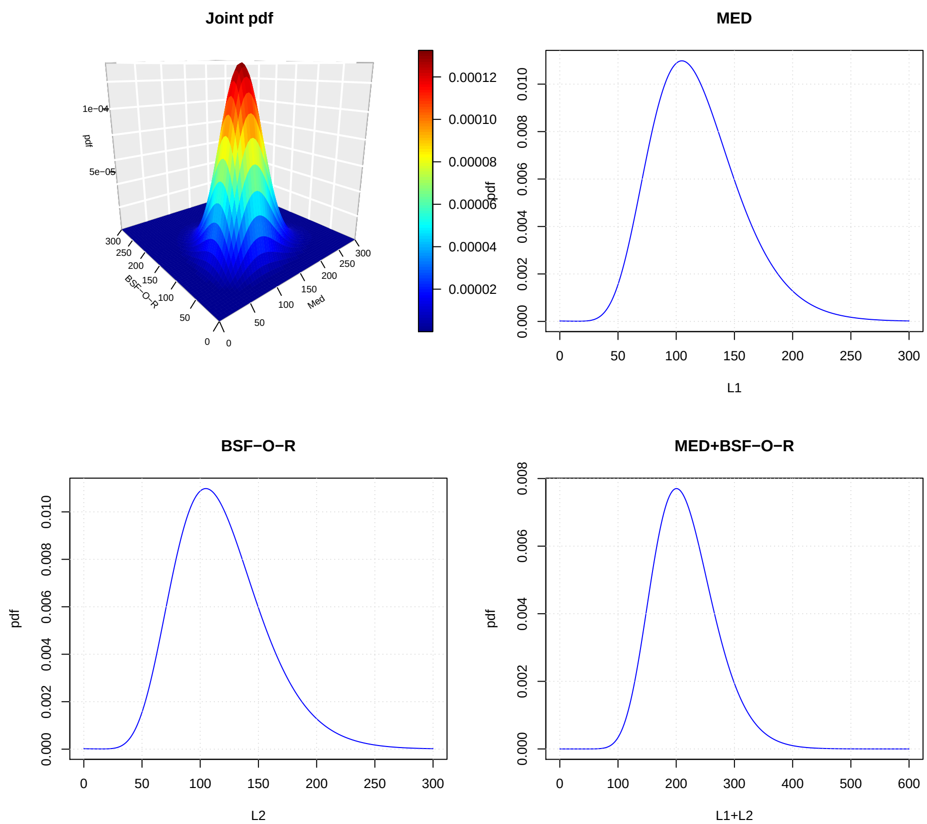

Figure 6: left upper plot: joint pdf of numbers of cyber-attacks targeting MED and BSF-O-R sectors, over 1 year. Other plots: marginal pdf's and pdf of the sum.

In practice, computing the joint pdf of more than 2 claim processes is time consuming. Nevertheless, in high dimension, it is still possible to calculate the marginal pdf's or the pdf of aggregated claim processes with a one dimensional Fourier transform. Let $\boldsymbol{\chi}$ be a m-vector taking its value in the set of orthonormal basis vectors $\{\boldsymbol{e}_1, ...\boldsymbol{e}_m\}$ of $\mathbb{R}^m$ or being equal to the unit vector $\boldsymbol{e} = (1, ..., 1)^\top$. We denote by $f_{t,s}^{\boldsymbol{\chi}}(.)$ the univariate pdf of the linear combination

[2] The $\boldsymbol{q}_l(.)$ are computed with the relation $\boldsymbol{q}_l(t-\Delta, s) = i\, b_l e^{i\, b_l\, \Delta} \boldsymbol{q}_\lambda(t, s)\, \Delta + e^{i\, b_l\, \Delta} \boldsymbol{q}_l(t, s)$ coming from Equation (19) instead of using the related ODE's.

$\boldsymbol{\chi}^\top \tilde{\boldsymbol{L}}_s | \mathcal{F}_t$. This density is equal to the following univariate Fourier transform:

$$f_{t,s}^{\boldsymbol{\chi}}(x) \;=\; \frac{1}{2\pi}\int_{-\infty}^{+\infty} \Upsilon_{t,s}(i\,\omega\,\boldsymbol{\chi}) e^{-i\omega x} d\omega\,.$$

From the previous proposition, this integral can be discretized in a compatible expression with the one dimensional DFFT as detailed in the next corollary.

Corollary 8.2. *Let M be the number of steps used in the Discrete Fast Fourier Transform (DFFT) and $\Delta_x = \frac{x_{max}}{M}$, be the step of discretization. Let us denote $\Delta_\omega = \frac{2\pi}{M\,\Delta_x}$ and $\omega_j = -\frac{M}{2}\Delta_\omega + (j-1)\Delta_\omega$ for $j = 1, ..., M+1$. The values of $f_{t,s}^{\boldsymbol{\chi}}(.)$ at points $x_k = (k-1)\Delta_x$ for $k = 1, ..., M$ are approached by*

$$\begin{aligned} f_{t,s}^{\boldsymbol{\chi}}(x_k) \;&=\; \Delta_\omega \sum_{j=1}^{M+1} \delta_j \Upsilon_{t,s}\left(i(\omega_j\,\boldsymbol{\chi})\right) \exp\left(i\left((k-1)\pi\right)\right) \qquad (34) \\ &\times \exp\left(-i(k-1)(j-1)\frac{2\pi}{M}\right)\,, \end{aligned}$$

where$\delta_j = \left(\frac{1}{2}\right)^{1_{\{j_1=1\}}} + 1_{\{j\neq 1\}}$.

Marginal and aggregated pdf's are then computable in a reasonable time (less than one minute for $M = 2^8$ on a standard laptop) with a 1D DFFT applied to

$$\Delta_\omega \delta_j \Upsilon_{t,s}\left(i(\omega_j\,\boldsymbol{\chi})\right) \exp\left(i\left((k-1)\pi\right)\right)\,.$$

The upper right and lower left, right plots of Figure 6 respectively present the pdf of attacks numbers targeting MED, BSF-O-R and both sectors over a period of one year (with Laplace kernels). These pdf's are obtained with a 1D DFFT and 2^8 discretization steps. The expectations, standard deviations (st. dev.), 5% and 95% percentiles of these distributions are reported in Table 2. The same statistics computed with two independent compound Poisson (CP) processes are reported in Table 3. The comparison of these figures reveals that the CP model overestimates the expected numbers of cyber-attacks but underestimates their standard deviations. This case study confirms the usefulness of results proposed in this article for computing the pdf of self-excited non-Markov claim processes.

Laplace kernels			
	MED	BSF-O-R	TOTAL
Expectation	118.30	96.86	215.30
St. dev.	38.55	35.83	53.81
5% perc.	62.35	45.88	134.12
95% perc.	188.24	162.35	310.59
x_{max}	300	300	600
M	2^8	2^8	2^8

Table 2: Expectations, standard deviations (st. dev.), 5% and 95% percentiles of the number of cyber-attacks over 1 year targeting MED, BSF-O-R and both sectors. x_{max} and M are the DFFT parameters. Laplace memory kernel.

Compound Poisson			
	MED	BSF-O-R	TOTAL
Expectation	148.01	128.76	276.78
St. dev.	21.52	19.08	28.76
5% perc.	112.94	97.25	228.23
95% perc.	185.09	229.01	324.70
x_{max}	300	300	600
M	2^8	2^8	2^8

Table 3: Expectations, standard deviations (st. dev.), 5% and 95% percentiles of the number of cyber-attacks over 1 year targeting MED, BSF-O-R and both sectors. x_{max} and M are the DFFT parameters. Compound Poisson process.

9 Conclusions

This article proposes a closed form expression for the moment generating function of self-excited claim processes with memory kernels that admit a Fourier's representation. This encompasses processes with power, quadratic exponential, hyperbolic sine and exponential decreasing memory functions. We show that this category of models may be reformulated as an infinite dimensional Markov process in the complex plane. Results are found by analysing the asymptotic properties of an approximation with a finite dimensional process. This finite dimensional approximation also offers a solution for the numerical computation of the mgf. We reconciliate the general expression of the mgf proposed in this article with standard one derived in the Markov case.

To illustrate the usefulness of our results, we fit proposed models to time-series of cyber-attacks in the USA, targeting medical and other services from 2014 to 2018. We confirm the results of Bessy-Roland et al. (2020) who find that considering self and mutual contagion between attacks improves the goodness of fit. We detail and apply the method for computing the joint and marginal pdf's of claim processes by DFFT. This allows us to compare moments and percentiles to these of a compound Poisson model. Notice that without the results of this article, the only alternative for approaching the pdf's consists to perform Monte-Carlo simulations.

This article opens the way to further research. Firstly, we can probably extend the family of memory functions casting the framework proposed in this work. Secondly, it would be interesting to investigate the impact of contagion between claims on utility prices. The mgf can also be used for defining an equivalent Esscher measure. Nevertheless, to the best of our knowledge, the dynamic of claims is undetermined under this new measure.

References

[1] Barsotti F., Milhaud X., Salhi Y. 2016. Lapse risk in life insurance: Correlation and contagion effects among policyholders' behaviors. Insurance: Mathematics and Economics (71), 317–331.

[2] Bessy-Roland Y., Boumezoued A., Hillairet C. 2020. Multivariate Hawkes process for cyber insurance. Annals of Actuarial Sciences 15 (1), pp. 14-39.

[3] Dassios A., Jiwook J. 2003. Pricing of catastrophe reinsurance and derivatives using the Cox process with shot noise intensity. Finance and stochastics, 7 (1), pp. 73-95.

[4] Embrechts P., Liniger T., Lu L., 2011, Multivariate Hawkes processes: an application to financial data. Journal of Applied Probability, 48 (A), 367–378.

[5] Hainaut D., 2017. Contagion modeling between the financial and insurance markets with time changed processes. Insurance: Mathematics and Economics 74, pp. 63-77.

[6] Hawkes A., 1971(a). Point sprectra of some mutually exciting point processes. Journal of the Royal Statistical Society Series B, 33, 438-443.

[7] Hawkes A., 1971(b). Spectra of some self-exciting and mutually exciting point processes. Biometrika, 58, 83–90.

[8] Hawkes A., Oakes D., 1974. A cluster representation of a self-exciting process. Journal of Applied Probability, 11, 493-503.

[9] Jiwook J., Dassios A. 2013. A bivariate shot noise self-exciting process for insurance. Insurance: Mathematics and Economics, 53 (3), pp. 524–532.

[10] Jiwook J., Dassios A., Hongbiao Z. 2018. Moments of renewal shot-noise processes and their applications. Scandinavian Actuarial Journal 8, pp. 727-752.

[11] Musmeci F. , Vere-Jones D. 1992. A space-time clustering model for historical earthquakes. Annals of the Institute of Statistical Mathematics, 44, pp 1–11.

[12] Muzy J-F, Delattre S., Hoffmann M., Bacry E. 2013. Some limit theorems for Hawkes processes and application to financial statistics. Stochastic Processes and their Applications, 123(7), 2475-2499.

[13] Ogata Y. 1998. Space-time point process models for earthquake occurences. Annals of the Institute of Statistical Mathematics, 50 (2), pp 397-402.

[14] Porter M.D., White G. 2012. Self-Exciting hurdle models for terrorist activity. The Annals of Applied Statistics, 6 (1) pp 106-124.

[15] Reinhart A., 2018. A Review of self-exciting spatio-temporal point processes and their applications. Statistical Science, 33(3), pp 299–318.

[16] Stabile G. , Torrisi G.L. 2010. Risk processes with non-stationary hawkes claims arrivals. Methodology and Computing in Applied Probability, 12 (3), pp 415–429.

[17] Musmeci F. , Vere-Jones D. 1992. A space-time clustering model for historical earthquakes. Annals of the Institute of Statistical Mathematics, 44, pp 1–11.