A FRACTIONAL HAWKES PROCESS FOR ILLIQUIDITY MODELING

Jean-Loup Dupret, Donatien Hainaut

LIDAM Discussion Paper ISBA
2023 / 01

UCLouvain

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

ISBA
Voie du Roman Pays 20 - L1.04.01
B-1348 Louvain-la-Neuve
Email : lidam-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/isba/publication.html

A fractional Hawkes process for illiquidity modeling

Jean-Loup Dupret[a,*], Donatien Hainaut[a]

[a] *LIDAM-ISBA, Université Catholique de Louvain, Voie du Roman Pays 20, Louvain-La-Neuve, 1348, Belgium*

Abstract

The Amihud illiquidity measure has proven to be very popular in the empirical literature for measuring the illiquidity process of stocks and indices. Many econometric models in discrete time have then been proposed for this Amihud measure. Such models are however not adapted for reproducing peaks of illiquidity with long-memory, for risk management and for pricing liquidity-related derivatives. This paper therefore proposes a new paradigm for modeling illiquidity via a continuous-time process with jumps exhibiting long-range dependence. More precisely, we first introduce a new fractional Hawkes process in which the intensity process is ruled by a modified Mittag-Leffler excitation function. Working with a mean-reverting jump model for the (log-)Amihud measure where jumps follow this modified fractional Hawkes process then allows to easily reproduce the observed peaks of illiquidity in financial markets while introducing long-range dependence and tractability in the model. Indeed, thanks to this modified Mittag-Leffler kernel, we show that our model for the (log)-Amihud measure admits a characteristic function in semi-closed form while having a long-memory of past events, which is not achievable with the existing Hawkes processes. We can therefore use this model to perform risk management on illiquidity as well as to introduce and price illiquidity derivatives on the Amihud measure. We hence provide with this paper new tools for a better understanding and management of the illiquidity risk in financial markets.

Keywords: Illiquidity modeling, Amihud measure, Hawkes process, Mittag-Leffler function, Illiquidity derivatives, Risk management

1. Introduction

Hawkes processes appear as good candidates for modeling the arrival of financial random events exhibiting a self-exciting behavior. They have been heavily used in finance for modeling *e.g.* the occurrence of trades at high frequency (see Bacry and Muzy (2014), Bacry et al. (2015), Hardiman et al. (2013), Lee and Seo (2017)), the firm defaults in a credit risk portfolio (see Da Fonseca and Zaatour (2014), Errais et al. (2010)) or the jumps of asset prices in financial markets (see Aït-Sahalia et al. (2015), Branger et al. (2014), Hawkes (2022)). The self-exciting behavior of Hawkes processes is characterized by the so-called excitation kernel, which is a decreasing function that provides the contribution of past jumps to the instantaneous probability of future jumps' occurrence. Several functions have been proposed in the literature as excitation kernel. The most common one is the exponential kernel which allows for a tractable and analytical treatment of Hawkes processes with explicit formulae for the moments, autocorrelation function and characteristic function of the number of jumps, see Aït-Sahalia et al. (2015), Da Fonseca and Zaatour (2014), Errais et al.

*Corresponding author, Email address: `jean-loup.dupret@uclouvain.be`

(2010). However, this exponential kernel cannot take into account the long-range dependence of past events on the number of jumps, which is typical in many financial applications (cfr Hainaut (2022) for a review and Bacry et al. (2015) in the context of high-frequency trading). Therefore, another widely-used excitation kernel proposed in the literature is the power-law one since it allows for a slower decay than the exponential kernel and can thus better reproduce the observed long-term effects in the number of jumps. This power-law kernel has also been extensively used in high-frequency modeling where the order flow is known to exhibit long-range dependence, as investigated in Bacry et al. (2012, 2015), Hardiman et al. (2013). However, unlike the exponential kernel, the Hawkes process with power-law excitation function does not admit an analytical characteristic function of the number of jumps and thus lacks of tractability for many financial applications : valuation and risk management of portfolio credit derivatives as in Errais et al. (2010), option pricing in rough volatility models as in El Euch and Rosenbaum (2019), option pricing for self-exciting jump diffusion models as in Hawkes (2022) or bond pricing under a self-exciting Hawkes dynamic for the short rate process as in Hainaut (2016b), etc. Hence, several authors have tried to fill this gap by introducing Hawkes models based on fractional dynamics. Hainaut (2020) proposes a fractional Hawkes process defined by a subdiffusive dynamic for the intensity process while Njike Leunga (2022) and Chen et al. (2021) consider respectively the Mittag-Leffler function and the derivative of the Mittag-Leffler function as excitation kernel. In the first part of this work, we hence review these fractional Hawkes processes and show that none of them are fully satisfying for the modeling of financial events. We then introduce a new fractional Hawkes process based on a modified Mittag-Leffler kernel that yields a tractable characteristic function while exhibiting long-memory. We study its properties and show that this new fractional Hawkes process is best suited for financial applications with long-range dependence, especially when the characteristic function for the number of jumps is required. Indeed, following Bäuerle and Desmettre (2020), we show that we can rewrite our fractional Hawkes process with modified Mittag-Leffler kernel as a superposition of an infinite number of Markov processes. Discretizing the obtained representation, a semi-closed form solution is derived for the conditional transform of the number of jumps.

From the theoretical results developed in the first part of the paper, we provide a practical application of this modified fractional Hawkes model in the context of illiquidity modeling and shows that it outperforms existing self-exciting models. Several econometric indicators have been proposed for measuring the illiquidity process of stocks and indices using daily (or weekly) frequency data, see Goyenko et al. (2009) for a review. We here focus on the Amihud illiquidity measure as introduced in Amihud (2002). The Amihud measure has proven to be very popular in the empirical literature and to influence the cross-sectional asset returns through the so-called illiquidity premium, see the review of Amihud and Mendelson (2015). This illiquidity premium comes from the fact that investors care about illiquidity costs and price them in the expected returns. Numerous econometric models in discrete time have then been proposed for modeling this Amihud measure, cfr Amihud and Mendelson (2015), Brennan et al. (2013) or more recently Hafner et al. (2022). In this work, we shift of paradigm and consider a continuous-time process instead. More precisely, we model the (log-)Amihud measure as a mean-reverting jump diffusion process where the jumps follow a fractional Hawkes model with the modified Mittag-Leffler excitation function, as introduced in the first part of the paper. This continuous-time approach simplifies many calculations and helps us to obtain the characteristic function satisfied by the Amihud measure. Therefore, this model allows to easily perform risk management on illiquidity as well as to introduce and price financial instruments related to the Amihud measure. This way, we provide market participants with new tools to better manage and assess their risk of liquidity, to reduce their exposition to this risk and to express views on future liquidity conditions. Moreover, adding fractional Hawkes jumps in the model allows to reproduce the observed peaks of illiquidity in financial markets and to show that these peaks effectively exhibit long-memory. Finally, by considering a multivariate setting for our fractional Hawkes process, we can study the contagion between the up and down shocks of this Amihud measure. We can also adapt this multivariate setting such as to simultaneously model the illiquidity process of multiple indices in financial markets as well as their contagion.

The goal of this paper is therefore fivefold. We first introduce a new fractional Hawkes process with a modified Mittag-Leffler kernel and show that it exhibits better properties for the modeling of financial events with long-range dependence, compared with existing Hawkes models. We then propose a new paradigm based on continuous-time processes for studying the illiquidity of financial assets via the Amihud illiquidity measure. Working with a mean-reverting jump model for the (log-)Amihud measure where jumps follow this modified fractional Hawkes process allows to easily reproduce the observed peaks of illiquidity in financial markets while introducing long-range dependence and tractability in the model. In a third step, we show via a Peaks Over Treshold procedure with likelihood maximization how to estimate the parameters of our mean-reverting fractional Hawkes model. By proposing a change of measure and deriving the characteristic function satisfied by the Amihud measure in this model, we can then introduce and price new financial instruments related to the illiquidity process of stocks and indices. Finally, we show in the last section of this paper how to perform risk management on illiquidity by providing an efficient way to compute the distribution, moments and risk measures of the Amihud measure so as to help market participants to better assess and manage their risk of liquidity. We hence provide with this paper new tools for a better understanding and assessment of the liquidity risk in financial markets.

2. Hawkes processes

The *Hawkes process* is a mathematical model for self-exciting processes. More precisely, it is a counting process that models a sequence of arrivals of some type over time such as earthquakes, gang violence, trade orders, price jumps or firm defaults. Each arrival excites the process in the sense that the chance of a subsequent arrival is increased for some time period after the initial event. As such, it is a non-Markov extension of the Poisson process.

Consider a sequence of stopping times $0 < \tau_1 < \tau_2 < \ldots$ corresponding to the arrivals of jumps which represent some market events such as peaks of illiquidity, defaults in a portfolio of firms or the occurrence of a trade. These stopping times are defined on a complete probability space $(\Omega, \mathcal{F}, P)$ with right-continuous and complete information filtration $\mathbb{F} = (\mathcal{F})_{t\geq 0}$. The size of the jump at τ_j is given by a random variable $O_j \in \mathcal{F}_{\tau_j}$. The sequence (τ_j, O_j) generates a nonexplosive counting process N given by $N_t = \sum_{j\geq 1} \mathbb{1}_{\{\tau_j \leq t\}}$ and a jump point process L defined by

$$L_t = \sum_{j\geq 1} O_j \, \mathbb{1}_{\{\tau_j \leq t\}}. \tag{2.1}$$

We propose to specify the processes N and L directly through a conditional arrival rate or intensity λ and a distribution ν on $(0,\infty)$ for the loss size O_j (assumed to be strictly positive in this paper). We suppose that the jump transform $\int_0^\infty e^{\xi o}\, d\nu(o)$ exists, is finite for complex ξ and admits a finite derivative $\int_0^\infty o\, e^{\xi o}\, d\nu(o)$. The intensity follows a strictly positive stochastic process that describes the conditional mean arrival rate such that

$$\begin{aligned} \lambda_t \mid \mathcal{F}_t = \lim_{h\to 0} \frac{\mathbb{E}[N_{t+h}|\mathcal{F}_t] - N_t}{h} &= \lambda_0(t) + \eta \int_0^t f(t-u)\, dL_u \\ &= \lambda_0(t) + \eta \sum_{j\,:\,\tau_j<t} f(t-\tau_j)\, O_j \end{aligned} \tag{2.2}$$

where $\eta > 0$ and where $\lambda_0(t) > 0$ is the background intensity function, describing the arrival of events triggered by a deterministic exogenous source. The function $f : \mathbb{R} \to \mathbb{R}^+$, called the self-excitation kernel of the process, is a decreasing function that provides the contribution to the intensity λ of a jump that occurred at a previous time $u < t$. Hence, each event increases the intensity by $\eta \lim_{h\to 0} f(h)O$, which then decays according to the excitation function $\phi := \eta f$ until the next event occurs to push it up again. This excitation kernel ϕ is decreasing so that more recent events have higher influence on the current intensity compared to events having occurred further away in the past. From Bacry et al. (2015), we have that the two-dimensional process $J = (L, N)^\top$ is a Hawkes process provided that the excitation kernel ϕ is

1. **Positive** : $\phi(t) \geq 0\ \forall t$.
2. **Causal** : $\phi(t) = 0$ for $t < 0$.
3. $\mathbf{L^1}$**-integrable** : $\int_0^\infty \phi(s)\, ds < \infty$.

From equation (2.2), we clearly have that the higher the jump size O_j, the higher the effect of the default on the intensity. Moreover, the intensity λ governs the common event time of N and L. However, while the jumps of the counting process N are unit-sized, the jumps of the point process L are drawn from the distribution ν. Finally, it can be shown that $N - \int_0^\cdot \lambda_s\, ds$ is a local martingale relative to $\mathbb{F}$ and P, cfr Errais et al. (2010).

A nice interpretation of Hawkes processes is given in terms of a population process where migrants arrive according to an inhomogeneous Poisson process with rate $\lambda_0(t)$. Then, each migrant gives birth to children according to another inhomogeneous Poisson process with intensity function ϕ, these children also giving birth to children according to the same inhomogeneous Poisson process. In our financial context, migrants can be seen as exogenous financial events whereas children are viewed as the events triggered by other previous events (defaults, jumps or trades). Hence, as explained in Bacry et al. (2015), the L^1 norm of ϕ (denoted $||\phi||_1 := \int_0^\infty \phi(t)\, dt = \eta\, ||f||_1$) can be interpreted as a branching ratio, *i.e.*, the number of events generated by any parent event. Moreover, $||\phi||_1$ provides a direct measure of the fraction of endogenously triggered events within the whole population of events and is thus a measure of market reflexivity. Finally, even though the process defined by equation (2.2) is well-defined for any choice of kernel ϕ satisfying the three conditions stated above, a stationary Hawkes process can be characterized as follows. The counting process N has asymptotically stationary increments and the intensity λ is asymptotically stationary if the kernel ϕ satisfies $||\phi||_1 < 1$. In the multivariate extension of the Hawkes process that will be introduced in Section 3.4, the L^1 norm required for this stationary condition is replaced by the spectral radius, as explained in Bacry et al. (2015).

This setting is largely used in the financial literature (see Bacry et al. (2015)) to model the arrival of trade orders at high-frequency as the counting process N and the associated microstructure price as the loss process L (where O_j then denotes the order size of the j^{th} trade). Similarly, firm defaults in a portfolio of credit risk can be modeled by the counting process N and the loss-given default by the random variable O_j, as in Errais et al. (2010) or Da Fonseca and Zaatour (2014). Finally, jumps in asset prices can also be studied based on these Hawkes processes, cfr Aït-Sahalia et al. (2015) or Hawkes (2022).

In a first step, we review the two most frequent choices of the decay function f in the literature.

- **Exponential kernel :**

$$f(t) = e^{-\alpha t}, \; t > 0\,, \, \alpha > 0 \tag{2.3}$$

where $\alpha > 0$ is the rate at which the impact of an event decays over time and where the L^1 norm of ϕ, $||\phi||_1$, is equal to η/α. More precisely, a widely-used intensity process is given by

$$\lambda_t = \lambda_\infty + e^{-\alpha t}(\lambda_0 - \lambda_\infty) + \eta \int_0^t e^{-\alpha(t-u)}\, dL_u \tag{2.4}$$

$$= \lambda_0 + \int_0^t f(t-u)\, \alpha\, (\lambda_\infty - \lambda_0)\, du + \eta \int_0^t f(t-u)\, dL_u\,, \tag{2.5}$$

which can be rewritten by differentiating the previous equation as

$$d\lambda_t = \alpha\, (\lambda_\infty - \lambda_t)\, dt + \eta\, dL_t\,. \tag{2.6}$$

The process does not blow up because the drift becomes negative whenever the intensity is above $\lambda_\infty > 0$ and prevents any explosion. The point process $J = (L, N)$ is not Markov since it depends upon λ but the pair $X = (\lambda, J)^\top$ is well Markov in the state space $D = \mathbb{R}_+ \times (\mathbb{R}_+ \times \mathbb{N})$. Therefore, the infinitesimal generator of X, denoted $\mathcal{L}$, is the operator acting on a sufficiently regular[1] function $f : D \to \mathbb{R}$ such that **(changer f en g car confusion avec decay function f above)**

$$\mathcal{L}f(x) = \lim_{h\to 0} \frac{\mathbb{E}_t[f(X_{t+h})] - f(x)}{h}\,,$$

where $\mathbb{E}_t = \mathbb{E}[\,.\mid \mathcal{F}_t]$ and $X_t = x = (\lambda_t, (L_t, N_t))^\top$. In the case of the Hawkes process (2.4), we can write

$$\mathcal{L}f(x) = \alpha(\lambda_\infty - \lambda)\frac{\partial f}{\partial \lambda}(x) + \lambda \mathbb{E}\left[f\left(\lambda + \eta O,\, L + O,\, N+1\right) - f(x)\right]$$

$$= \alpha(\lambda_\infty - \lambda)\frac{\partial f}{\partial \lambda}(x) + \lambda \int_0^\infty \left(f\left(\lambda + \eta o,\, L + o,\, N+1\right) - f(x)\right) \nu(do)\,.$$

Hence, Da Fonseca and Zaatour (2014) obtain the following Dynkin formula for this Hawkes process

$$\mathbb{E}_s\left[f(X_t)\right] = f(X_s) + \mathbb{E}_s\left[\int_s^t \mathcal{L}f(X_u)\, du\right]. \tag{2.7}$$

[1] with continuous partial derivative $\partial f/\partial\lambda$.

Furthermore, since the process $X = (\lambda, J)^\top$ is an affine Markov process, we can express for $u = (u_\lambda, u_L, u_N) \in \mathbb{C}^3$ and $t \le T$ its conditional transform $\mathbb{E}_t\left[e^{u\cdot X_T}\right]$ as $f(t,T,X_t)$, where f is a complex-valued function f on $[0,T]\times D$. Since f must satisfy the PIDE

$$0 = \frac{\partial f}{\partial t}(t,T,x) + \mathcal{L}f(t,T,x)\,, \tag{2.8}$$

with boundary condition $f(T,T,X_T) = e^{u\cdot X_T} = e^{u_\lambda \lambda_T + u_L L_T + u_N N_T}$, we have from Errais et al. (2010) that the conditional transform of the point process $X = (\lambda, J)^\top = (\lambda, (L,N))^\top$ is given by

$$\mathbb{E}_t\left[e^{u\cdot X_T}\right] = \exp\left(a(t,T) + b(t,T)\lambda_t + u'\cdot J_t\right),$$

where $u' = (u_L, u_N) \in \mathbb{C}^2$ and the coefficient functions $a(t,T) := a(u',t,T)$ and $b(t,T) := b(u',t,T)$ satisfy the following ODEs

$$\begin{aligned}
\partial_t b(t,T) &= \alpha\, b(t,T) + 1 - \theta\left(\eta\, b(t,T) + u'\cdot(1,0)^\top\right) e^{u'\cdot(0,1)^\top}\,,\\
\partial_t a(t,T) &= -\alpha\lambda_\infty b(t,T)\,,
\end{aligned}$$

with boundary conditions $a(T,T) = 0$, $b(T,T) = u_\lambda$ and where $\theta(.)$ is the jump transform

$$\theta(\xi) = \int_0^\infty e^{\xi o}\, d\nu(o),\ \ \xi \in \mathbb{C}\,.$$

This exponential kernel is used *e.g.* in Errais et al. (2010) in the context of portfolio credit risk, where the correlation of firm defaults is the main concern. Such specification of the Hawkes process allows to valuate derivatives on debt portfolio in a highly tractable way thanks to the existence of closed-form formulae for its moments and characteristic function. However, in this framework of exponential excitation kernel, the autocovariance of the intensity decays exponentially with time. This feature is not necessary adapted for modeling real phenomena that exhibit a long-term memory of past events such as trade occurrence, default events or illiquidity peaks (as we will see in Section 4). For example, it has been shown empirically in Hardiman et al. (2013) and in Bacry et al. (2012) that the exponential kernel is not the best suited for reproducing the occurrence of trades at high-frequency. Hence, other excitation functions need to be considered in order to better reproduce this long-memory property (such as the following power-law kernel).

- **Power-law kernel :**

$$f(t) = \frac{\alpha}{(1+\alpha t)^{1+\gamma}},\ \ \alpha > 0,\ \ t > 0. \tag{2.9}$$

The L^1 norm $||\phi||_1$ of the power-law kernel is now equal to η/γ. This kernel makes the Hawkes process non-Markov and we hence cannot rely on the above theory of infinitesimal generators to write the characteristic function of (λ, J). In the case of portfolio credit derivatives, this prevents from valuating efficiently these derivative contracts as done in Errais et al. (2010) with the exponential kernel (2.3) above. Similarly for illiquidity modeling, this characteristic function is required for efficient risk management and option pricing as we will see in Section 4. Other applications requiring this characteristic function includes the pricing of options for self-exciting jump diffusion model as in Aït-Sahalia et al. (2015) or the pricing of bonds under a Hawkes dynamic for the short rate process as in Hainaut (2016b). Nevertheless, the main benefit of this power-law kernel is to allow for a slower form of decay than the exponential kernel so that it can be used for better modeling long-term effects. This long-term power-law decay is for example observed at high-frequency with $\gamma \approx 0$ (and thus with a power-law exponent close to $1+\gamma \approx 1$) as described in Hardiman et al. (2013), which hence allows to better model the long-range nature of offer and demand. We therefore want to build a kernel exhibiting such long-range dependence while having a characteristic function available in (semi-)closed form so that we can better manage the illiquidity risk and write options on liquidity-related instruments. Fractional Hawkes processes offer ways to achieve this.

2.1. Fractional Hawkes processes

Several alternatives based on fractional kernels and fractional dynamics have been proposed in the literature to enhance the two previous models. However, none of them fully satisfies the empirical properties observed in financial markets. We first review them and then propose an alternative specification based on a modified Mittag-Leffler kernel allowing to exhibit long-term memory in the intensity process while having a characteristic function available in semi-closed form.

- **Mittag-Leffler kernel :**

$$f_\gamma(t) = E_\gamma(-t^\gamma)\,, \qquad t > 0,\ \gamma \in (0,1]. \tag{2.10}$$

where $E_\gamma(z) := \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(n\gamma+1)}$ is the one-parameter Mittag-Leffler (ML) function with $\gamma \in (0,1]$. If computed on $z = -t^\gamma$ for $t \geq 0$, the one-parameter ML function $E_\gamma(-t^\gamma)$ has the meaning of survival function for a positive random variable T with infinite mean. This function in fact interpolates between a stretched exponential for small times and a power-law with index $\gamma \in (0,1]$ for large times as we can see with the following expansion. The choice of this kernel is motivated in a more general setting in Njike Leunga (2022) by the behavior of the ML function at short and long term. Indeed, it is commonly known that $E_\gamma(-t^\gamma)$ matches for $t \to 0$ with a stretched exponential with infinite negative derivative, whereas it matches as $t \to \infty$ with a negative power-law, cfr Mainardi (2020) :

$$E_\gamma(-t^\gamma) \sim \begin{cases} \exp\left[-\frac{t^\gamma}{\Gamma(1+\gamma)}\right], & t \to 0, \\ \\ \frac{t^{-\gamma}}{\Gamma(1-\gamma)} = \frac{\sin(\gamma\pi)}{\pi}\frac{\Gamma(\gamma)}{t^\gamma}, & t \to \infty. \end{cases} \tag{2.11}$$

As a consequence of the long-term power-law asymptotic, the process turns to be no longer Markov but of long-memory type. However, the L^1 norm is infinite which prevents from using this excitation function for modeling the arrival of financial random events. Indeed, using the probabilistic interpretation of the ML function as survival random variable T, we find

$$||f_\gamma(t)||_1 = \int_0^\infty E_\gamma(-u^\gamma)\,du = \int_0^\infty P(T \geq u)\,du = \mathbb{E}[T] = \infty$$

since from Theorem 3 (ii) in Lin (1998), the expected value of the the ML distribution T is infinite. In this case, each exogenous event generates infinitely many endogenous events and the average intensity explodes at large time scales, making the choice of the ML kernel not conceivable for modeling financial events. Indeed, this excitation function does not even define a Hawkes process since the third condition of Definition (2.2) is not satisfied.

- **Mittag-Leffler's derivative kernel :**

$$f_\gamma(t) = -\frac{dE_\gamma(-t^\gamma)}{dt} = t^{\gamma-1}E_{\gamma,\gamma}(-t^\gamma)\,, \qquad t>0,\ \gamma\in(0,1]. \tag{2.12}$$

where $E_{\gamma,\delta}(z) := \sum_{n=0}^{\infty}\frac{z^n}{\Gamma(n\gamma+\delta)}$ is the two-parameter Mittag-Leffler function. This kernel (2.12) is in fact the sign-changed first derivative of the ML kernel (2.10) and hence, it is now the pdf of the survival random variable T introduced above. This kernel was first considered in Chen et al. (2021) for defining a so-called *fractional Hawkes process*, which is a non-Markov Hawkes process with long-memory. This kernel is thus again positive and causal. One can prove easily that the L^1 norm of this kernel $||f||_1$ is equal to 1 (since it is the pdf of the survival random variable T) and also that it decays for large times as $K\,t^{-\gamma-1}$, with K a positive constant. Indeed, we have from Mainardi (2020) that

$$t^{\gamma-1}E_{\gamma,\gamma}(-t^\gamma) \sim \begin{cases} +\infty & t\to 0^+, \\ \\ \frac{-t^{-\gamma-1}}{\Gamma(-\gamma)} = \frac{\sin(\gamma\pi)}{\pi}\frac{\Gamma(\gamma+1)}{t^{\gamma+1}}, & t\to\infty. \end{cases} \tag{2.13}$$

The three conditions of Definition (2.2) are therefore satisfied and the Hawkes process defined by the ML's derivative kernel (2.12) is thus well defined. However, we directly see in equation (2.12) with $\gamma < 1$ and in (2.13) that this kernel tends to infinity when $t \to 0^+$ and hence cannot capture the right behavior of the Hawkes process right after its jumps (given by $\eta \lim_{t\to 0^+} f_\gamma(t)O$). Moreover, the developments we will make for deriving the characteristic function of (λ, J) in Section 3.3 will not apply to this kernel (2.12)[2]. As this characteristic function is required to price options and perform risk management on illiquidity (cfr Section 4), this prevents from using the ML's derivative kernel.

- **Subordinated kernel with inverse α-stable process :**

Another fractional Hawkes process is introduced in Hainaut (2020) by subordinating the intensity process (2.6) with the inverse of a α-stable Lévy process, *i.e.*

$$d\lambda_{S_t} = \alpha(\lambda_\infty - \lambda_{S_t}) + \eta\, dL_{S_t}$$

where $S_t = \inf\{\tau > 0 : U_\tau \geq t\}$ and U_t is an α-stable process with transform $\mathbb{E}[e^{-\xi U_t}] = e^{-t\xi^\alpha}$. Such Hawkes process ruled by an intensity which exhibits periods of constant values is again non-Markov. Moreover, this process also allows to take into account long-memory in a financial context thanks to a sub-exponential autocovariance function for its intensity λ_{S_t}, cfr Hainaut (2020). However, this constant piecewize intensity function is not justified empirically for modeling such financial effects since no market events could explain the freeze at some random value of the instantaneous probability λ_{S_t}. Furthermore, the generalization to a multivariate Hawkes process (which will be required in Section 3.4) has not been investigated yet in the literature due to the difficulty to define a multivariate inverse subordinated process. Similarly, dependence between illiquidity shocks (or between firms default) cannot be modeled in this framework due to the need to work with multiple correlated inverse subordinated processes. Finally, even though this kernel allows for a sub-exponential decay of the autocovariance function and hence long-memory, it cannot reproduce the empirical power-law decay observed in many financial applications, cfr Bacry et al. (2015), Hardiman et al. (2013). We now introduce the modified ML kernel and show that its properties are best suited for modeling such financial events with long memory.

[2] see Section 3.3, in this case the equation (B.2) explodes and becomes $\lim_{n\to\infty}\sum_{\omega=1}^{n} m_\omega^{(n)} = +\infty$.

3. A modified Mittag-Leffler fractional Hawkes process

3.1. Properties

We now introduce a new fractional Hawkes process based on the following excitation kernel, defined as the difference of a ML kernel and its derivative, namely

$$f_\gamma(t) = E_\gamma(t^\gamma) - t^{\gamma-1} E_{\gamma,\gamma}(t^\gamma)\,, \quad \gamma \in [1,2],\ t > 0, \tag{3.1}$$

which we call the modified Mittag-Leffler (mML) kernel. It is important to note that we now impose $\gamma \in [1,2]$, whereas for kernels (2.10) and (2.12), this parameter was restricted to $\gamma \in (0,1]$. Moreover, the ML functions in (3.1) are now evaluated at t^γ and not $-t^\gamma$ anymore. We can directly expand the continuity of (3.1) in zero and hence obtain the right limit $\lim_{t\to 0^+} f_\gamma(t) = f_\gamma(0+) = 1$, ensuring a non-explosion and reasonable behavior of the Hawkes process at the occurrence of a jump. From the expansion 18.1(22) in Bateman (1953), we can also easily derive the following asymptotics for our mML kernel

$$f_\gamma(t) \sim \begin{cases} 1 - \frac{t^{\gamma-1}}{\Gamma(\gamma)} \sim \exp\left[-\frac{t^{\gamma-1}}{\Gamma(\gamma)}\right] & t \to 0^+, \\ \\ -\frac{t^{-\gamma}}{\Gamma(1-\gamma)} & t \to \infty. \end{cases} \tag{3.2}$$

Hence, we see that it behaves as a stretched exponential for short time, exactly as the ML kernel (2.10). Indeed, we obtain the same short-time asymptotic (2.11) as the ML kernel given that now $\gamma \in [1,2]$, whereas $\gamma \in (0,1]$ in the excitation function (2.10). Secondly, we observe that at large times this mML kernel behaves exactly as the ML's derivative kernel (2.12) since we again have from the condition $\gamma \in [1,2]$ the same long-term asymptotic (2.13), for which $\gamma \in (0,1]$. Therefore, we have that our newly introduced mML kernel interpolates between a stretched exponential for small times and a power-law with index $\gamma \in (1,2]$ for large times as we can see with the previous expansions. This kernel thus handles in an unified way the properties of the exponential kernel and the power law kernel, which will lead to a better modeling and fine tuning of the short and long-term properties of the Hawkes process. We now show that the L^1 norm of the excitation kernel $\phi_\gamma := \eta f_\gamma$ is equal to η and does not explode anymore as in the classical ML kernel. The parameter η hence controls the clustering property of the mML fractional Hawkes process.

Proposition 3.1. *The L^1 norm of the modified Mittag-Leffler kernel* (3.1) *is given by*

$$||f_\gamma||_1 = \int_0^\infty f_\gamma(t)\,dt = \int_0^\infty \left(E_\gamma(t^\gamma) - t^{\gamma-1}E_{\gamma,\gamma}(t^\gamma)\right)dt = 1\,, \tag{3.3}$$

which thus leads to the L^1 norm $||\phi_\gamma||_1 = \eta||f_\gamma||_1 = \eta < \infty$. Since ϕ_γ is also positive and causal, the three conditions of Definition 2.2 are verified and the mML fractional Hawkes process ruled by the excitation function (3.1) *is thus well-defined.*

Proof. To derive this norm, we first recall some probabilistic results derived from Kyprianou (2006) (page 216 and Theorem 6.16). Let us first define an α-stable process without negative jumps with parameter $\alpha \in (1,2]$ denoted U and we then denote $\widetilde{U} = -U$, an α-stable process without positive jumps. We can now define the supremum $\overline{U}_t := \sup\{U_s,\ s \le t\} = -\inf\{\widetilde{U}_s,\ s \le t\}$ and $\mathbf{e_1}$ an exponential random variable with parameter 1, independent from U. Therefore, we have from Kyprianou (2006) that the Wiener-Hopf factorization leads for $\lambda \ge 0$ to

$$\mathbb{E}[e^{-\lambda \overline{U}_{\mathbf{e_1}}}] = \frac{\lambda - 1}{\lambda^\alpha - 1}\,, \tag{3.4}$$

and hence, integrating by parts, to

$$\int_0^{\infty} e^{-\lambda t} P\left(\overline{U}_{\mathbf{e_1}} \geq t\right) dt = \frac{\lambda^{\alpha-1}}{\lambda^{\alpha}-1} - \frac{1}{\lambda^{\alpha}-1}\,.$$

Moreover, it is well-known from Mainardi (2020) that the function $E_\gamma(t^\gamma)$ has the following explicit Laplace transform

$$\mathcal{L}\left[E_\gamma(t^\gamma)\right](\lambda) = \int_0^{\infty} e^{-\lambda t} E_\gamma(t^\gamma)\, dt = \frac{\lambda^{\gamma-1}}{\lambda^{\gamma}-1}\,,$$

and integrating by parts, it comes for the modified Mittag-Leffler kernel (3.1) :

$$\mathcal{L}\left[f_\gamma(t)\right](\lambda) = \int_0^{\infty} e^{-\lambda t}\left(E_\gamma(t^\gamma) - t^{\gamma-1}E_{\gamma,\gamma}(t^\gamma)\right) dt = \frac{\lambda^{\gamma-1}}{\lambda^{\gamma}-1} - \frac{1}{\lambda^{\gamma}-1} = \int_0^{\infty} e^{-\lambda t} P\left(\overline{U}_{\mathbf{e_1}} \geq t\right) dt\,, \tag{3.5}$$

when identifying γ and α. Hence, we finally find by inverting the Laplace transform (3.5) :

$$P\left(\overline{U}_{\mathbf{e_1}} \geq t\right) = E_\gamma(t^\gamma) - t^{\gamma-1}E_{\gamma,\gamma}(t^\gamma)\,.$$

We therefore find the following L^1 norm for the modified Mittag-Leffler kernel

$$||f_\gamma||_1 = \int_0^{\infty} P\left(\overline{U}_{\mathbf{e_1}} \geq t\right)\, dt = \mathbb{E}\left[\overline{U}_{\mathbf{e_1}}\right] = 1.$$

The expectation is obtained by differentiating the moment generating function (mgf) (3.4) with respect to λ and evaluating the derivative at $\lambda = 0$. □

Finally, the following proposition will help us to derive an expression for the conditional transform of (λ, J) in the mML fractional Hawkes model, which is fundamental for the sequel of this paper. From now on, we will hence rely heavily on the following Proposition 3.2.

Proposition 3.2. *The modified Mittag-Leffler kernel* (3.1) *is completely monotone*[3] *if and only if* $\gamma \in [1, 2]$. *This is equivalent to the existence of a representation of the kernel* (3.1) *in terms of a Laplace-Stieltjes integral with non-decreasing density and non-negative measure given by*

$$f_\gamma(t) = \int_0^{\infty} e^{-\omega t}\mu(d\omega) \tag{3.6}$$

where

$$\mu(d\omega) = \frac{-\sin(\gamma\pi)\,\omega^{\gamma-1}(1+\omega)}{\pi\left(\omega^{2\gamma} - 2\omega^{\gamma}\cos(\gamma\pi) + 1\right)} d\omega \tag{3.7}$$

Proof. This proof is direct from Titchmarsh formula for inversion of Laplace transforms, see Gross (1950). We first have that the Laplace transform of the kernel (3.1) is given by equation (3.5) :

$$\mathcal{L}_{f_\gamma}(\lambda) = \int_0^{\infty} e^{-\lambda t} f_\gamma(t) dt = \frac{\lambda^{\gamma-1}-1}{\lambda^{\gamma}-1}\,. \tag{3.8}$$

We have that this function is analytic such that $\mathcal{L}_{f_\gamma}(\lambda) \to 0$ as $|\lambda| \to \infty$ and $\mathcal{L}_{f_\gamma}(\lambda) = o(|\lambda|^{-1})$ as $|\lambda| \to 0$ uniformly in every sector $|\arg \lambda| < \pi$, and that for $r > 0$ and $\theta \in (-\pi, \pi)$:

$$\Re\left(e^{i\theta/2}\mathcal{L}_{f_\gamma}(re^{i\theta}) + e^{-i\theta/2}\mathcal{L}_{f_\gamma}(re^{-i\theta})\right) = 2\,\frac{\cos(\theta/2)(\lambda^{2\gamma-1}+1) - \cos(\theta\alpha)(\lambda^{\gamma-1}+\lambda^{\gamma})}{\lambda^{2\gamma} - 2\lambda^{\gamma}\cos(\theta\gamma) + 1} \geq 0\,,$$

[3] A function $f : (0, +\infty) \to \mathbb{R}$ is completely monotone if it is C^∞ and if $(-1)^n f^{(n)}(x) \geq 0\,, \quad \forall n \in \mathbb{N}$.

with $\alpha = \gamma - 1/2 \in [1/2, 3/2]$. Therefore, the assumptions of Hirschman and Widder (2012) are satisfied and we then have the following "*restoring pair property*". Under these assumptions, the knowledge of $\mathcal{L}_{f_\gamma}(\lambda)$ enables one to find the inverse transform $\mu(\omega)$ of $f_\gamma(t)$, that is

$$f_\gamma(t) = \int_0^\infty e^{-\omega t} \mu(\omega)\, d\omega\,. \tag{3.9}$$

Following Gross (1950), we have from (3.8) and (3.9) that

$$\mathcal{L}_{f_\gamma}(\lambda) = \int_0^\infty \frac{\mu(\omega) d\omega}{\omega + \lambda}\,. \tag{3.10}$$

This is a Stieltjes equation, which is inverted by

$$\begin{aligned}\mu(\omega) = \frac{1}{\pi}\mathrm{Im}\left[\mathcal{L}_{f_\gamma}(\omega e^{-i\pi})\right] &= \frac{\omega^{2\gamma-1}\sin(-\pi(\gamma-1)+\pi\gamma) - \omega^{\gamma-1}\sin(-\pi\gamma+\pi) + \omega^\gamma \sin(-\pi\gamma)}{\pi\left(\omega^{2\gamma} - 2\omega^\gamma\cos(\gamma\pi) + 1\right)} \\ &= \frac{-\sin(\gamma\pi)\,\omega^{\gamma-1}(1+\omega)}{\pi\left(\omega^{2\gamma} - 2\omega^\gamma\cos(\gamma\pi)+1\right)}\,. \end{aligned} \tag{3.11}$$

□

3.2. Final model and spectral representation

Our final model is obtained by extending the exponential intensity model (2.5) with the mML kernel (3.1) in the following way

$$\begin{aligned}\lambda_t &= \lambda_0 + \int_0^t f_\gamma(t-u)\,\alpha\,(\lambda_\infty - \lambda_u)\,du + \int_0^t f_\gamma(t-u)\,\eta\,dL_u \\ &= \lambda_0 + \int_0^t f_\gamma(t-u)\left(\alpha\,(\lambda_\infty - \lambda_u)\,du + \ \eta\,dL_u\right). \end{aligned} \tag{3.12}$$

This form is in fact inspired by the affine Volterra structure of Abi Jaber et al. (2019) with an Ornstein-Uhlenbeck (OU) drift coefficient. Such mML kernel with affine Volterra structure for the intensity process will allow us to obtain in the sequel a semi-closed form solution for the conditional transform of N and L. Moreover, from the spectral representation (3.6) and using Fubini theorem, we can rewrite the intensity (3.12) as

$$\begin{aligned}\lambda_t &= \lambda_0 + \int_0^t \left(\int_0^\infty e^{-\omega(t-u)}\mu(d\omega)\right)\left(\alpha\,(\lambda_\infty - \lambda_u)\,du + \ \eta\,dL_u\right) \\ &= \lambda_0 + \int_0^\infty \left(\int_0^t e^{-\omega(t-u)}\left(\alpha\,(\lambda_\infty - \lambda_u)\,du + \ \eta\,dL_u\right)\right)\mu(d\omega) \\ &= \lambda_0 + \int_0^\infty Y_t^{(\omega)}\mu(d\omega)\,. \end{aligned} \tag{3.13}$$

In the last line, we denote by $Y_t^{(\omega)}$ a Markov process satisfying $\forall \omega \in \mathbb{R}^+$ the SDE

$$dY_t^{(\omega)} = \left(-\omega Y_t^{(\omega)} + \alpha\,(\lambda_\infty - \lambda_t)\right)dt + \eta\, O dN_t\,, \quad Y_0^{(\omega)} = 0\,, \tag{3.14}$$

with solution

$$Y_t^{(\omega)} = \int_0^t e^{-\omega(t-u)}\left(\alpha(\lambda_\infty - \lambda_u)\,du + \eta\,dL_u\right). \tag{3.15}$$

From (3.13), we have that the intensity at time t is the sum of a background rate λ_0 and the expectation of arrival rates $\left(Y_t^{(\omega)}\right)_{\omega\in\mathbb{R}^+}$ with respect to a measure $\mu(d\omega)$. We see this way that the fractional Hawkes process defined by the mML kernel (3.1) can be written as a superposition of infinitely many Markov factors $\left(Y_t^{(\omega)}\right)_{\omega\in\mathbb{R}^+}$ mean reverting at different speeds ω. Hence, the mML fractional Hawkes process can be represented as an infinite-dimensional Markov process $\left(\lambda_t, \left(Y_t^{(\omega)}\right)_{\omega\in\mathbb{R}^+}, N_t\right)_{t\geq 0}$ which allows to use the standard tools of stochastic calculus to define its conditional transform using the affine property of equation (3.14). Note that a similar spectral representation has been derived in the context of rough volatility models by Abi Jaber (2019).

3.3. Discretization and conditional transform

Following Bäuerle and Desmettre (2020), we now approximate the measure $\mu(d\omega)$ derived in (3.7) as a discrete measure with a finite number of atoms, allowing to move from an infinitely dimensional Markov process to a framework with finite dimensional Markovianity. For this purpose, we consider a partition $\mathcal{E}^{(n)} := \left\{0 < \xi_0^{(n)} < \xi_1^{(n)} < \ldots < \xi_n^{(n)} < \infty\right\}$. The barycenter of μ on each interval $(\xi_i^{(n)}, \xi_{i+1}^{(n)})$ for $i = 0, \ldots, n-1$ is given by

$$b_{i+1}^{(n)} = \frac{\int_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} \omega\, \mu(d\omega)}{\int_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} \mu(d\omega)}\,. \tag{3.16}$$

Moreover, for $i = 0, \ldots, n-1$, on each interval $\left(\xi_i^{(n)}, \xi_{i+1}^{(n)}\right)$, the mass of any atom is

$$m_{i+1}^{(n)} = \int_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} \mu(d\omega) = \int_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} \frac{-\sin(\gamma\pi)\,\omega^{\gamma-1}(1+\omega)}{\pi\,(\omega^{2\gamma} - 2\omega^\gamma\cos(\gamma\pi) + 1)}\, d\omega\,.$$

We cannot obtain an analytical expression for $m_{i+1}^{(n)}$ and $b_{i+1}^{(n)}$. However, since

$$\frac{-\sin(\gamma\pi)}{\pi}\int \frac{\omega^{\gamma-1}}{\omega^{2\gamma} - 2\omega^\gamma\cos(\gamma\pi) + 1} d\omega = -\frac{1}{\gamma\pi}\arctan\left(\frac{\omega^\gamma - \cos(\gamma\pi)}{\sin(\gamma\pi)}\right).$$

and using integration by parts in which the integral $-\sin(\gamma\pi)\pi^{-1}\int \omega^\gamma\left(\omega^{2\gamma} - 2\omega^\gamma\cos(\gamma\pi) + 1\right)^{-1} d\omega$ is approached with the trapezoidal method, we obtain the approximation

$$\begin{aligned}
m_{i+1}^{(n)} \approx & -\frac{1}{\gamma\pi}\arctan\left(\frac{\left(\xi_{i+1}^{(n)}\right)^\gamma - \cos(\gamma\pi)}{\sin(\gamma\pi)}\right) + \frac{1}{\gamma\pi}\arctan\left(\frac{\left(\xi_i^{(n)}\right)^\gamma - \cos(\gamma\pi)}{\sin(\gamma\pi)}\right) \\
& - \left(\frac{\xi_{i+1}^{(n)} + \xi_i^{(n)}}{2\gamma\pi}\right)\left[\arctan\left(\frac{\left(\xi_{i+1}^{(n)}\right)^\gamma - \cos(\gamma\pi)}{\sin(\gamma\pi)}\right) - \arctan\left(\frac{\left(\xi_i^{(n)}\right)^\gamma - \cos(\gamma\pi)}{\sin(\gamma\pi)}\right)\right] \\
\approx & \left[-\frac{1}{\gamma\pi}\arctan\left(\frac{\left(\xi_{i+1}^{(n)}\right)^\gamma - \cos(\gamma\pi)}{\sin(\gamma\pi)}\right) + \frac{1}{\gamma\pi}\arctan\left(\frac{\left(\xi_i^{(n)}\right)^\gamma - \cos(\gamma\pi)}{\sin(\gamma\pi)}\right)\right]\left(1 + \frac{\xi_{i+1}^{(n)} + \xi_i^{(n)}}{2}\right).
\end{aligned}$$

A similar approximation can be obtained for the barycenter $b_{i+1}^{(n)}$, $i = 0, \ldots, n-1$. Anyway, the exact expressions for $b_i^{(n)}$ and $m_i^{(n)}$ are not useful in what follows. The discrete measure for a partition of size n is then defined as follows

$$\mu^{(n)}(\omega) := \sum_{i=1}^{n} m_i^{(n)}\delta_{b_i^{(n)}}(\omega)\,, \tag{3.17}$$

where $\delta_{b_i^{(n)}}(\omega)$ is the Dirac measure located at point ω such that $\lim_{n\to\infty}\mu^{(n)}(\omega) = \mu(\omega)$. We consider that the following assumption holds for the partition $\mathcal{E}^{(n)}$:

- $\xi_0^{(n)} \to 0$ and $\xi_n^{(n)} \to \infty$ when $n \to \infty$.
- $\max \left|\xi_{i+1}^{(n)} - \xi_i^{(n)}\right| \to 0$ when $n \to \infty$.
- $\mathcal{E}^{(n)} \subset \mathcal{E}^{(n+1)}$.

Considering the previous discretization setting, for any integrable function $f \in L^1(\mu)$ bounded on compact intervals, we have

$$\lim_{n\to\infty} \int_0^\infty f(\omega)\mu^{(n)}(d\omega) = \int_0^\infty f(\omega)\mu(d\omega)\,. \tag{3.18}$$

Detailed convergence results in this setting can be found in Lemma 3.1. and Theorem 3.2. of Bäuerle and Desmettre (2020). We choose n and adopt the following notation

$$Y_t^{(\omega)} := Y_t^{\left(b_\omega^{(n)}\right)},\ \omega = 1, \dots, n.$$

From equation (3.13), the discretized intensity based on a partition of size n is given by

$$\lambda_t^{(n)} = \lambda_0 + \sum_{\omega=1}^n m_\omega^{(n)} Y_t^{(\omega)}\,,$$

with $N_t^{(n)}$ a counting process of intensity $\lambda_t^{(n)}$ and

$$dY_t^{(\omega)} = \left(\alpha\left(\lambda_\infty - \lambda_t^{(n)}\right) - b_\omega^{(n)} Y_t^{(\omega)}\right) + \eta\, O\, dN_t^{(n)}\,.$$

From Bäuerle and Desmettre (2020), we indeed have that $\lambda_t^{(n)} \to \lambda_t$ almost surely for $n \to \infty$. Finally, we obtain

$$d\lambda_t^{(n)} = \sum_{\omega=1}^n m_\omega^{(n)} dY_t^{(\omega)} = \sum_{\omega=1}^n m_\omega^{(n)} \left(\alpha\left(\lambda_\infty - \lambda_t^{(n)}\right) - b_\omega^{(n)} Y_t^{(\omega)}\right) + \eta\, O \sum_{\omega=1}^n m_\omega^{(n)} \mathrm{d}N_t^{(n)}.$$

Now that we have decomposed the mML fractional Hawkes process into a finite number of related Markov processes, we are able to study its conditional transform.

Proposition 3.3. *The transform of* $X_t^{(n)} = \left(\lambda_t^{(n)}, J_t^{(n)}\right)^\top = \left(\lambda_t^{(n)}, \left(L_t^{(n)}, N_t^{(n)}\right)\right)^\top$ *conditionally to the information at time* $t \le T$ *with* $u = (u_\lambda, u_L, u_N) \in \mathbb{C}^3$ *is given by*

$$\begin{aligned}\mathbb{E}_t\left[e^{u\cdot X_T^{(n)}}\right] &= \mathbb{E}_t\left[e^{u_\lambda \lambda_T^{(n)} + u_L L_T^{(n)} + u_N N_T^{(n)}}\right] \\ &= \exp\left(a(t,T) + b(t,T)\lambda_t^{(n)} + u' \cdot {}^\top J_t^{(n)} + \sum_{\omega=1}^n c_\omega(t,T,b_\omega^{(n)})\, Y_t^{(\omega)}\right),\end{aligned} \tag{3.19}$$

where $u' = (u_L, u_N) \in \mathbb{C}^2$ *and where the coefficients* $a(t,T) := a(u',t,T)$, $b(t,T) := b(u',t,T)$ *and* $c_\omega(t,T,b_\omega^{(n)}) := c_\omega(u',t,T,b_\omega^{(n)})$ *satisfy the following system of equations*

$$\begin{aligned}a(t,T) &= \alpha\,\lambda_\infty \sum_{\omega=1}^n m_\omega^{(n)} \left(\int_t^T e^{-b_\omega^{(n)}(v-t)} b(v,T)dv\right), \\ \frac{\partial b(t,T)}{\partial t} &= \alpha\left(\sum_{\omega=1}^n m_\omega^{(n)} b(t,T) + \sum_{\omega=1}^n c_\omega(t,T,b_\omega^{(n)})\right) \\ &\quad - \theta\left(b(t,T)\,\eta \sum_{\omega=1}^n m_\omega^{(n)} + u'\cdot(1,0)^\top + \sum_{\omega=1}^n c_\omega(t,T,b_\omega^{(n)})\,\eta\right) e^{u'\cdot(0,1)^\top} + 1\,, \\ c_\omega(t,T,b_\omega^{(n)}) &= -b_\omega^{(n)}\, m_\omega^{(n)} \int_t^T e^{-b_\omega^{(n)}(v-t)}\, b(v,T)\, dv\,, \quad \forall \omega = 1, \dots n.\end{aligned}$$

with boundary condition $b(T,T) = u_\lambda$ *and where* $\theta(\xi) = \int e^{\xi o}\, d\nu(o)$, $\xi \in \mathbb{C}$ *is the transform of the jump distribution.*

The proof of this proposition is given in Appendix B.1. Passing to the limit leads to the following corollary.

Corollary 3.4. *The conditional transform of* $X_t = (\lambda_t, J_t)^\top = (\lambda_t, (L_t, N_t))^\top$ *conditionally to the information at time* $t \le T$ *with* $u = (u_\lambda, u_L, u_N) \in \mathbb{C}^3$ *is given by*

$$\mathbb{E}_t\left[e^{u\cdot X_T}\right] = \mathbb{E}_t\left[e^{u_\lambda \lambda_T + u_L L_T + u_N N_T}\right] = \tag{3.20}$$
$$\exp\left(\alpha\,\lambda_\infty \int_t^T f_\gamma(v-t)\, b(v,T)\, dv + b(t,T)\lambda_t + u'\cdot J_t^\top - \int_0^\infty \left(\int_t^T e^{-\omega(v-t)} b(v,T)\, dv\right) \omega\, Y_t^{(\omega)} \mu(d\omega)\right),$$

where $u' = (u_L, u_N) \in \mathbb{C}^2$ *and where the coefficient function* $b(t,T)$ *satisfies the following integro-differential equation*

$$\frac{\partial b(t,T)}{\partial t} = -\alpha\,\frac{\partial}{\partial t}\int_t^T f_\gamma(v-t)\, b(v,T)\, dv - \theta\left(u'\cdot(1,0)^\top - \eta\frac{\partial}{\partial t}\int_t^T f_\gamma(v-t)\, b(v,T)\, dv\right) e^{u'\cdot(0,1)^\top} + 1\,,$$

with boundary condition $b(T,T) = u_\lambda$.

The proof of this corollary is again relegated in Appendix B.2. Based on the conditional transform (3.20), we can now obtain the moments of the process $X = (\lambda, J)^\top$ given in the following corollary.

Corollary 3.5. *The conditional expectation of* X_T *at time* $T \ge t$ *is given by*

$$\mathbb{E}_t\left[w\cdot X_T\right] = \alpha\lambda_\infty \int_t^T f_\gamma(v-t)\, b'(0,v,T)\, dv + b'(0,t,T)\lambda_t + w'\cdot J_t \tag{3.21}$$
$$- \int_0^\infty \left(\int_t^T e^{-\omega(v-t)} b'(0,v,T)\, dv\right) \omega\, Y_t^{(\omega)} \mu(d\omega)\,,$$

for $w \in \mathbb{R}^3$. *We denote* $w' = (w_L, w_N) \in \mathbb{R}^2$ *and* $\chi = \int_0^\infty o\,\nu(do)$. *The coefficient function* $b'(0,t,T) := \left.\frac{\partial b(u',t,T)}{\partial u'}\right|_{u'=(0,0)}$ *satisfies*

$$b'(0,t,T) = w_\lambda + (\chi\eta - \alpha)\int_t^T f_\gamma(v-t)\, b'(0,v,T)\, dv + (T-t)\, w'\cdot(\chi, 1)^\top.$$

See Appendix B.3 for the proof of this corollary. Similarly, higher order moments and autocovariance function can be obtained by successive conditioning and differentiating.

3.4. A multivariate extension

We now want to extend the previous univariate setting by considering a multivariate mML fractional Hawkes process where the intensity $\lambda_t^{(i)}$ of each component process $i = 1,\ldots,d$ may also depend on several of the other components of this Hawkes process. Moreover, we now let each component $\lambda_t^{(i)}$ to be influenced by exogenous risk factors represented by a diffusive stochastic process. This is important for empirical applications, in which the intensity model is estimated from a time series of market prices or market measure (such as the Amihud measure studied in Section 4). In these applications, the model must replicate the diffusive fluctuation of market prices, requiring the presence of such diffusive risk factor. We hence consider the $\mathbb{R}^d$-valued counting process N defined on a state space $D_N = \mathbb{N}^d \times \mathbb{R}_+$ and the point process L defined on a state space $D_L = \mathbb{R}_+^d \times \mathbb{R}_+$

where the conditional mean $\lambda_t^{(i)}$ of the i^{th} loss component $L_t^{(i)} := \sum_{j\geq 1} O_j^{(i)} \mathbb{1}_{\{\tau_j^{(i)} \leq t\}}$ is given by the affine structure

$$\lambda_t^{(i)} = \lambda_0^{(i)} + \int_0^t f_{\gamma^{(i)}}(t-u) \left(\alpha^{(i)}(\lambda_\infty^{(i)} - \lambda_u^{(i)})\, du + \eta^{(i)} \cdot dL_u^\top + \sigma^{(i)} \sqrt{\lambda_u^{(i)}} dW_u^{(i)} \right) \tag{3.22}$$

where[4] $\eta^{(i)} = (\eta^{(i,1)}, \eta^{(i,2)} \ldots, \eta^{(i,d)})$ and where $W_t^{(i)}$ is a one-dimensional standard Brownian motion independent from the jump processes $L_t^{(i)}$ $(i = 1, \ldots, d)$ and from the other Brownian motions $W_t^{(j)}, \forall\, j \neq i$. The loss size $O^{(i)}$ of the i^{th} component $L_t^{(i)}$ follows a distribution $\nu^{(i)}$ on $(0, \infty)$. This specification enables the jump occurrence of one component process to have an impact on the intensities of other component processes, which facilitates the modeling of cross-excitation phenomena. For example, each component process $L_t^{(i)}$ can model the default process of a particular firm, where each default also has an impact on the default intensity of other firms. The setting (3.22) also allows to model peaks of (il)liquidity through the Amihud measure, as we will see in the next section. Finally, this multivariate extension is required at high-frequency in order to model both up and down mid changes in prices as well as the arrival of buy and sell orders, cfr Bacry and Muzy (2014) and Bacry et al. (2015).

As in Section 3.3, we can rewrite the conditional intensity (3.22) as

$$\lambda_t^{(i)} = \lambda_0^{(i)} + \int_0^\infty Y_t^{(i,\omega)}\, \mu^{(i)}(d\omega)\,, \tag{3.23}$$

where

$$\mu^{(i)}(d\omega) = \frac{-\sin(\gamma^{(i)}\pi)\, \omega^{\gamma^{(i)}-1}(1+\omega)}{\pi \left(\omega^{2\gamma} - 2\omega^{\gamma^{(i)}} \cos(\gamma^{(i)}\pi) + 1\right)}\, d\omega\,,$$

and

$$dY_t^{(i,\omega)} = \left(-\omega Y_t^{(i,\omega)} + \alpha^{(i)} \left(\lambda_\infty^{(i)} - \lambda_t^{(i)}\right)\right) dt + \eta^{(i)} \cdot dL_t^\top + \sigma^{(i)} \sqrt{\lambda_t^{(i)}} dW_t^{(i)}, \quad Y_0^{(i,\omega)} = 0\,. \tag{3.24}$$

From the affine structure of equation (3.22) and the representation of each conditional intensity in terms of infinitely many factors $Y_t^{(i,\omega)}$, we can apply directly the same reasoning as in the univariate case (3.20). The conditional transform of the process $Z = (\lambda, J)^\top = (\lambda, (L, N))^\top$ defined on a state space $D \subseteq \mathbb{R}^{3d} \times \mathbb{R}_+$ is then defined with $u_\lambda, u_L, u_N \in \mathbb{C}^d$ and $u = (u_\lambda, u_L, u_N) \in \mathbb{C}^{3d}$ by

$$\mathbb{E}_t \left[e^{u \cdot Z_T}\right] := \mathbb{E}_t \left[e^{u_\lambda \cdot \lambda_T^\top + u_L \cdot L_T^\top + u_N \cdot N_T^\top}\right],$$

and can be obtained as above via discretization and thanks to the infinitesimal generator of Z.

Proposition 3.6. *The transform of $Z_t = (\lambda_t, (L_t, N_t))^\top$ conditionally to the information at time $t \leq T$ with $u = (u_\lambda, u_L, u_N) \in \mathbb{C}^{3d}$ is given by*

$$\mathbb{E}_t \left[e^{u \cdot Z_T}\right] = \mathbb{E}_t \left[e^{u_\lambda \cdot \lambda_T^\top + u_L \cdot L_T^\top + u_N \cdot N_T^\top}\right] = \exp \Bigg[\beta(t,T) \cdot \lambda_t^\top + u_L \cdot L_t^\top + u_N \cdot N_t^\top \tag{3.25}$$

$$+ \sum_{i=1}^d \left(\alpha^{(i)} \lambda_\infty^{(i)} \int_t^T f_{\gamma^{(i)}}(v-t)\, \beta^{(i)}(v,T)\, dv - \int_0^\infty \left(\int_t^T e^{-\omega(v-t)} \beta^{(i)}(v,T)\, dv \right) \omega\, Y_t^{(i,\omega)} \mu^i(d\omega) \right) \Bigg]$$

[4]Note that we can only select $m \leq d$ components by setting $\eta^{(i,j)} = 0$ for some j.

where the boundary condition $\beta(T,T) = u_\lambda$ and where each component $\beta^{(i)}(t,T)$, $i = 1,\dots,d$, of the vector of coefficient functions $\beta(t,T)$ satisfies the following integro-differential equation

$$\begin{aligned}\frac{\partial \beta^{(i)}(t,T)}{\partial t} =& -\alpha^{(i)} \frac{\partial}{\partial t} \int_t^T f_{\gamma^{(i)}}(v-t)\, \beta^{(i)}(v,T)\, dv - \frac{1}{2} {\sigma^{(i)}}^2 \left(\frac{\partial}{\partial t} \int_t^T f_{\gamma(i)}(v-t)\, \beta^{(i)}(v,T)\, dv \right)^2 \\ & - \theta^{(i)} \left(u_{L,i} - \sum_{j=1}^{d} \eta^{(j,i)} \frac{\partial}{\partial_t} \int_t^T f_{\gamma^{(j)}}(v-t)\, \beta^{(j)}(v,T)\, dv \right) e^{u_{N,i}} + 1\,, \end{aligned} \tag{3.26}$$

where $\theta^{(i)}(\xi) = \int_0^\infty e^{\xi o}\, d\nu^{(i)}(o)$, $\xi \in \mathbb{C}$ is the transform of the loss size distribution of component i.

The proof of this proposition is relegated in Appendix B.4.

4. Illiquidity modeling

4.1. Amihud measure

Liquidity is a fundamental property of a well-functioning market and lack of liquidity is generally at the heart of many financial crises and disasters. Common ways of measuring liquidity using high frequency data include bid-ask spreads, effective spreads, realized spreads, depth and transaction volume. There is a big literature that uses such measures to compare market quality across time, sectors and before and after interventions of various sorts. Moreover, understanding and incorporating the behavior of liquidity in asset management and investment decision making is crucial for fund managers since their performance are strongly affected by the liquidity condition of their fund. It is indeed shown in Cao et al. (2013) that fund managers have the ability to time market liquidity and hence to adjust their portfolios' market exposure based on their insight of future liquidity conditions so as to bring substantially more return and performance to their fund. As proxy for illiquidity, we decide to focus on the Amihud illiquidity measure as proposed in Amihud (2002). This measure has proven to be very popular in the empirical literature. It is easy to implement and by all accounts relatively robust. It has been shown to influence the cross-sectional asset returns through the so-called illiquidity premium, see the review of Amihud and Mendelson (2015). This illiquidity premium indeed comes from the fact that investors care about illiquidity costs and therefore price them in the expected returns.

Many econometric models in discrete time have been studied in the literature for modeling illiquidity via the Amihud measure, cfr Amihud and Mendelson (2015), Brennan et al. (2013), Lou and Shu (2017) or more recently Hafner et al. (2022). We propose in this work a new approach by introducing instead a continuous-time process for modeling the log-Amihud measure. More precisely, we decide to model this log-Amihud measure as a mean-reverting jump diffusion process where the jumps follow a fractional Hawkes process driven by a mML intensity with the multidimensional affine structure (3.22). This way, we want to show that peaks of illiquidity in financial markets effectively exhibit long-memory and that there is contagion between the up and down jumps of the log-Amihud measure. This model also allows to study simultaneously the illiquidity process of multiple stocks and indices in financial markets. We will finally show that using such continuous-time process allows to easily perform risk management on illiquidity and to introduce and price financial instruments related to this Amihud measure.

The Amihud illiquidity measure A_t of a stock at time t is defined as

$$A_t = \frac{1}{n_t}\sum_{j=1}^{n_t} l_{t_j}\,, \quad l_{t_j} = \frac{|R_{t_j}|}{V_{t_j}} \tag{4.1}$$

where R_t is the stock return and V_t is the trading volume (in dollar) at time t. We therefore see that the Amihud measure captures the fact that a stock is less liquid if a given trading volume generates a larger move in its price. The measure is typically computed over periods ranging from a day to a year by averaging the daily illiquidity ratio l_{t_j} over the corresponding period n_t. We consider for the remaining of this work a daily sampling period with $n_t = 1$ and $\Delta t = t_j - t_{j-1} = 1/251\,,\ \forall j = 1,\ldots,n$ with $t_n = T$ and $t_0 = 0$. Due to the very small order of magnitude of the Amihud measure ($\sim 10^{-16} - 10^{-12}$ for the most popular indices), it is customary to work with the logarithm of the Amihud measure instead which allows to standardize its values, cfr Amihud (2002), Brennan et al. (2013), Lou and Shu (2017). We then denote $a_t := \log A_t$ in the sequel of the paper. Typical sample paths of the log-Amihud measure a_{t_j} and of $\Delta a_{t_j} = a_{t_j} - a_{t_j-1}$ $(j = 1,\ldots,n)$ are pictured for the FTSE 100 on Figure A.1 in Appendix A with $n = 3266$ daily observations over a total period T of 13 years. These plots show that Δa_t, the log-Amihud measure in difference, randomly fluctuates around zero and displays some clustering with serial dependence[5]. We also observe during these periods some shocks that do not display any clear trend: negative abrupt movements followed by abrupt movements of either sign. Finally, the Histogram A.2 of the log-Amihud increments Δa_t for the FTSE 100 shows that they are close to Gaussian. These observations justify to model the Amihud difference Δa_t as a mean-reverting Ornstein-Uhlenbeck (OU) process with two jump components (for up and down jumps) and to link the frequency of jumps to their size. Moreover, it is shown in Lou and Shu (2017) that the Amihud illiquidity measure is mainly driven by variations in the trading volume component. Since virtually all empirical studies show a significant predictive power of traded volume in explaining volatility/market activity (Ané and Geman (2000), Gallant et al. (1992), Karpoff (1987)), it confirms that it is judicious to model the log-Amihud measure in the same way as log-volatility, which is typically described by a mean-reverting OU process, see e.g. Andersen et al. (2001) or Gatheral et al. (2018). Adding jumps with a mML intensity finally allows to reproduce the observed illiquidity shocks as well as the long-memory behavior of the trading volume component, as highlighted in Lobato and Velasco (2000). We therefore here consider the following modified Mittag-Leffler Ornstein-Uhlenbeck (mML OU) process defined by the following dynamic of the infinitesimal increments of the log-Amihud measure $a_t = \log A_t$:

$$\begin{aligned} da_t &= \theta_1\,(\theta_2 - a_t)\,dt + \theta_3\,dW_t^{(a)} + dL_t^{(1)} - dL_t^{(2)} \\ &= \theta_1\,(\theta_2 - a_t)\,dt + \theta_3\,dW_t^{(a)} + O^{(1)}dN_t^{(1)} - O^{(2)}dN_t^{(2)}\,, \end{aligned} \tag{4.2}$$

where $L_t^{(1)}$ and $L_t^{(2)}$ are the loss processes defined by equation (2.1) of the up jump and down jump components, respectively. The conditional intensity of the jumps $N_t^{(1)}$, $N_t^{(2)}$ is given by equation (3.22) and $W_t^{(a)}$ is a standard Brownian motion, independent from the two jump components and from $W_t^{(i)}$ $(i = 1,2)$. Recall that the variable $O^{(i)}$ representing the jump size of each type is

[5]The fact that a OU process captures the same serial dependence as the log-Amihud measure a_t of the FTSE 100 index can be easily confirmed with ACF plots which exhibit the exact same pattern of autocorrelation (plots available upon request).

strictly positive with distribution $\nu^{(i)}(.)$ and support $(0, \infty)$. Note first that modeling the log-Amihud measure as a mML OU model does guarantee the positivity of the Amihud measure and allows an efficient pricing of liquidity-related instruments by providing semi-closed form formula for its conditional transform, as we will see in Section 4. Another feasible alternative would have been to model directly the Amihud measure A_t as a Cox-Ingersoll-Ross process with mML fractional Hawkes jumps but considering this time only up jumps in order to ensure the positivity of A_t. Indeed, what matters the most in practice is to manage the illiquidity peaks and hence, modeling only up-jumps would be satisfying enough. However, considering the log-Amihud measure provides a more general setting and gives more stable results since we avoid working with extremely small values. Secondly, this model (4.2) allows to study the peaks of (il)liquidity of various assets while having a dependence and contagion between up and down shocks of the log-Amihud measure. Finally, as already mentioned, our mML kernel (3.1) allows to capture the long-range dependence in these (il)liquidity shocks while ensuring good properties of the intensity process.

We now model the log-Amihud dynamic of the FTSE 100 index thanks to equation (4.2). Without loss of generality, we do not consider the Brownian components in the intensities $\lambda_t^{(1)}$ and $\lambda_t^{(2)}$ of the up and down jump processes $N_t^{(1)}$ and $N_t^{(2)}$ driving the log-Amihud measure. This leads to

$$\lambda_t^{(i)} = \lambda_0^{(i)} + \int_0^t f_{\gamma^{(i)}}(t-u) \left(\alpha^{(i)} (\lambda_\infty^{(i)} - \lambda_u^{(i)})\, du + \eta^{(i,1)} dL_u^{(1)} + \eta^{(i,2)} dL_u^{(2)} \right), \quad i = 1, 2. \tag{4.3}$$

Note that since $dL_t^{(2)} = O^{(2)}\, dN_t^{(2)}$ is strictly positive, we have indeed that a down jump increases the probability of a subsequent up jump (as well as the probability of another down jump). The largest the size of this down jump, the more likely is the occurrence of the following jump. For estimating this mML OU process, we use a modified Peaks Over Threshold (POT) procedure, as described in Hainaut and Moraux (2018). This method is divided into three steps. In the first step, from the time series of daily log-Amihud measures a_{t_j} $(j = 1, \ldots, n)$ of the FTSE 100 index, we estimate a OU process (*i.e.* the parameters $\theta_1,\ \theta_2,\ \theta_3$) via maximum likelihood without considering the jump size components $L_t^{(1)}$ and $L_t^{(2)}$. The second step consists in detecting jumps given that jumps occur when the log-Amihud measure is above (for up jumps) or below (for down jumps) some thresholds. These two thresholds $\mathrm{Tr}_j^{(1)}$ and $\mathrm{Tr}_j^{(2)}$ are given for each observation $j = 1, ..., n$ by

$$\mathrm{Tr}_j^{(1)} = a_{t_j} + \theta_1 \left(\theta_2 - a_{t_j} \right) \Delta t + \theta_3\, \phi^{-1}(\alpha_1) \sqrt{\Delta t}\,, \tag{4.4}$$

$$\mathrm{Tr}_j^{(2)} = a_{t_j} + \theta_1 \left(\theta_2 - a_{t_j} \right) \Delta t + \theta_3\, \phi^{-1}(\alpha_2) \sqrt{\Delta t}\,, \tag{4.5}$$

where $\phi^{-1}(\alpha_1)$ and $\phi^{-1}(\alpha_2)$ are Normal quantiles at levels α_1 and α_2. The levels α_1 and α_2 are chosen such that the sample without jumps $\left\{ a_{t_j} \mid a_{t_j} \in \left[\mathrm{Tr}_j^{(1)}\ ,\ \mathrm{Tr}_j^{(2)} \right] \right\}$ has a skewness and a kurtosis as closed as possible to the theoretical skewness and kurtosis of a OU process, which are respectively equal to 0 and 3 since the log-Amihud measure is Gaussian in this model :

$$\alpha_1,\, \alpha_2 = \arg\min \left(\left(Skew \left\{ a_{t_j} \mid a_{t_j} \in \left[\mathrm{Tr}_j^{(1)}\ ,\ \mathrm{Tr}_j^{(2)} \right] \right\} \right)^2 + \left(Kurt \left\{ a_{t_j} \mid a_{t_j} \in \left[\mathrm{Tr}_j^{(1)}\ ,\ \mathrm{Tr}_j^{(2)} \right] \right\} - 3 \right)^2 \right).$$

The set of the $n_T^{(1)}$ up jumps is then given by $\left\{ \tau_1^{(1)}, \ldots, \tau_{n_T^{(1)}}^{(1)} \right\} = \left\{ t_j \mid a_{t_j} > \mathrm{Tr}_j^{(1)} \right\}$ and the set of the $n_T^{(2)}$ down jumps by $\left\{ \tau_1^{(2)}, \ldots, \tau_{n_T^{(2)}}^{(2)} \right\} = \left\{ t_j \mid a_{t_j} < \mathrm{Tr}_j^{(2)} \right\}$. In the case of the FTSE 100, we find the following Figure 4.1 where we have in black the sample of log-Amihud differences Δa_t with jumps and in red the sample without jumps (with $\alpha_1 = 0.9695$ and $\alpha_2 = 0.0191$).

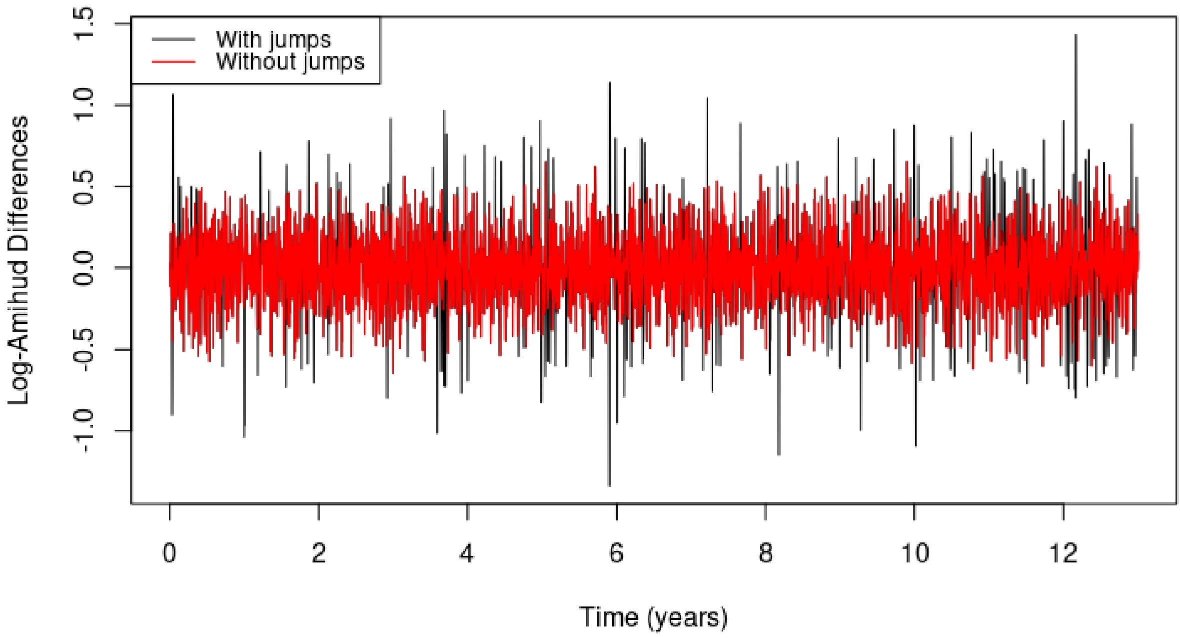

Figure 4.1: Daily log-Amihud measure in difference Δa_{t_j} $(j = 1, \ldots, 3266)$ for the FTSE 100 index over a period T of 13 years. In red, sample path without jumps and in black, sample path with jumps.

Considering these thresholds, we re-estimate the parameter θ_3 such that the sample without jumps follows a OU process, *i.e.* $\Delta a_{t_j} = \widehat{\theta}_1 \left(\widehat{\theta}_2 - a_{t_j} \right) \Delta t + \theta_3 W_\Delta$ for all $a_{t_j} \in \left[Tr_j^{(1)}, Tr_j^{(2)} \right]$ with $\widehat{\theta}_1$ and $\widehat{\theta}_2$ the estimators found at the previous calibration step. We now let $\left\{ \tau_1^{(i)}, \ldots, \tau_{n_T^{(i)}}^{(i)} \right\}$ the set of the $n_T^{(i)}$ observed up jumps if $i = 1$ and down jumps if $i = 2$ over the interval $[0, T]$. When a jump is detected at time $\tau_j^{(i)}$, the jump size $\Delta L_{\tau_j^{(i)}}^{(i)}$ is equal to the absolute value of the variation of the log-Amihud measure, *i.e.*

$$\Delta L_{\tau_j^{(i)}}^{(i)} = \left| \Delta a_{\tau_j^{(i)}} - \theta_1 \left(\theta_2 - a_{\tau_j^{(i)}-1} \right) \Delta t \right| .$$

We also assume that the jump sizes of up and down jumps are independent and identically distributed exponential random variables with density $\nu^{(i)}(z)$ on $\mathbb{R}^+$, $i = 1, 2$:

$$\nu^{(i)}(z) = \rho^{(i)} \exp\left(-\rho^{(i)} z \right), \tag{4.6}$$

with $\rho^{(i)} \geq 0$. The jump parameter $\rho^{(i)}$ is then estimated by log-likelihood maximization

$$\rho^{(i)} = \arg\max \sum_{l=1}^{n_T^{(i)}} \ln \nu \left(a_{\tau_l^{(i)}} - a_{\tau_l^{(i)}-1} - \theta_1 \left(\theta_2 - a_{\tau_l^{(i)}-1} \right) \Delta t \,\middle|\, \rho^{(i)} \right). \tag{4.7}$$

Based on the jump times $\left\{ \tau_1^{(i)}, \ldots, \tau_{n_T^{(i)}}^{(i)} \right\}$ of each type i, we can obtain via Embrechts et al. (2011) the log-likelihood of the mML fractional Hawkes process with intensity (4.3), which is given by

$$\log L \left(\sum_{i=1}^{d} \sum_{l=1}^{n_T^{(i)}} \tau_l^{(i)} \,\middle|\, \theta \right) = - \sum_{i=1}^{d} \int_0^T \lambda_s^{(i)} \, ds + \sum_{i=1}^{d} \sum_{l=1}^{n_T^{(i)}} \log \left(\lambda_{\tau_l^{(i)}}^{(i)} \right), \tag{4.8}$$

where $\theta = \left(\gamma^{(1)}, \alpha^{(1)}, \lambda_\infty^{(1)}, \lambda_0^{(1)}, \gamma^{(2)}, \alpha^{(2)}, \lambda_\infty^{(2)}, \lambda_0^{(2)}, \eta^{(1,1)}, \eta^{(1,2)}, \eta^{(2,1)}, \eta^{(2,2)}\right)$. For the exponential kernel (2.3), the expression (4.8) can be drastically simplified as in Errais et al. (2010), which speeds up computation. However, knowing the up and down jumps $\left\{\tau_1^{(i)}, \ldots, \tau_{n_T^{(i)}}^{(i)}\right\}_{i=1,2}$, as well as their jump size $\Delta L_{\tau_j^{(i)}}^{(i)}$ allows us to use the spectral representation (3.23) and (3.24) in order to compute efficiently their conditional intensity $\lambda_t^{(i)}$ and hence obtain the log-likelihood (4.8). Indeed, we can discretize for $i = 1, 2$ the integral (3.23) with a sum of $M + 1$ factors $\omega = 0, \Delta\omega, 2\Delta\omega, \ldots, M\Delta\omega$:

$$\lambda_{t_j}^{(i)} = \lambda_0^{(i)} + \sum_{k=0}^{M} Y_{t_j}^{(i,\,k\Delta\omega)} \mu^{(i)}(\Delta\omega)\,, \tag{4.9}$$

where $Y_0^{(i,k\Delta\omega)} = 0$ and for $j = 0, \ldots, n-1$:

$$Y_{t_{j+1}}^{(i,\,k\Delta\omega)} = Y_{t_j}^{(i,\,k\Delta\omega)} + \left(-k\Delta\omega\, Y_{t_j}^{(i,\,k\Delta\omega)} + \alpha^{(i)}\left(\lambda_\infty^{(i)} - \lambda_{t_j}^{(i)}\right)\right)\Delta t + \eta^{(i,1)}\Delta L_{t_j}^{(1)} + \eta^{(i,2)}\Delta L_{t_j}^{(2)}\,. \tag{4.10}$$

Taking $M = 50\,000$ with $\Delta\omega = 0.001$ gives very accurate results and allows to drastically decrease the computing time. A similar application of such technique with detailed convergence results can be found in Abi Jaber (2019). Note that since this discretized representation (4.9) cannot be written for the power-law kernel (2.9), the computation of the log-likelihood (4.8) is much more slow in this case. Finally, the estimated parameters of the mML intensity are then such that

$$\theta = \arg\max \log L\left(\sum_{i=1}^{d}\sum_{l=1}^{n_T^{(i)}} \tau_l^{(i)} \,\middle|\, \theta\right)$$

The full calibration procedure is summarized with the following table.

	OU	Std. Error	mML OU	Std. Error
θ_1	27.789	2.187	27.789	2.187
θ_2	-34.290	1.701	-34.290	1.701
θ_3	4.196	0.054	4.313	0.055
Log-likelihood	-117.678		-197.832	
$\rho^{(1)}$			1.545	0.177
$\rho^{(2)}$			1.464	0.165
Log-likelihood			-91.85	
$\gamma^{(1)}$			1.045	0.064
$\lambda_0^{(1)}$			1.611	0.68
$\gamma^{(2)}$			1.028	0.021
$\lambda_0^{(2)}$			1.719	0.64
$\eta^{(1,1)}$			14.229	5.063
$\eta^{(1,2)}$			0.168	0.221
$\eta^{(2,1)}$			2.265	1.012
$\eta^{(2,2)}$			19.799	7.762
Log-likelihood			136.188	

Table 4.1: Comparison of the parameter estimates and standard errors of the mML OU model with the classical OU model for the log-Amihud measure. This table is obtained via a POT procedure for the FTSE 100 index based on a 13-year period.

The first striking observation is that the mean-reverting parameters $\alpha^{(i)}$ and $\lambda_\infty^{(i)}$ $(i = 1, 2)$ of the intensity process are absent from the table above both for up and down jumps. This is due to the fact that the AIC criterion of the mML fractional Hawkes process is equal to -256.02 when considering the mean-reverting parameters $\alpha^{(i)}$ and $\lambda_\infty^{(i)}$ $(i = 1, 2)$ and equal to -256.38 without these parameters. We hence only consider here the simplified model without $\alpha^{(i)}$ and $\lambda_\infty^{(i)}$ (but these parameters could be relevant for other indices than the FTSE 100). We also see that the parameters $\gamma^{(1)}$, $\gamma^{(2)}$ capturing the long-memory of the (il)liquidity peaks in the log-Amihud measure are both very close to 1. Hence, this confirms the strong long-memory property of these illiquidity peaks. Moreover, based on equation (3.2), we have at large time scales the exact same power-law nature with exponent close to 1 for these illiquidity jumps as the one observed in Hardiman et al. (2013) or in Bacry and Muzy (2014) for mid-price changes and trade arrivals at high frequency based on the power-law kernel (2.9) with $\gamma \approx 0$ (and hence also with a power-law exponent close to $\gamma + 1 \approx 1$). Concerning the self-exciting parameters $\eta^{(1,1)}$ and $\eta^{(2,2)}$, we see that they are positive which confirms the presence of clustering effects in the intensity of jumps. Moreover, these parameters are higher than the cross-excitation parameters $\eta^{(1,2)}$ and $\eta^{(2,1)}$ (which are also positive) leading to a predominant effect of the self-excitation. Testing for stationarity in our bivariate setting (cfr Section 2), we obtain that the spectral radius $||.||$ of the matrix $\left(\eta^{(i,j)}\right)_{i,j=1,2}$ is equal to $||\phi|| = 19.87 > 1$. The stationary condition is therefore not satisfied based on our data. Finally, we see that we obtain high values of the speed of mean-reversion θ_2 and of the volatility θ_3. Computing the log-Amihud measure based on a weekly or monthly time window would allow us to obtain less extreme values. Another alternative frequently used in the literature consists in averaging this measure between different stocks/indices, as done in Amihud and Mendelson (2015). The large standard errors of the self and cross-exciting parameters in Table 4.1 are explained by the limited number of up and down jumps (rare events) used for estimating the model, which thus leads to numerical instability.

Our multidimensional mML fractional Hawkes model also allows to study simultaneously the illiquidity process of multiple stocks and indices in financial markets as well as their contagion. For each index $i \in \{1, \ldots, d\}$, the infinitesimal increments of the log-Amihud measure can be extended in the following way

$$da_t^{(i)} = \theta_1^{(i)} \left(\theta_2^{(i)} - a_t^{(i)}\right) dt + \theta_3^{(i)} \, dW_t^{(a,i)} + dL_t^{(i,up)} - dL_t^{(i,down)} ,$$

where $\left\{W_t^{(a,i)}\right\}_{i=1,\ldots,d}$ are independent standard Brownian motions (as well as independent from the jump components) and where the conditional intensity of the jump component $L_t^{(i,k)}$ of stock $i = 1, \ldots, d$ and type $k = \{up, down\}$ is given by

$$\lambda_t^{(i,k)} = \lambda_0^{(i,k)} + \int_0^t f_{\gamma^{(i,k)}}(t-u) \left(\alpha^{(i,k)}(\lambda_\infty^{(i,k)} - \lambda_u^{(i,k)})\, du + \eta^{(i,up)} \cdot dL_u^{(up)} + \eta^{(i,down)} \cdot dL_u^{(down)}\right),$$

where $\eta^{(i,k)} = (\eta^{(i,1,k)}, \ldots, \eta^{(i,d,k)})$ and $L^{(k)} \in \mathbb{R}^d$. Therefore, in this model, the dependence between the log-Amihud measure of different stocks lies in the peaks of their illiquidity process and not from their diffusive factor which can be considered as independent. A shock in the log-Amihud measure of one index has an impact on the intensity of the (up and down) jump component of the log-Amihud measure of the other indices. We can then apply the same POT procedure as above with the two thresholds (4.4) and (4.5) defined for each index i.

4.2. Option pricing

We now introduce and price two different types of European options on the Amihud measure so as to provide market participants with a way to reduce their exposition to the illiquidity risk of an index or a stock. These so-called illiquidity options also allow market operators to directly express views on illiquidity and hedge themselves again this risk. We first define a European call and put directly on the Amihud measure A_T at maturity T with the following payoff function

$$g(a_T) = (e^{a_T} - e^k)_+ = (A_T - K)_+ \quad \text{or} \quad g(a_T) = (e^k - e^{a_T})_+ = (K - A_T)_+\,, \tag{4.11}$$

where K is the liquidity strike and $k := \log K$. The valuation of an illiquidity call $C(.)$ is then classically given by the following formula under an equivalent risk neutral measure Q :

$$C(t,T,k) = \mathbb{E}^Q \left[e^{-r(T-t)} \left(A_T - K\right)_+ \Big| \mathcal{F}_t \right] \tag{4.12}$$

$$= \mathbb{E}^Q \left[e^{-r(T-t)} \left(e^{a_T} - e^k\right)_+ \Big| \mathcal{F}_t \right]. \tag{4.13}$$

A class of equivalent measures Q will be introduced in Propositions 4.1 and 4.2 below such as to compute this expectation under the risk-neutral measure Q. Based on the characteristic function of a_t which will be derived below in Corollary 4.3 with equation (4.27), we can use standard Fourier techniques to compute this European call, cfr Carr and Madan (1999) or Lian et al. (2014). Indeed, the discounted risk-neutral expectation of a contingent with general payoff $g(X_T)$ can be computed with $\alpha > 0$ as

$$\mathbb{E}^Q \left[e^{-r(T-t)} g(X_T) \Big| \mathcal{F}_t \right] = \frac{e^{-r(T-t)}}{\pi} \int_0^\infty \mathrm{Re} \left[G(\alpha + iu)\, \mathbb{E}^Q \left[e^{(\alpha+iu)X_T} | \, \mathcal{F}_t \right] \right] du\,, \tag{4.14}$$

where $\mathbb{E}^Q[e^{\xi X_T} | \, \mathcal{F}_t]$ $(\xi = \alpha + iu \in \mathbb{C})$ is the conditional characteristic function under Q of the log-Amihud random variable $X_T = a_T$, which will be denoted by $\varphi(t,T,\xi)$ in the sequel (cfr equation (4.27)). Moreover, $G(\xi)$ is the generalized Fourier transform of $g(X_T)$ and is defined for the payoff (4.11) above by the following equation with $\alpha > 1$:

$$G(\xi) := \int_{-\infty}^{\infty} e^{-\xi x}\, g(x)\, dx = \int_k^{\infty} e^{-\xi x}\, (e^x - e^k)\, dx = -\frac{e^{(1-\xi)k}}{(1-\xi)\xi} = -\frac{e^{(1-\alpha-iu)k}}{(1-\alpha-iu)(\alpha+iu)}\,.$$

By setting $\alpha' = \alpha - 1$ with $\alpha' > 0$, we find the classical Carr-Madan formula for the call price (4.12) :

$$C(t,T,k) = \frac{e^{-r(T-t)}\, e^{-\alpha' k}}{\pi} \int_0^\infty \mathrm{Re} \left[\frac{e^{-iuk}\, \varphi(t,T,\alpha'+1+iu)}{(\alpha')^2 + \alpha' - u^2 + i(2\alpha'+1)u} \right] du\,. \tag{4.15}$$

In a similar way as for variance swaps, we now introduce a second type of illiquidity derivative based on a new measure called the realized Amihud measure. From the definition of the realized variance, it appears natural to define the realized Amihud measure as the following arithmetic average

$$RA(n) = \frac{AF}{n} \sum_{j=1}^{n} A_{t_j}\,, \tag{4.16}$$

where AF is the annualization factor. The squared log-returns approximating the variability of the asset are indeed here replaced by the Amihud measure as proxy for illiquidity. The sampling frequency being every day trading, we have $AF = 251$ since there are 251 trading days for the FTSE 100 index. The number of observations n is again equal to 3266, which is consistent with a time horizon T of 13 years. We assume equally-spaced discrete observations over the contract life $[0, T]$ so that the annualized factor is equal to $AF = n/T$. When $n \to \infty$, we have then in the continuous case

$$RA(\infty) = \lim_{n\to\infty} RA(n) = \frac{1}{T}\int_0^T A_s\, ds\,.$$

However, the valuation of a European option based on this payoff requires the characteristic function of $\frac{1}{T}\int_0^T A_s\, ds = \frac{1}{T}\int_0^T e^{a_s} ds$, which does not admit a semi-closed form expression since we lose the affine structure of a_t due to the exponential function in this integral. We hence rather consider the following Asian payoff written on A_t based on a geometric average

$$RA(\infty) = \exp\left(\frac{1}{T}\int_0^T \log(A_s)\, ds\right) = \exp\left(\frac{1}{T}\int_0^T a_s\, ds\right).$$

We can then consider European-type options with the following payoff g for an illiquidity call (and respectively an illiquidity put) :

$$g(X_T) = \left(e^{X_T} - K_A\right)_+ = \left(e^{X_T} - e^{k_A}\right)_+ \quad \text{or} \quad g(X_T) = \left(K_A - e^{X_T}\right)_+ = \left(e^{k_A} - e^{X_T}\right)_+ ,$$

and where $X_T = \frac{1}{T}\int_0^T a_s\, ds$, where K_A is the liquidity strike and where $k_A = \log K_A$. The valuation of an illiquidity call $C_A(.)$ is then classically given by the following formula under an equivalent risk-neutral measure Q :

$$C_A(t, T, k_A) = \mathbb{E}^Q\left[e^{-r(T-t)}\left(\exp\left(\frac{1}{T}\int_0^T a_s\, ds\right) - e^{k_A}\right)_+ \middle| \mathcal{F}_t\right]. \tag{4.17}$$

Similarly, the formula for a European illiquidity put follows directly from (4.17).

We again use formula (4.14) to valuate this option with $X_T = \frac{1}{T}\int_0^T a_s\, ds$, which keeps an affine structure such that it is possible to compute its characteristic function. Hence, we have that $\varphi_A(t, T, \xi)$ now represents the characteristic function of the continuous arithmetic mean X_T of the log-Amihud measure (which will be also derived below in Corollary 4.3 with equation (4.25)). We then again have the classical Carr-Madan formula for the call price (4.17) with $\alpha' > 0$:

$$C_A(t, T, k_A) = \frac{e^{-r(T-t)}\, e^{-\alpha' k_A}}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-iuk_A}\,\varphi_A(t, T, \alpha' + 1 + iu)}{(\alpha')^2 + \alpha' - u^2 + i(2\alpha' + 1)u}\right] du\,. \tag{4.18}$$

4.3. Change of measure

Before computing the exact expressions for the characteristic functions $\varphi(t, T, \xi)$ and $\varphi_A(t, T, \xi)$, we need to define an equivalent risk-neutral measure that can be used by investors. We consider in this work the family of exponential affine changes of measure and we show that the dynamic of the Amihud measure is preserved under this class of measure for the mML fractional Hawkes process defined by the intensities $\lambda_t^{(i)}$ $(i = 1, ..., d)$ of equation (4.3). Note that without loss of generality, the Brownian components $\sqrt{\lambda_t^{(i)}}\, W_t^{(i)}$ in each intensity process (3.22) are again omitted in the sequel (cfr Zhang et al. (2009) for a detailed change of measure including them). These equivalent measures are induced by an exponential martingale of the form

$$M_t^{(n)}(\boldsymbol{\varepsilon}, \psi) := \exp\left(\sum_{i=1}^d \left(h^{(i)}(\varepsilon^{(i)})\int_0^t \lambda_s^{(i,n)}\, ds + \varepsilon^{(i)} L_t^{(i,n)}\right) - \frac{1}{2}\int_0^t \psi(s)^2\, ds - \int_0^t \psi(s)\, dW_s\right), \tag{4.19}$$

where $\psi(.)$ and $h^{(i)}(.)$ are suitable functions that will be defined later and $\boldsymbol{\varepsilon} = \left(\varepsilon^{(1)}, \ldots, \varepsilon^{(d)}\right)$. It is shown in Zhang et al. (2009) and in Hainaut (2016a) that $\left(M_t^{(n)}(\boldsymbol{\varepsilon}, \psi)\right)_{t\geq 0}$ is a local martingale provided that $h^{(i)}(\varepsilon^{(i)}) = 1 - \theta^{(i)}(\varepsilon^{(i)})$ for any parameter $\varepsilon^{(i)}$ and for $i = 1, ..., d$ (where $\theta^{(i)}(.)$ is again the transform of the loss size distribution of component i). When $M_t^{(n)}$ is actually a martingale (cfr Zhang et al. (2009)), we may define an equivalent (non-unique) probability measure Q by

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = \frac{M_t^{(n)}(\boldsymbol{\varepsilon}, \psi)}{M_0^{(n)}(\boldsymbol{\varepsilon}, \psi)} . \tag{4.20}$$

We now want to identify the dynamic of the intensities $\lambda_t^{Q,(i)}$ and of the log-Amihud measure a_t under Q. We will show that under the change of measure (4.20) their dynamic is preserved ($i.e$ a multivariate mML fractional Hawkes process for $\lambda_t^{Q,(i)}$ and a mML OU process for a_t).

Proposition 4.1. *Let $N_t^{Q,(i)}$, $i = 1, ..., d$ be counting processes with respective intensities defined under the equivalent measure Q by*

$$\lambda_t^{Q,(i)} := \theta^{(i)}(\varepsilon^{(i)})\, \lambda_t^{(i)} . \tag{4.21}$$

We then also denote $\lambda_\infty^{Q,(i)} := \theta^{(i)}(\varepsilon^{(i)})\, \lambda_\infty^{(i)}$, $\lambda_0^{Q,(i)} := \theta^{(i)}(\varepsilon^{(i)})\, \lambda_0^{(i)}$ and $\eta^{Q,(j,i)} := \theta^{(j)}(\varepsilon^{(j)})\, \eta^{(j,i)}$. On the other hand, if $O^{Q,(i)}$ denote the loss size of component i defined by the following jump transform

$$\theta^{Q,(i)}(\xi) = \mathbb{E}\left(e^{\xi\, O^{Q,(i)}}\right) := \frac{\theta^{(i)}\left(\xi + \varepsilon^{(i)}\right)}{\theta^{(i)}\left(\varepsilon^{(i)}\right)} , \tag{4.22}$$

and if $L_t^{Q,(i)}$, $i = 1, ..., d$ are defined by the jump processes

$$L_t^{Q,(i)} := \sum_{k=1}^{N_t^{Q,(i)}} O_k^{Q,(i)} ,$$

then, the intensities $\lambda_t^{Q,(i)}$ are driven by the following SDE under Q :

$$d\lambda^{Q,(i)} = \sum_{\omega=1}^{n} m_\omega^{(i,n)} dY_t^{Q(i,\omega)} ,$$

where

$$dY_t^{Q,(i,\omega)} = \left(-\omega Y_t^{Q,(i,\omega)} + \alpha^{(i)} \left(\lambda_\infty^{Q,(i)} - \lambda_t^{Q,(i)}\right)\right) dt + \sum_{j=1}^{d} \eta^{Q,(i,j)}\, O^{Q,(j)} dN_t^{Q,(j)} . \tag{4.23}$$

The proof of this proposition is given in Appendix B.5. We are now interested in the dynamic of the log-Amihud measure a_t under the measure Q. We again show that the OU dynamic is preserved provided a special form of the function $\psi(.)$.

Proposition 4.2. *If the market price of illiquidity risk is of the form $\psi(t) = \lambda$, for $\lambda \in \mathbb{R}$, then the dynamic of the log-Amihud measure under Q is given by*

$$da_t = \theta_1 \left(\widetilde{\theta}_2 - a_t\right) dt + \theta_3\, dW_t^{(a,Q)} + dL_t^{Q,(1)} - dL_t^{Q,(2)} , \tag{4.24}$$

where $\widetilde{\theta}_2 = \theta_2 - \frac{\lambda\theta_3}{\theta_1}$.

See Appendix B.6 for the proof of this proposition.

4.4. Characteristic function of the log-Amihud measure

We finally have to derive the characteristic function $\varphi(t,T,\xi)$ and $\varphi_A(t,T,\xi)$ of the random variables a_T and $\frac{1}{T}\int_0^T a_s\,ds$ under Q, which are defined respectively by $\varphi(t,T,\xi) := \mathbb{E}^Q\left[\exp\left(\xi\, a_T\right) \,|\, \mathcal{F}_t\right]$ and $\varphi_A(t,T,\xi) := \mathbb{E}^Q\left[\exp\left(\frac{\xi}{T}\int_0^T a_s\,ds\right) \Big|\, \mathcal{F}_t\right]$. We consider here the dynamic (4.2) proposed in the previous section where $i=1$ for up jumps and $i=2$ for down jumps. We can then apply exactly the same technique as for Proposition 3.6 and we therefore obtain the following corollary.

Corollary 4.3. *The characteristic function under Q of the random variable $X_T = \frac{1}{T}\int_0^T a_s\,ds$ conditionally to the information at time $t \le T$ with $\xi \in \mathbb{C}$ is given by*

$$\begin{aligned}
\varphi_A(t,T,\omega) &= \mathbb{E}^Q\left[\exp\left(\frac{\xi}{T}\int_0^T a_s\,ds\right) \Big|\, \mathcal{F}_t\right]\\
&= \exp\left(\frac{\xi}{T}\int_0^t a_s\,ds\right) \times \exp\Bigg[A(t,T) + B(t,T)\cdot {}^{\top}\lambda_t^Q + D(t,T)\,a_t \qquad (4.25)\\
&+ \sum_{i=1}^{2}\left(\alpha^{(i)}\lambda_\infty^{Q,(i)}\int_t^T f_{\gamma^{(i)}}(v-t)\,B^{(i)}(v,T)\,dv - \int_0^\infty\left(\int_t^T e^{-\omega(v-t)}B^{(i)}(v,T)\,dv\right)\omega\,Y_t^{Q,(i,\omega)}\mu^i(d\omega)\right)\Bigg],
\end{aligned}$$

where the dynamic of $Y_t^{Q,(i,\omega)}$ is given by equation (4.23). The boundary conditions are given by $A(T,T)=0$, $D(T,T)=0$ and $B(T,T)=(0,0)$. Each component $B^{(i)}(t,T)$ of the vector coefficient function $B(t,T)$ satisfies the following integro-differential equation

$$\begin{aligned}
\frac{\partial B^{(i)}(t,T)}{\partial t} &= -\,\alpha^{(i)}\frac{\partial}{\partial t}\int_t^T f_{\gamma^{(i)}}(v-t)\,B^{(i)}(v,T)\,dv \qquad (4.26)\\
&-\theta^{(i)}\left(-\sum_{j=1}^{2}\eta^{Q,(j,i)}\frac{\partial}{\partial_t}\int_t^T f_{\gamma^{(j)}}(v-t)\,B^{(j)}(v,T)\,dv + (-1)^{i-1}D(t,T)\right) + 1\ , \quad i=1,2.
\end{aligned}$$

The coefficient $D(t,T)$ is equal to

$$D(t,T) = -\frac{\xi\left(e^{\theta_1(t-T)}-1\right)}{T\theta_1}\,.$$

Finally, the coefficient $A(t,T)$ satisfies

$$\begin{aligned}
A(t,T) &= \int_t^T\left(\theta_1\widetilde{\theta}_2\,D(v,T) + \frac{\theta_3^2}{2}\,D(v,T)^2\right)dv\\
&= \frac{\widetilde{\theta}_2\,\xi}{T\theta_1}\left(e^{\theta_1(t-T)}-1\right) + \frac{\xi\,\widetilde{\theta}_2}{T}(T-t) + \frac{\theta_3^2\,\xi^2}{2T^2\theta_1^2}\left[\frac{1}{2\theta_1}\left(1-e^{2\theta_1(t-T)}\right) - \frac{2}{\theta_1}\left(1-e^{\theta_1(t-T)}\right) + (T-t)\right].
\end{aligned}$$

Similarly, the characteristic function of a_T at time $T \ge t$ under Q is defined by $\varphi(t,T,\xi) := \mathbb{E}^Q[e^{\xi\,a_T}\,|\,\mathcal{F}_t]$ with $\xi \in \mathbb{C}$ and has the same form as equation (4.25) :

$$\begin{aligned}
\varphi(t,T,\xi) =&\exp\Bigg[A(t,T) + B(t,T)\cdot {}^{\top}\lambda_t^Q + D(t,T)\,a_t \qquad (4.27)\\
&+ \sum_{i=1}^{2}\left(\alpha^{(i)}\lambda_\infty^{Q,(i)}\int_t^T f_{\gamma^{(i)}}(v-t)\,B^{(i)}(v,T)\,dv - \int_0^\infty\left(\int_t^T e^{-\omega(v-t)}B^{(i)}(v,T)\,dv\right)\omega\,Y_t^{Q,(i,\omega)}\mu^i(d\omega)\right)\Bigg].
\end{aligned}$$

The coefficient vector $B(t,T)$ also satisfies equation (4.26) with $B(T,T)=(0,0)$. However, the coefficient $D(t,T)$ now satisfies

$$D(t,T) = \xi\,e^{\theta_1(t-T)}\,,$$

and hence the function $A(t,T) = \int_t^T \left(\theta_1 \widetilde{\theta}_2 \, D(v,T) + \frac{\theta_3^2}{2} \, D(v,T)^2\right) dv$ *is equal to*

$$A(t,T) = \xi \, \widetilde{\theta}_2 \left[1 - e^{\theta_1 (t-T)}\right] + \frac{\theta_3^2 \, \xi^2}{4 \, \theta_1} \left[1 - e^{2\theta_1 (t-T)}\right].$$

Proof. This proposition is directly derived from the proof of Proposition 3.6 and from the PIDE (2.8). The infinitesimal generator of the discretized process $\left(\lambda^{Q,(n)}, \frac{\xi}{T} \int_\cdot^T a_s^{(n)} \, ds\right)$ under Q is now given by

$$\begin{aligned}\mathcal{L}f(z) = \sum_{i=1}^{2} \Bigg\{ & \frac{\partial f}{\partial \lambda^{Q,(i,n)}}(z) \sum_{\omega=1}^{n} m_\omega^{(i,n)} \left(\alpha^{(i)} \left(\lambda_\infty^{Q,(i)} - \lambda^{Q,(i,n)}\right) - b_\omega^{(n)} Y^{Q,(i,\omega)}\right) \\ & + \sum_{\omega=1}^{n} \frac{\partial f}{\partial Y^{Q,(i,\omega)}}(z) \left[\alpha^{(i)} \left(\lambda_\infty^{Q,(i)} - \lambda^{Q,(i,n)}\right) - b_\omega^{(n)} Y^{Q,(i,\omega)}\right] + \lambda^{Q,(i,n)} \int_0^\infty \Bigg[f\Bigg(\left\{ \lambda^{Q,(j,n)} + \eta^{Q,(j,i)} \, o \sum_{\omega=1}^{n} m_\omega^{(j,n)} \right\}_{j=1,2} \\ & , \left\{ Y^{Q,(j,\omega)} + \eta^{Q,(j,i)} o \right\}_{\substack{j=1,2 \\ \omega=1,\dots,n}}, a^{(n)} + (-1)^{i-1} o \Bigg) - f(z) \Bigg] \nu^{(i)}(do) \Bigg\} + \frac{\partial f}{\partial a^{(n)}}(z) \, \theta_1 (\widetilde{\theta}_2 - a) + \frac{1}{2} \frac{\partial^2 f}{\partial (a^{(n)})^2}(z) \, \theta_3^2\end{aligned} \tag{4.28}$$

where $z = \left(\lambda_t^{Q,(n)}, \frac{\xi}{T} \int_t^T a_s^{(n)} \, ds\right)$ and where $da_t^{(n)} = \theta_1 \left(\widetilde{\theta}_2 - a_t^{(n)}\right) dt + \theta_3 \, dW_t^{(a,Q)} + dL_t^{Q,(1,n)} - dL_t^{Q,(2,n)}$. As in Proposition 3.6, $L_t^{Q,(1,n)}$ and $L_t^{Q,(2,n)}$ are the discretized jump. The rest of the proof follows closely this proposition 3.6. The same justification is also used for the characteristic function of a_T. □

5. Results and discussion

5.1. Option pricing

From the theoretical results above, we are now ready to price options related to the illiquidity process of the FTSE 100 index via its Amihud illiquidity measure. More precisely, we focus here on European calls written on the realized Amihud measure $RA(n)$, defined by equation (4.16), since we are interested in the average behavior of the illiquidity process over some period of time (in a similar way as variance swaps). The results of this section can of course be easily extended to calls and puts directly on the Amihud measure A_T with formulae (4.12) and (4.15), instead.

Based on Propositions 4.1 and 4.2, we need to evaluate the parameters $\varepsilon^{(i)}$ $(i = 1,2)$ and λ in order to obtain an equivalent risk-neutral measure Q under which we can price our illiquidity options. These parameters should be calibrated based on observed illiquidity calls and puts. Indeed, we want to find the parameters $\varepsilon^{(1)}$, $\varepsilon^{(2)}$ and λ such that the model prices $C_A(0,T,k_A)$ match as best as possible to the observed market prices of these options (conditionally to the parameter estimates of Table 4.1 obtained from the econometric estimation under P). However, since these liquidity derivatives do not exist in financial markets, there is no risk-neutral measure available and we hence consider in this work a minimal change of measure by setting $\varepsilon^{(i)} = 0$ for $i = 1,2$ and $\lambda = 0$, which amounts to use the estimated parameters of Table 4.1.

Moreover, in order to evaluate the illiquidity call on the realized Amihud measure via formula (4.18), we need to obtain the characteristic function $\varphi_A(0,T,\xi)$ given by the expression (4.25). This requires to solve the integro-differential equation (4.26) satisfied by the coefficients $B^{(i)}(0,T)$. This equation is solved numerically using a finite difference scheme, which appears to give stable results for the FTSE 100 index. Since the order of magnitude of the Amihud measure is very small ($\sim 10^{-15}$ for the FTSE 100), we standardize the obtained call prices by multiplying them with a constant equal to $e^{-a_0} = 3.7 \times 10^{14}$ for a better readability. This leads to the following figures of illiquid call prices on the realized Amihud measure for different maturities and strike prices.

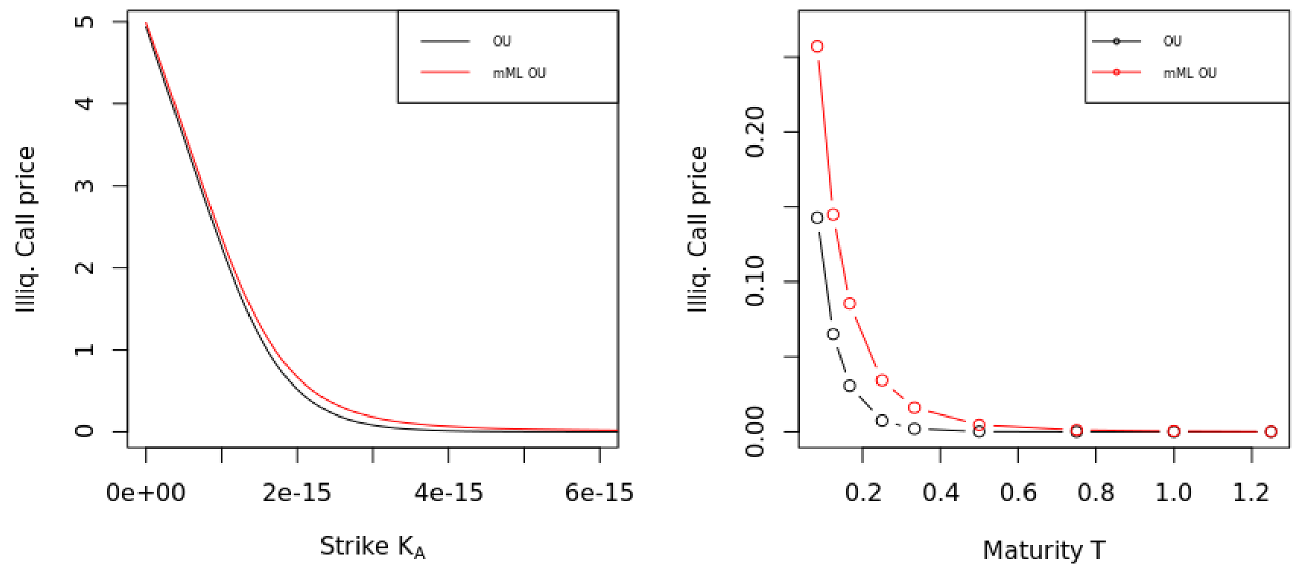

Figure 5.1: Left plot : Illiquid option prices $C_A(0, T, K_A)$ in function of strike prices K_A at maturity $T = 1/12$ (one month). Right plot : ATM illiquid option prices $C_A(0, T, e^{a_0})$ in function of the maturity T. Call prices in both figures are multiplied by the constant e^{-a_0} and are given for the classical OU model (black line) as well as for the mML OU model (red line).

We clearly see that the difference in illiquid option prices between the two models is increased for ATM options (around $K_A = e^{a_0} = 2.7 \times 10^{-15}$) and for short-maturity options. Indeed, due to the high speed of mean-reversion θ_1, the effect of jumps quickly vanishes when the maturity increases.

5.2. Risk Management

We first use Proposition 4.3 with the obtained transform (4.27) to compute the moments of the Amihud measure A_T at time T for the FTSE 100 as well as several risk measures. Such measures can indeed help various financial institutions to better manage and assess their risk of liquidity. They can also help fund managers to determine which stock/index to buy as well as the optimal timing to do so. The different centered moments can be obtained easily thanks to the conditional transform $\varphi(t, T, k) = \mathbb{E}_t\left[e^{k a_T}\right] = \mathbb{E}_t\left[A_T^k\right]$ with $k \in \mathbb{R}$. The transform $\varphi(t, T, \xi)$ is obtained thanks to equation (4.27) by replacing the risk-neutral parameters $(\widetilde{\theta}_2, \lambda_\infty^{(i,Q)}, \eta^{Q,(i,j)}, \lambda_0^{Q,(i)})$ with their real-world counterpart from Table 4.1.

Different risk measures for the Amihud measure A_T can also be computed via the probability density function (pdf) of a_T, which can be obtained efficiently via its Fourier representation. By setting $\xi = iz$, $z \in \mathbb{R}$ and by inverting the characteristic function

$$\mathbb{E}_t\left[e^{iza_T}\right] = \int_0^\infty e^{iza_1}\, p(T, a_1 \,|\, t, a_2)\, da_1\,,$$

it is indeed possible to determine the conditional probability density $p(T, a_1 \,|\, t, a_2)$ of a_T, which is defined by

$$p(T, a_1 \,|\, t, a_2) := \frac{\partial}{\partial a_1} P\left(a_T \le a_1 \mid a_t = a_2\right).$$

We can then compute numerically the Value-at-Risk (VaR) and Tail Value-at-Risk (TVaR). The VaR at level $\varepsilon \in (0,1)$ is defined as the shortfall value q such that $P(A_T \le q) = \varepsilon$. We can rewrite it in function of the pdf of a_T by

$$\int_{-\infty}^{\log q} p(T, a_1 \mid 0, a_2)\, da_1 = \varepsilon$$

We now use the following definition for the TVaR of the Amihud measure at level ε, $tv := \mathbb{E}\left[A_T | A_T > q\right]$, where q is the VaR at level ε. It can be rewritten in the following way

$$tv = \frac{1}{1-\varepsilon} \int_{\log q}^{+\infty} e^{a_1}\, p(T, a_1 \mid 0, a_2)\, da_1$$

The inversion of the Fourier transform is usually performed accurately and fast using a Discrete Fast Fourier Transform (DFFT) algorithm. Interested readers may refer to Dupret and Hainaut (2021) or Albanese et al. (2004) for a review. With the model parameters estimated from the POT procedure above in Table 4.1, we obtain in Table 5.1 the moments and risk measures of the Amihud measure A_T for $T = 1$ year. A comparison with the classical OU model based on the same parameters $(\theta_1, \theta_2, \theta_3)$ of Table 4.1 is also performed and the density of the Amihud and log-Amihud measures is depicted for each model in Figure 5.2 below.

	Expectation	Variance	Skewness	Kurtosis	VaR	TVaR
OU	1.52×10^{-15}	9.14×10^{-31}	2.14	12.12	5.69×10^{-15}	6.81×10^{-15}
mML OU	1.58×10^{-15}	1.82×10^{-30}	6.38	91.53	7.86×10^{-15}	1.24×10^{-14}

Table 5.1: The first columns of the table give the first four central moments of the Amihud measure A_T under the classical OU model and the mML OU model. The two last columns give the Value-at-Risk and Tail Value-at-Risk at level $\varepsilon = 99.5\%$.

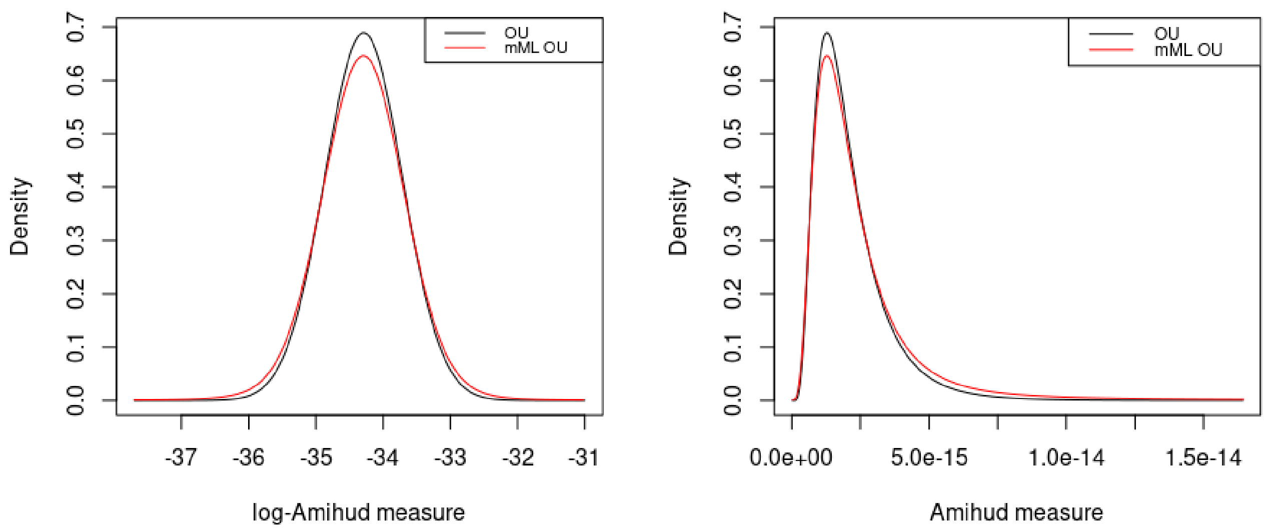

Figure 5.2: Density of the log-Amihud measure a_T (left) and of the Amihud measure A_T (right) for $T = 1$ year under the classical OU model (black line) and the mML OU model (red line).

We clearly see from Table 5.1 and Figure 5.2 that the mML OU model leads to a higher variability, a higher skewness and a higher kurtosis compared with the classical OU model. Since the intensity and size between up and down jumps are very close, the effect of the mML OU model is relatively small on the skewness of the obtained distributions. However, the presence of these jumps leads to a strong effect on the kurtosis, the VaR and the TVaR, making this mML OU model more risky.

6. Conclusion

This paper proposes a new procedure for modeling illiquidity in financial markets. More precisely, we introduce a new fractional Hawkes process in which the intensity process is ruled by a modified Mittag-Leffler excitation function. The nice properties of the mML kernel in terms of short and long-term asymptotics and in terms of L^1 norm makes this fractional Hawkes process better suited for the modeling of financial events with long-range dependence, compared with the existing self-exciting models (exponential, power-law, ML or ML's derivative Hawkes processes). Moreover, thanks to its spectral representation which admits a useful discretization in terms of a finite number of Markov processes, the mML kernel offers sufficient tractability to our fractional Hawkes process so as to have a characteristic function available in semi-closed form. Finally, a multivariate extension of the mML fractional Hawkes process can also be derived easily so as to study the contagion between the components of this Hawkes process. Therefore, this modified fractional Hawkes setting is perfectly suited in the context of illiquidity modeling.

More precisely, we consider in this work the Amihud illiquidity measure, one of the most widely used proxy in illiquidity modeling, which is known to influence the asset returns through a liquidity premium that compensates for price impact. We hence provide a new paradigm for illiquidity modeling based on a mean-reverting (OU) jump process for the log-Amihud measure where the jumps follow a mML fractional Hawkes process. This so-called mML OU model allows to simplify many computations and to develop new tools for managing and reducing the illiquidity risk, while reproducing the observed peaks of illiquidity in financial markets and their long-memory property. In particular, after having specified a suited change of measure preserving the dynamics of our mML OU model, we managed to introduce new illiquidity derivatives based on the Amihud measure and on the realized Amihud measure. Such illiquidity options allow market participants to express views on illiquidity and to better hedge against this risk. Moreover, having the characteristic function of this model in semi-closed form makes it also extremely convenient for risk management on illiquidity so as to better assess and manage this risk. Finally, after having estimated our mML OU model via a POT procedure on FTSE 100's historical data, we show how this model leads to a more risky distribution of the Amihud measure and how it impacts the price of illiquid option prices, compared with the classical OU model.

In further research, we need to verify and generalize these results for a larger number of assets and indices while including a diffusive factor in their intensity process. We could also apply and test this new fractional Hawkes process in the context of credit risk portfolio, for option pricing in jump diffusion models and for the modeling of trade orders at high-frequency.

Appendix A. Figures

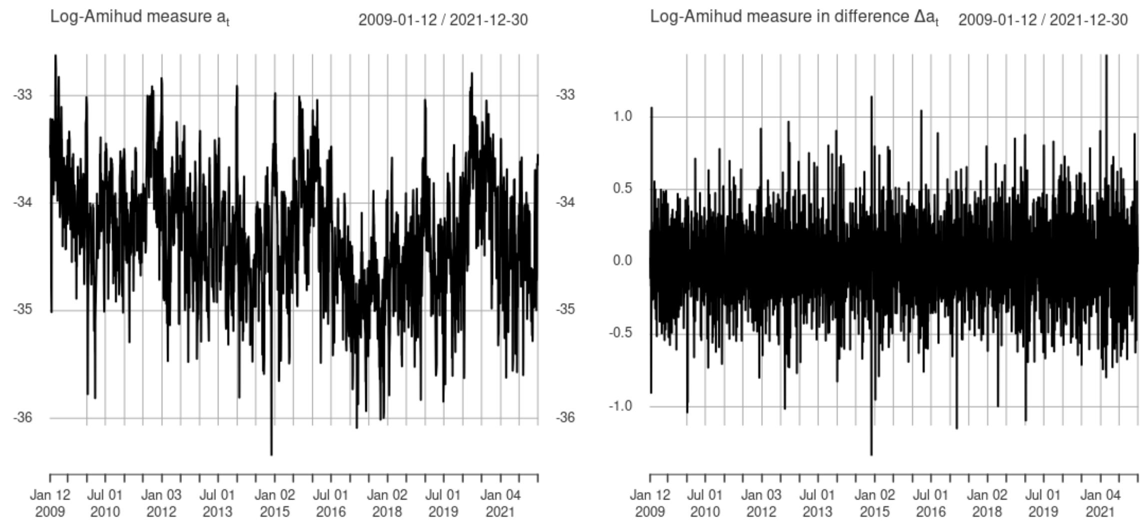

Figure A.1: Daily log-Amihud measure in level a_{t_j} (left) and in difference Δa_{t_j} (right) with $j = 1, \ldots, 3266$ for the FTSE 100 index over a period T of 13 years.

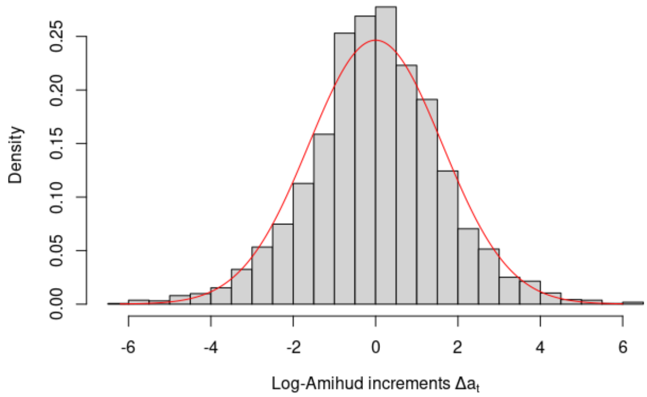

Figure A.2: Histogram of the daily log-Amihud increments Δa_{t_j} ($j = 1, \ldots, 3266$) for the FTSE 100 index based on a 13-year period with the Gaussian fit on these data superimposed in red.

Appendix B. Proofs

Appendix B.1. Proof of Proposition 3.3.

We first express the conditional transform $\mathbb{E}_t\left[e^{u\cdot X_T^{(n)}}\right] = \mathbb{E}_t\left[e^{u_\lambda \lambda_T^{(n)} + u_L L_T^{(n)} + u_N N_T^{(n)}}\right]$ as $f\left(t, T, X_t^{(n)}, \left(Y_t^{(\omega)}\right)_{\omega=1,\ldots,n}\right)$. Moreover, the infinitesimal generator of $X^{(n)}$ is given by

$$\mathcal{L}f(x) = \frac{\partial f}{\partial \lambda^{(n)}}(x) \sum_{\omega=1}^{n} m_\omega^{(n)} \left(\alpha\left(\lambda_\infty - \lambda^{(n)}\right) - b_\omega^{(n)} Y^{(\omega)}\right) + \sum_{\omega=1}^{n} \frac{\partial f}{\partial Y^{(\omega)}}(x) \left(\alpha\left(\lambda_\infty - \lambda^{(n)}\right) - b_\omega^{(n)} Y^{(\omega)}\right)$$
$$+ \lambda^{(n)} \int_0^\infty \left(f\left(\lambda^{(n)} + \eta\, o \sum_{\omega=1}^{n} m_\omega^{(n)},\ L^{(n)} + o,\ N^{(n)} + 1, \left(Y^{(\omega)} + \eta o\right)_{\omega=1,\ldots,n}\right) - f(x)\right) \nu(do)\,, \quad \text{(B.1)}$$

where $X_t^{(n)} = x = (\lambda_t^{(n)}, (L_t^{(n)}, N_t^{(n)}))^\top$. Since f must again satisfy the PIDE (2.8) with boundary condition $f(T, T, X_T^{(n)}) = e^{u\cdot X_T^{(n)}} = e^{u_\lambda \lambda_T^{(n)} + u_L L_T^{(n)} + u_N N_T^{(n)}}$, we have that the conditional transform of the point process $X_t^{(n)}$ is given by the exponential function

$$f\left(t,T,X_t^{(n)},\left(Y_t^{(\omega)}\right)_{\omega=1,\dots,n}\right)=\exp\left(a(t,T)+b(t,T)\lambda_t^{(n)}+u'\cdot{}^{\top}J_t^{(n)}+\sum_{\omega=1}^{n}c_\omega(t,T,b_\omega^{(n)})\,Y_t^{(\omega)}\right),$$

where $u'=(u_L,u_N)$. The coefficient functions $a(t,T):=a(u',t,T)$ and $b(t,T):=b(u',t,T)$ are time dependent functions with boundary conditions $a(T,T)=0$ and $b(T,T)=u_\lambda$. The function $c_\omega(t,T,b_\omega^{(n)}):=c_\omega(u',t,T,b_\omega^{(n)})$ is a function that depends again on time but also on the position $\omega=1,\dots,n$, within the partition $\mathcal{E}^{(n)}$, and which satisfies the boundary condition $c_\omega(T,T,b_\omega^{(n)})=0$, $\forall\omega\in 1,\dots,n$. Differentiating f with respect to t, $\lambda^{(n)}$ and $Y^{(\omega)}$ leads to

$$\frac{\partial f}{\partial t}=f\left(\frac{\partial a(t,T)}{\partial t}+\frac{\partial b(t,T)}{\partial t}\lambda^{(n)}+\sum_{\omega=1}^{n}\frac{\partial c_\omega(t,T,b_\omega^{(n)})}{\partial t}\,Y^{(\omega)}\right),$$

$$\frac{\partial f}{\partial\lambda^{(n)}}=f\,b(t,T)\,,\quad\frac{\partial f}{\partial Y^{(\omega)}}=f\,c_\omega(t,T,b_\omega^{(n)})\quad\forall\omega=1,\dots,n,$$

and the variation of f due to the occurrence of a jump of size o is given by

$$f\,\left(e^{b(t,T)\,\eta\,o\sum_{\omega=1}^{n}m_\omega^{(n)}+u'\cdot(1,0)^\top o+u'\cdot(0,1)^\top+\sum_{\omega=1}^{n}c_\omega(t,T,b_\omega^{(n)})\,\eta\,o}-1\right).$$

Injecting these equations in (B.1) and in the PIDE (2.8), we have

$$\begin{aligned}0=&\frac{\partial a(t,T)}{\partial t}+\frac{\partial b(t,T)}{\partial t}\lambda^{(n)}+\sum_{\omega=1}^{n}\frac{\partial c_\omega(t,T,b_\omega^{(n)})}{\partial t}\,Y^{(\omega)}+b(t,T)\sum_{\omega=1}^{n}m_\omega^{(n)}\left(\alpha\left(\lambda_\infty-\lambda^{(n)}\right)-b_\omega^{(n)}Y^{(\omega)}\right)\\&+\sum_{\omega=1}^{n}c_\omega(t,T,b_\omega^{(n)})\left(\alpha\left(\lambda_\infty-\lambda^{(n)}\right)-b_\omega^{(n)}Y^{(\omega)}\right)\\&+\lambda^{(n)}\left[\left(\int e^{b(t,T)\,\eta\,o\sum_{\omega=1}^{n}m_\omega^{(n)}+u'\cdot(1,0)^\top o+\sum_{\omega=1}^{n}c_\omega(t,T,b_\omega^{(n)})\,\eta\,o}\,\nu(do)\right)e^{u'\cdot(0,1)^\top}-1\right].\end{aligned}$$

From which we find the following system of PDE

$$\begin{aligned}\frac{\partial a(t,T)}{\partial t}&=-b(t,T)\sum_{\omega=1}^{n}m_\omega^{(n)}\,\alpha\,\lambda_\infty-\sum_{\omega=1}^{n}c_\omega(t,T,b_\omega^{(n)})\,\alpha\,\lambda_\infty\,,\\\frac{\partial b(t,T)}{\partial t}&=b(t,T)\sum_{\omega=1}^{n}m_\omega^{(n)}\alpha+\sum_{\omega=1}^{n}c_\omega(t,T,b_\omega^{(n)})\,\alpha\\&\quad-\theta\left(b(t,T)\,\eta\sum_{\omega=1}^{n}m_\omega^{(n)}+u'\cdot(1,0)^\top+\sum_{\omega=1}^{n}c_\omega(t,T,b_\omega^{(n)})\,\eta\right)e^{u'\cdot(0,1)^\top}+1\,,\\\frac{\partial c_\omega(t,T,b_\omega^{(n)})}{\partial t}&=b(t,T)\,m_\omega^{(n)}\,b_\omega^{(n)}+c_\omega(t,T,b_\omega^{(n)})\,b_\omega^{(n)}\,,\quad\forall\omega=1,\dots n,\end{aligned}$$

where $\theta(.)$ is again the transform of the jump distribution. The previous system of equation finally gives with $a(T,T)=0$ and $c_\omega(T,T,b_\omega^{(n)})=0$:

$$\begin{aligned}a(t,T)&=\alpha\,\lambda_\infty\int_t^T\left(\sum_{\omega=1}^{n}m_\omega^{(n)}\,b(v,T)+\sum_{\omega=1}^{n}c_\omega(v,T,b_\omega^{(n)})\right)dv\,,\\\frac{\partial b(t,T)}{\partial t}&=\alpha\left(\sum_{\omega=1}^{n}m_\omega^{(n)}b(t,T)+\sum_{\omega=1}^{n}c_\omega(t,T,b_\omega^{(n)})\right)\\&\quad-\theta\left(b(t,T)\,\eta\sum_{\omega=1}^{n}m_\omega^{(n)}+u'\cdot(1,0)^\top+\sum_{\omega=1}^{n}c_\omega(t,T,b_\omega^{(n)})\,\eta\right)e^{u'\cdot(0,1)^\top}+1\,,\\c_\omega(t,T,b_\omega^{(n)})&=-b_\omega^{(n)}\,m_\omega^{(n)}\int_t^T e^{-b_\omega^{(n)}(v-t)}\,b(v,T)\,dv\,,\quad\forall\omega=1,\dots n,\end{aligned}$$

and $a(t,T)$ can be rewritten

$$\begin{aligned} a(t,T) &= \alpha\,\lambda_\infty \sum_{\omega=1}^{n} m_\omega^{(n)} \left(\int_t^T b(v,T)\,dv - b_\omega^{(n)} \int_t^T \int_v^T e^{-b_\omega^{(n)}(s-v)}\, b(s,T)\,ds\,dv \right) \\ &= \alpha\,\lambda_\infty \sum_{\omega=1}^{n} m_\omega^{(n)} \left(\int_t^T b(v,T)dv - b_\omega^{(n)} \int_t^T \frac{1}{b_\omega^{(n)}} \left(1 - e^{-b_\omega^{(n)}(s-t)}\right) b(s,T)\,ds \right) \\ &= \alpha\,\lambda_\infty \sum_{\omega=1}^{n} m_\omega^{(n)} \left(\int_t^T e^{-b_\omega^{(n)}(v-t)} b(v,T)\,dv \right). \end{aligned}$$

□

Appendix B.2. Proof of Corollary 3.4.

The conditional transform (3.20) is a direct consequence of the approximations (3.17) and (3.18). For deriving the coefficient function $b(t,T)$, we first note from equation (3.6) that

$$\lim_{n\to\infty} \sum_{\omega=1}^{n} m_\omega^{(n)} = \int_0^\infty \mu(d\omega) = f_\gamma(0+) = 1\,. \tag{B.2}$$

Hence, we find almost surely in the limit $n \to \infty$:

$$\begin{aligned} \frac{\partial b(t,T)}{\partial t} &= \alpha \left(b(t,T) - \int_0^\infty \left(\int_t^T e^{-\omega(v-t)} b(v,T)\,dv \right) \omega\,\mu(d\omega) \right) \\ &\quad - \theta \left(b(t,T)\,\eta + u' \cdot (1,0)^\top - \eta \int_0^\infty \left(\int_t^T e^{-\omega(v-t)} b(v,T)\,dv \right) \omega\,\mu(d\omega) \right) e^{u'\cdot(0,1)^\top} + 1\,. \end{aligned}$$

Then, from equation (3.6), we have

$$\begin{aligned} \frac{\partial b(t,T)}{\partial t} &= \alpha \left(b(t,T) - \int_t^T \frac{\partial f_\gamma(v-t)}{\partial t}\, b(v,T)\,dv \right) \\ &\quad - \theta \left(b(t,T)\,\eta + u' \cdot (1,0)^\top - \eta \int_t^T \frac{\partial f_\gamma(v-t)}{\partial t}\, b(v,T)\,dv \right) e^{u'\cdot(0,1)^\top} + 1\,, \end{aligned}$$

and finally from Leibniz integral rule to

$$\frac{\partial b(t,T)}{\partial t} = -\alpha\,\frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t)\,b(v,T)\,dv - \theta \left(u' \cdot (1,0)^\top - \eta \frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t) b(v,T)\,dv \right) e^{u'\cdot(0,1)^\top} + 1\,.$$

□

Appendix B.3. Proof of Corollary 3.5.

By differentiating the conditional transform (3.20) with respect to u and evaluating the derivative at $u = 0$, we find the following conditional expectation

$$\begin{aligned} \mathbb{E}_t\left[w \cdot X_T\right] &= \alpha\lambda_\infty \int_t^T f_\gamma(v-t)\,b'(0,v,T)\,dv + b'(0,t,T)\lambda_t + w' \cdot J_t \\ &\quad - \int_0^\infty \left(\int_t^T e^{-\omega(v-t)} b'(0,v,T)\,dv \right) \omega\, Y_t^{(\omega)} \mu(d\omega)\,, \end{aligned}$$

where $b'(0,t,T)$ is the derivative of the coefficient function $b(u',t,T)$ evaluated at $(0,0)$ and which satisfies

$$\begin{aligned} \frac{\partial b'(0,t,T)}{\partial t} &= -\alpha\,\frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t)\,b'(0,v,T)\,dv - \theta \left(-\eta \frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t)\,b(0,v,T)\,dv \right) w' \cdot (0,1)^\top \\ &\quad - \theta' \left(-\eta \frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t)\,b(0,v,T)\,dv \right) \left(w' \cdot (1,0)^\top - \eta \frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t)\,b'(0,v,T)\,dv \right), \end{aligned}$$

with boundary condition $b'(0,T,T) = w_\lambda$ and where $\theta'(\xi) = \int o\, e^{\xi o} d\nu(o)$ is the derivative of the jump transform. Since $u = 0$, we can choose $b(0,t,T) = 0$, $a(0,t,T) = 0$ and $c_\omega(0,t,T,b_\omega^{(n)}) = 0$ thanks to equation (3.19) and hence, the previous equation simplifies to

$$\begin{aligned}\frac{\partial b'(0,t,T)}{\partial t} &= -\alpha \frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t)\, b'(0,v,T)\, dv - \theta(0)\, w' \cdot (0,1)^\top \\ &\quad - \theta'(0) \left(w' \cdot (1,0)^\top - \eta \frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t)\, b'(0,v,T)\, dv \right).\end{aligned}$$

We then have

$$\begin{aligned}\frac{\partial b'(0,t,T)}{\partial t} &= -\alpha \frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t)\, b'(0,v,T)\, dv + \chi\eta \frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t)\, b'(0,v,T)\, dv - w' \cdot (\chi,1)^\top \\ &= (\chi\eta - \alpha) \frac{\partial}{\partial_t} \int_t^T f_\gamma(v-t)\, b'(0,v,T)\, dv - w' \cdot (\chi,1)^\top,\end{aligned} \tag{B.3}$$

where $\chi = \int o\, d\nu(o)$, which finally leads to

$$b'(0,t,T) = w_\lambda + (\chi\eta - \alpha) \int_t^T f_\gamma(v-t)\, b'(0,v,T)\, dv + (T-t)\, w' \cdot (\chi,1)^\top.$$

□

Appendix B.4. Proof of Proposition 3.6.

As for Proposition 3.3, we start by studying the discretized version of the conditional intensity (3.23) and hence, we have the following dynamics

$$dY_t^{(i,\omega)} = \left(-b_\omega^{(n)} Y_t^{(i,\omega)} + \alpha^{(i)} \left(\lambda_\infty^{(i)} - \lambda_t^{(i,n)}\right)\right) dt + \sum_{j=1}^d \eta^{(i,j)}\, O^{(j)} dN_t^{(j,n)} + \sigma^{(i)} \sqrt{\lambda_t^{(i,n)}} dW_t^{(i)},$$

and

$$d\lambda_t^{(i,n)} = \sum_{\omega=1}^n m_\omega^{(i,n)} dY_t^{(i,\omega)}.$$

The infinitesimal generator of $Z^{(n)} = \left(\lambda^{(n)}, J^{(n)}\right)^\top$ is given by

$$\begin{aligned}\mathcal{L}f(z) = \sum_{i=1}^d \Bigg\{ & \frac{\partial f}{\partial \lambda^{(i,n)}}(z) \sum_{\omega=1}^n m_\omega^{(i,n)} \left(\alpha^{(i)} \left(\lambda_\infty^{(i)} - \lambda^{(i,n)}\right) - b_\omega^{(n)} Y^{(i,\omega)} \right) + \sum_{\omega=1}^n \frac{\partial f}{\partial Y^{(i,\omega)}}(z) \Big[\alpha^{(i)} \left(\lambda_\infty^{(i)} - \lambda^{(i,n)}\right) \\ & - b_\omega^{(n)} Y^{(i,\omega)} \Big] + \frac{1}{2} \lambda^{(i,n)} {\sigma^{(i)}}^2 \left(\sum_{\omega=1}^n m_\omega^{(i,n)} \right)^2 \frac{\partial^2 f}{\partial {\lambda^{(i,n)}}^2}(z) + \frac{1}{2} \lambda^{(i,n)} {\sigma^{(i)}}^2 \sum_{\omega_1=1,\omega_2=1}^n \frac{\partial^2 f}{\partial Y^{(i,\omega_1)} \partial Y^{(i,\omega_2)}}(z) \\ & + \lambda^{(i,n)} \int_0^\infty \Bigg[f \Bigg(\left\{ \lambda^{(j,n)} + \eta^{(j,i)}\, o \sum_{\omega=1}^n m_\omega^{(j,n)} \right\}_{j=1,\ldots,d}, \left\{ Y^{(j,\omega)} + \eta^{(j,i)} o \right\}_{\substack{j=1,\ldots,d \\ \omega=1,\ldots,n}}, L^{(i,n)} + o, N^{(i,n)} + 1 \Bigg) \\ & - f(z) \Bigg] \nu^{(i)}(do) + \lambda^{(i,n)} {\sigma^{(i)}}^2 \left(\sum_{\omega=1}^n m_\omega^{(i,n)} \right) \sum_{\omega=1}^n \frac{\partial^2 f}{\partial \lambda^{(i,n)} \partial Y^{(i,\omega)}}(z) \Bigg\},\end{aligned} \tag{B.4}$$

with $Z_t^{(n)} = z = (\lambda_t^{(n)}, J_t^{(n)})^\top$ and for f a function with continuous partial derivatives in the domain of the infinitesimal generator of $Z^{(n)}$. We then express the conditional transform as

$$f\left(t, T, Z_t^{(n)}, \left\{Y^{(i,\omega)}\right\}_{\substack{i=1,\ldots,d \\ \omega=1,\ldots,n}}\right) := \mathbb{E}_t\left[e^{u \cdot Z_T^{(n)}}\right] = \mathbb{E}_t\left[e^{u_\lambda \cdot^\top \lambda_T^{(n)} + u_L \cdot^\top L_T^{(n)} + u_N \cdot^\top N_T^{(n)}}\right].$$

Since f must satisfy the PIDE (2.8) with boundary condition $f(T,T,Z_T^{(n)}) = e^{u_\lambda \cdot {}^\top\lambda_T^{(n)} + u_L \cdot {}^\top L_T^{(n)} + u_N \cdot {}^\top N_T^{(n)}}$, we have that the transform of $Z^{(n)}$ is given by :

$$f\left(t,T,Z_t^{(n)},\left\{Y^{(i,\omega)}\right\}_{\substack{i=1,\dots,d\\ \omega=1,\dots,n}}\right)$$
$$= \exp\left(\alpha(t,T) + \beta(t,T)\cdot {}^\top\lambda_t^{(n)} + u_L \cdot {}^\top L_t^{(n)} + u_N \cdot {}^\top N_t^{(n)} + \sum_{i=1}^{d}\sum_{\omega=1}^{n} c_\omega^{(i)}(t,T,b_\omega^{(n)})\, Y_t^{(i,\omega)}\right),$$

with boundary conditions $\alpha(T,T)=0,\ \beta(T,T)=u_\lambda$ and $c_\omega^{(i)}(T,T,b_\omega^{(n)})=0\,,\ \forall i \in 1,\dots d,\ \forall \omega \in 1,\dots n$. We hence find

$$0 = \frac{\partial \alpha(t,T)}{\partial t} + \sum_{i=1}^{d}\Bigg\{ \frac{\partial \beta^{(i)}(t,T)}{\partial t}\lambda^{(i,n)} + \sum_{\omega=1}^{n}\frac{\partial c_\omega^{(i)}(t,T,b_\omega^{(n)})}{\partial t}\, Y^{(i,\omega)}$$
$$+ \beta^{(i)}(t,T)\sum_{\omega=1}^{n} m_\omega^{(i,n)}\left(\alpha^{(i)}\left(\lambda_\infty^{(i)} - \lambda^{(i,n)}\right) - b_\omega^{(n)} Y^{(i,\omega)}\right) + \sum_{\omega=1}^{n} c_\omega^{(i)}(t,T,b_\omega^{(n)})\left(\alpha^{(i)}\left(\lambda_\infty^{(i)} - \lambda^{(i,n)}\right) - b_\omega^{(n)} Y^{(i,\omega)}\right)$$
$$+ \lambda^{(i,n)}\left[\left(\left(\int \exp\left(\sum_{j=1}^{d}\beta^{(j)}(t,T)\,\eta^{(j,i)} o \sum_{\omega=1}^{n} m_\omega^{(j,n)} + u_{L,i}\, o + \sum_{j=1}^{d}\sum_{\omega=1}^{n} c_\omega^{(j)}(t,T,b_\omega^{(n)})\,\eta^{(j,i)} o\right)\nu^{(i)}(do)\right) e^{u_{N,i}} - 1\right]$$
$$+ \frac{1}{2}\lambda^{(i,n)}\sigma^{(i)^2}\left(\sum_{\omega=1}^{n}\left(m_\omega^{(i,n)}\beta^{(i)}(t,T) + c_\omega^{(i)}(t,T,b_\omega^{(n)})\right)\right)^2 \Bigg\}.$$

From the previous equation, we can obtain the following system of equations

$$\frac{\partial \alpha(t,T)}{\partial t} = -\sum_{i=1}^{d}\left(\alpha^{(i)}\lambda_\infty^{(i)}\sum_{\omega=1}^{n}\left(\beta^{(i)}(t,T)\, m_\omega^{(i,n)} + c_\omega^{(i)}(t,T,b_\omega^{(n)})\right)\right),$$
$$\frac{\partial c_\omega^{(i)}\left(t,T,b_\omega^{(n)}\right)}{\partial t} = \beta^{(i)}(t,T)\, m_\omega^{(i,n)}\, b_\omega^{(n)} + c_\omega^{(i)}(t,T,b_\omega^{(n)})\, b_\omega^{(n)}\,, \quad \forall i=1,\dots,d\,,\ \forall \omega=1,\dots,n,$$
$$\frac{\partial \beta^{(i)}(t,T)}{\partial t} = \alpha^{(i)}\sum_{\omega=1}^{n}\left(m_\omega^{(i,n)}\beta^{(i)}(t,T) + c_\omega^{(i)}(t,T,b_\omega^{(n)})\right) - \frac{1}{2}\sigma^{(i)^2}\left(\sum_{\omega=1}^{n}\left(m_\omega^{(i,n)}\beta^{(i)}(t,T) + c_\omega^{(i)}(t,T,b_\omega^{(n)})\right)\right)^2$$
$$- \theta^{(i)}\left(\sum_{j=1}^{d}\eta^{(j,i)}\sum_{\omega=1}^{n}\left(\beta^{(j)}(t,T)\, m_\omega^{(j,n)} + c_\omega^{(j)}(t,T,b_\omega^{(n)})\right) + u_{L,i}\right) e^{u_{N,i}} + 1\,, \quad \forall i=1,\dots,d.$$

This finally leads for the coefficients $\beta^{(i)}(.)$ and $c_\omega^{(i)}(.)$ to

$$c_\omega^{(i)}(t,T,b_\omega^{(n)}) = -b_\omega^{(n)}\, m_\omega^{(i,n)}\int_t^T e^{-b_\omega^{(n)}(v-t)}\,\beta^{(i)}(v,T)\,dv\,, \quad \forall i=1,\dots,d,\ \forall \omega=1,\dots,n,$$

and

$$\alpha(t,T) = \sum_{i=1}^{d}\left(\alpha^{(i)}\lambda_\infty^{(i)}\sum_{\omega=1}^{n} m_\omega^{(i,n)}\left(\int_t^T \beta^{(i)}(v,T)\,dv - b_\omega^{(n)}\int_t^T\int_v^T e^{-b_\omega^{(n)}(s-v)}\,\beta^{(i)}(s,T)\,ds\,dv\right)\right)$$
$$= \sum_{i=1}^{d}\left(\alpha^{(i)}\lambda_\infty^{(i)}\sum_{\omega=1}^{n} m_\omega^{(i,n)}\left(\int_t^T e^{-b_\omega^{(n)}(v-t)}\beta^{(i)}(v,T)\,dv\right)\right).$$

From Corollary 3.4, passing to the limit $n \to \infty$ leads to the announced result. In particular, we have almost surely that

$$\lim_{n\to\infty}\sum_{\omega=1}^{n}\left(m_\omega^{(i,n)}\beta^{(i)}(t,T) + c_\omega^{(i)}(t,T,b_\omega^{(n)})\right) = -\frac{\partial}{\partial_t}\int_t^T f_{\gamma^{(i)}}(v-t)\,\beta^{(i)}(v,T)\,dv\,, \quad \forall i \in 1,\dots,d.$$

□

Appendix B.5. Proof of Proposition 4.1.

We start by studying the discretized version of the above processes (indexed by n). The conditional transform of $\lambda_t^{Q,(i,n)}$ under Q induced by the affine martingale $M_t^{(n)}$ is

$$\mathbb{E}^Q\left(\exp\left(\sum_{i=1}^d \xi_i\,\lambda_T^{Q,(i,n)}\right)\Bigg|\,\mathcal{F}_t\right) = e^{-U_t^{(n)}}\mathbb{E}\left(\exp\left(U_T^{(n)}+\sum_{i=1}^d \xi_i\,\theta^{(i)}\left(\varepsilon^{(i)}\right)\lambda_T^{(i,n)}\right)\Bigg|\,\mathcal{F}_t\right),$$

where $M_t^{(n)} = e^{U_t^{(n)}}$, $dU_t^{(n)} = \sum_{i=1}^d\left(h^{(i)}(\varepsilon^{(i)})\lambda_t^{(i,n)}\,dt + \varepsilon^{(i)}dL_t^{(i,n)}\right) - \frac{1}{2}\psi(t)^2\,dt - \psi(t)\,dW_t$ and $\xi_i\in\mathbb{C}$. We then write with $\lambda_t^{(n)} := \left\{\lambda_t^{(i,n)}\right\}_{i=1,\ldots,d}$ and $Y_t^{(\omega)} := \left\{Y_t^{(i,\omega)}\right\}_{i=1,\ldots,d}$:

$$f\left(t,T,U_t^{(n)},\lambda_t^{(n)},\left\{Y_t^{(\omega)}\right\}_{\omega=1,\ldots,n}\right) := \mathbb{E}\left(\exp\left(U_T^{(n)}+\sum_{i=1}^d \xi_i\,\theta^{(i)}\left(\varepsilon^{(i)}\right)\lambda_T^{(i,n)}\right)\Bigg|\,\mathcal{F}_t\right),$$

and as in Proposition 3.6, we assume an affine structure for f :

$$\begin{aligned} &f\left(t,T,U_t^{(n)},\lambda_t^{(n)},\left\{Y_t^{(\omega)}\right\}_{\omega=1,\ldots,n}\right)\\ &\quad=\exp\left(A(t,T)+\sum_{i=1}^d\theta^{(i)}(\varepsilon^{(i)})\,B^{(i)}(t,T)\,\lambda_t^{(i,n)}+\sum_{i=1}^d\sum_{\omega=1}^n c_\omega^{(i)}(t,T,b_\omega^{(n)})\,Y_t^{(i,\omega)}+D(t,T)\,U_t^{(n)}\right),\end{aligned}$$

where $A(T,T)=0$, $B^{(i)}(T,T)=\xi_i$, $c_\omega^{(i)}(T,T,b_\omega^{(n)})=0$ and $D(T,T)=1$. The infinitesimal generator of $\left(\lambda^{(n)},U^{(n)}\right)$ for a sufficiently regular function f is given by

$$\begin{aligned}\mathcal{L}f(x) = \sum_{i=1}^d&\left\{\frac{\partial f}{\partial\lambda^{(i,n)}}(x)\sum_{\omega=1}^n m_\omega^{(i,n)}\left(\alpha^{(i)}\left(\lambda_\infty^{(i)}-\lambda^{(i,n)}\right)-b_\omega^{(n)}Y^{(i,\omega)}\right)+\sum_{\omega=1}^n\frac{\partial f}{\partial Y^{(i,\omega)}}(x)\Big[\alpha^{(i)}\left(\lambda_\infty^{(i)}-\lambda^{(i,n)}\right)\right.\\ &-b_\omega^{(n)}Y^{(i,\omega)}\Big]+\lambda^{(i,n)}\int_0^\infty\Bigg[f\left(\left\{\lambda^{(j,n)}+\eta^{(j,i)}\,o\sum_{\omega=1}^n m_\omega^{(j,n)}\right\}_{j=1,\ldots,d},\left\{Y^{(j,\omega)}+\eta^{(j,i)}o\right\}_{\substack{j=1,\ldots,d\\ \omega=1,\ldots,n}},U^{(n)}+\varepsilon^{(i)}o\right)\\ &\left.-f(x)\Bigg]\,\nu^{(i)}(do)+\frac{\partial f}{\partial U^{(n)}}(x)\,h^{(i)}\left(\varepsilon^{(i)}\right)\lambda^{(i,n)}\right\}-\frac{\partial f}{\partial U^{(n)}}(x)\frac{1}{2}\psi(t)^2+\frac{1}{2}\frac{\partial^2 f}{\partial U^{(n)2}}(x)\,\psi(t)^2\,,\end{aligned}$$

with $x=\left(\lambda_t^{(n)},U_t^{(n)}\right)$. Using the PIDE (2.8) and computing the corresponding partial derivatives, we find

$$\begin{aligned}0=\;&\frac{\partial A(t,T)}{\partial t}+\frac{\partial D(t,T)}{\partial t}U^{(n)}+\sum_{i=1}^d\left\{\theta^{(i)}(\varepsilon^{(i)})\frac{\partial B^{(i)}(t,T)}{\partial t}\lambda^{(i,n)}+\sum_{\omega=1}^n\frac{\partial c_\omega^{(i)}(t,T,b_\omega^{(n)})}{\partial t}\,Y^{(i,\omega)}\right.\\ &+\theta^{(i)}(\varepsilon^{(i)})B^{(i)}(t,T)\sum_{\omega=1}^n m_\omega^{(i,n)}\left(\alpha^{(i)}\left(\lambda_\infty^{(i)}-\lambda^{(i,n)}\right)-b_\omega^{(n)}Y^{(i,\omega)}\right)+\sum_{\omega=1}^n c_\omega^{(i)}(t,T,b_\omega^{(n)})\left(\alpha^{(i)}\left(\lambda_\infty^{(i)}-\lambda^{(i,n)}\right)-b_\omega^{(n)}Y^{(i,\omega)}\right)\\ &+\lambda^{(i,n)}\left[\theta^{(i)}\left(\sum_{j=1}^d\theta^{(j)}(\varepsilon^{(j)})B^{(j)}(t,T)\,\eta^{(j,i)}\sum_{\omega=1}^n m_\omega^{(j,n)}+\sum_{j=1}^d\sum_{\omega=1}^n c_\omega^{(j)}(t,T,b_\omega^{(n)})\,\eta^{(j,i)}+D(t,T)\,\varepsilon^{(i)}\right)-1\right]\\ &\left.+D(t,T)\,h^{(i)}(\varepsilon^{(i)})\,\lambda^{(i,n)}\right\}+\frac{1}{2}\psi(t)^2\left(-D(t,T)+D(t,T)^2\right).\end{aligned}$$

From the previous equation, we can obtain the following system

$$\frac{\partial A(t,T)}{\partial t} = -\sum_{i=1}^{d}\left(\alpha^{(i)}\lambda_\infty^{(i)} \sum_{\omega=1}^{n}\left(\theta^{(i)}(\varepsilon^{(i)})B^{(i)}(t,T)\, m_\omega^{(i,n)} + c_\omega^{(i)}(t,T,b_\omega^{(n)})\right)\right) + \frac{1}{2}\psi(t)^2\left(D(t,T) - D(t,T)^2\right),$$

$$\frac{\partial c_\omega^{(i)}\left(t,T,b_\omega^{(n)}\right)}{\partial t} = \theta^{(i)}(\varepsilon^{(i)})B^{(i)}(t,T)\, m_\omega^{(i,n)}\, b_\omega^{(n)} + c_\omega^{(i)}(t,T,b_\omega^{(n)})\, b_\omega^{(n)}, \quad \forall i = 1,\dots,d,\ \forall \omega = 1,\dots,n.$$

$$\theta^{(i)}(\varepsilon^{(i)})\frac{\partial B^{(i)}(t,T)}{\partial t} = \alpha^{(i)}\sum_{\omega=1}^{n}\left(m_\omega^{(i,n)}\theta^{(i)}(\varepsilon^{(i)})B^{(i)}(t,T) + c_\omega^{(i)}(t,T,b_\omega^{(n)})\right) - D(t,T)\, h^{(i)}(\varepsilon^{(i)})$$
$$- \theta^{(i)}\left(\sum_{j=1}^{d}\eta^{(j,i)}\sum_{\omega=1}^{n}\left(\theta^{(j)}(\varepsilon^{(j)})\, B^{(j)}(t,T)\, m_\omega^{(j,n)} + c_\omega^{(j)}(t,T,b_\omega^{(n)})\right) + D(t,T)\,\varepsilon^{(i)}\right) + 1, \quad \forall i = 1,\dots,d.$$

$$\frac{\partial D(t,T)}{\partial t} = 0. \tag{B.5}$$

Since $D(T,T) = 1$, $D(t,T) = 1$ solves the last PDE. For the coefficients $c_\omega^{(i)}(.)$, we have

$$c_\omega^{(i)}(t,T,b_\omega^{(n)}) = -b_\omega^{(n)}\, m_\omega^{(i,n)}\theta^{(i)}(\varepsilon^{(i)})\int_t^T e^{-b_\omega^{(n)}(v-t)}\, B^{(i)}(v,T)\, dv, \quad \forall i = 1,\dots,d,\ \forall \omega = 1,\dots n.$$

and for the coefficient $A(t,T)$:

$$A(t,T) = \sum_{i=1}^{d}\left(\alpha^{(i)}\lambda_\infty^{(i)}\,\theta^{(i)}(\varepsilon^{(i)})\sum_{\omega=1}^{n} m_\omega^{(i,n)}\left(\int_t^T B^{(i)}(v,T)\, dv - b_\omega^{(n)}\int_t^T\int_v^T e^{-b_\omega^{(n)}(s-v)}\, B^{(i)}(s,T)\, ds\, dv\right)\right)$$
$$= \sum_{i=1}^{d}\left(\alpha^{(i)}\lambda_\infty^{Q,(i)}\sum_{\omega=1}^{n} m_\omega^{(i,n)}\left(\int_t^T e^{-b_\omega^{(n)}(v-t)}B^{(i)}(v,T)\, dv\right)\right).$$

The PDE satisfied by the coefficients $B^{(i)}(t,T)$, $i = 1,...,d$ can be rewritten thanks to the martingale condition $h^{(i)}(\varepsilon^{(i)}) = 1 - \theta^{(i)}(\varepsilon)$:

$$\theta^{(i)}(\varepsilon^{(i)})\frac{\partial B^{(i)}(t,T)}{\partial t} = \theta^{(i)}(\varepsilon^{(i)})\,\alpha^{(i)}\sum_{\omega=1}^{n}\left(m_\omega^{(i,n)}B^{(i)}(t,T) - b_\omega^{(n)}\, m_\omega^{(i,n)}\int_t^T e^{-b_\omega^{(n)}(v-t)}\, B^{(i)}(v,T)\, dv\right)$$
$$- \theta^{(i)}\left[\sum_{j=1}^{d}\eta^{(j,i)}\theta^{(j)}(\varepsilon^{(j)})\sum_{\omega=1}^{n}\left(B^{(j)}(t,T)\, m_\omega^{(j,n)} - b_\omega^{(n)}\, m_\omega^{(j,n)}\int_t^T e^{-b_\omega^{(n)}(v-t)}\, B^{(j)}(v,T)\, dv\right) + \varepsilon^{(i)}\right] + \theta^{(i)}(\varepsilon^{(i)}).$$

Using equation (4.22), we have with $\eta^{Q,(j,i)} := \theta^{(j)}(\varepsilon^{(j)})\,\eta^{(j,i)}$,

$$\frac{\partial B^{(i)}(t,T)}{\partial t} = \alpha^{(i)}\sum_{\omega=1}^{n}\left(m_\omega^{(i,n)}B^{(i)}(t,T) - b_\omega^{(n)}\, m_\omega^{(i,n)}\int_t^T e^{-b_\omega^{(n)}(v-t)}\, B^{(i)}(v,T)\, dv\right)$$
$$- \theta^{Q,(i)}\left[\sum_{j=1}^{d}\eta^{Q,(j,i)}\sum_{\omega=1}^{n}\left(B^{(j)}(t,T)\, m_\omega^{(j,n)} - b_\omega^{(n)}\, m_\omega^{(j,n)}\int_t^T e^{-b_\omega^{(n)}(v-t)}\, B^{(j)}(v,T)\, dv\right)\right] + 1.$$

Passing to the limit $n \to \infty$, we have almost surely that

$$\mathbb{E}^Q\left(\exp\left(\sum_{i=1}^{d}\xi_i\,\lambda_T^{Q,(i)}\right)\middle|\,\mathcal{F}_t\right) = \exp\left[\sum_{i=1}^{d}\left(\alpha^{(i)}\,\lambda_\infty^{Q,(i)}\int_t^T f_{\gamma^{(i)}}(v-t)B^{(i)}(v,T)dv + B^{(i)}(t,T)\lambda_t^{Q,(i)}\right.\right.$$
$$\left.\left. - \int_0^\infty\left(\int_t^T e^{-\omega(v-t)}B^{(i)}(v,T)dv\right)\omega\, Y_t^{Q(i,\omega)}\mu^{(i)}d\omega\right)\right], \tag{B.6}$$

where $\lambda_t^{Q,(i)} := \theta^{(i)}(\varepsilon^{(i)})\,\lambda_t^{(i)}$, $Y_t^{Q,(i,\omega)} := \theta^{(i)}(\varepsilon^{(i)})\,Y_t^{(i,\omega)}$ and where the coefficients $B^{(i)}(t,T)$ satisfy for $i = 1,\dots,d$:

$$\begin{aligned}\frac{\partial B^{(i)}(t,T)}{\partial t} =& -\alpha^{(i)}\frac{\partial}{\partial t}\int_t^T f_{\gamma^{(i)}}(v-t)\,B^{(i)}(v,T)\,dv \\ & -\theta^{Q,(i)}\left(\sum_{j=1}^d \eta^{Q,(j,i)}\frac{\partial}{\partial_t}\int_t^T f_{\gamma^{(j)}}(v-t)\,B^{(j)}(v,T)\,dv\right)+1\,. \end{aligned} \tag{B.7}$$

This conditional transform has the exact same form as the one of Proposition 3.6 but with parameters $\eta^{Q,(j,i)}$, $\lambda_\infty^{Q,(i)}$ and transform $\theta^{Q,(i)}(.)$. Hence, we have that the processes $Y_t^{Q,(i,\omega)}$ satisfy

$$Y_t^{Q,(i,\omega)} = \int_0^t e^{-\omega(t-u)}\left(\alpha(\lambda_\infty^{Q,(i)}-\lambda_u^{Q,(i)})\,du + \sum_{j=1}^d \eta^{Q,(i,j)}\,dL_t^{Q,(j)}\right), \forall i = 1,\dots,d,\ \forall\omega = 1,\dots,n,$$

which is indeed equal to $\theta^{(i)}(\varepsilon^{(i)})Y_t^{(i,\omega)}$ and which satisfies as stated in Proposition 4.1,

$$dY_t^{Q,(i,\omega)} = \left(-\omega\,Y_t^{Q,(i,\omega)} + \alpha\left(\lambda_\infty^{Q,(i)}-\lambda_t^{Q,(i)}\right)\right)dt + \sum_{j=1}^d \eta^{Q,(i,j)}\,dL_t^{Q,(j)}\,,\quad Y_0^{Q(i,\omega)} = 0\,. \tag{B.8}$$

The dynamic of the conditional intensities $\lambda_t^{Q,(i)}$ is therefore well preserved under the equivalent measure Q. □

Appendix B.6. Proof of Proposition 4.2.

We first show that under the measure Q defined by (4.20), the process

$$W_t^{(a,Q)} = W_t^{(a)} + \int_0^t \psi(s)\,ds$$

is a Brownian motion. We indeed have

$$\begin{aligned}\mathbb{E}^Q\left[e^{\xi\,W_t^{(a,Q)}}\middle|\,\mathcal{F}_0\right] &= \mathbb{E}\left[M_t^{(n)}\exp\left(\xi\,W_t^{(a)} + \int_0^t \xi\,\psi(s)\,ds\right)\middle|\,\mathcal{F}_0\right],\\ &= \mathbb{E}\left[\exp\left(\sum_{i=1}^d h^{(i)}(\varepsilon^{(i)})\int_0^t \lambda_s^{(i,n)}ds + \varepsilon^{(i)}L_t^{(i,n)}\right)\right]\\ &\quad\times e^{\xi^2 t/2}\,\mathbb{E}\left[\exp\left(-\frac{1}{2}\int_0^t(\psi(s)-\xi)^2ds - \int_0^t(\psi(s)-\xi)\,ds\right)\right].\end{aligned}$$

Since $h^{(i)}(\varepsilon^{(i)}) = 1-\theta^{(i)}(\varepsilon^{(i)})$, we know from above that the first expectation is a martingale. The second expectation is a Doleans-Dade exponential, which is a martingale. We hence obtain

$$\mathbb{E}^Q\left[e^{\xi\,W_t^{(a,Q)}}\middle|\,\mathcal{F}_0\right] = e^{\xi^2\,t/2}\,,$$

and we recognize the characteristic function of a Brownian motion. Finally, if $\psi(t) = \lambda$ (constant price of illiquidity risk), we directly have with $\widetilde{\theta}_2 = \theta_2 - \frac{\lambda\,\theta_3}{\theta_1}$:

$$\begin{aligned}da_t &= \theta_1\,(\theta_2 - a_t)\,dt + \theta_3\left(dW_t^{(a,Q)} - \lambda\,dt\right) + dL_t^{Q,(1)} - dL_t^{Q,(2)}\\ &= \theta_1\left(\widetilde{\theta}_2 - a_t\right)dt + \theta_3\,dW_t^{(a,Q)} + dL_t^{Q,(1)} - dL_t^{Q,(2)}\,.\end{aligned}$$

□

Data availability statement The datasets and code generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Disclosure statement The authors declare that they have no conflict of interest.

Funding details This work was supported by the FNRS, Belgium under Grant no. 33658713.

References

Eduardo Abi Jaber. Lifting the heston model. *Quantitative Finance*, 19(12):1995–2013, 2019.

Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine volterra processes. *The Annals of Applied Probability*, 29(5):3155–3200, 2019.

Yacine Aït-Sahalia, Julio Cacho-Diaz, and Roger JA Laeven. Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117(3):585–606, 2015.

Claudio Albanese, Ken Jackson, and Petter Wiberg. A new fourier transform algorithm for value-at-risk. *Quantitative Finance*, 4(3):328, 2004.

Yakov Amihud. Illiquidity and stock returns: cross-section and time-series effects. *Journal of financial markets*, 5(1):31–56, 2002.

Yakov Amihud and Haim Mendelson. The pricing of illiquidity as a characteristic and as risk. *Multinational Finance Journal*, 19(3):149–168, 2015.

Torben G Andersen, Tim Bollerslev, Francis X Diebold, and Heiko Ebens. The distribution of realized stock return volatility. *Journal of financial economics*, 61(1):43–76, 2001.

Thierry Ané and Hélyette Geman. Order flow, transaction clock, and normality of asset returns. *The Journal of Finance*, 55(5):2259–2284, 2000.

Emmanuel Bacry and Jean-François Muzy. Hawkes model for price and trades high-frequency dynamics. *Quantitative Finance*, 14(7):1147–1166, 2014.

Emmanuel Bacry, Khalil Dayri, and Jean-François Muzy. Non-parametric kernel estimation for symmetric hawkes processes. application to high frequency financial data. *The European Physical Journal B*, 85(5):1–12, 2012.

Emmanuel Bacry, Iacopo Mastromatteo, and Jean-François Muzy. Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(01):1550005, 2015.

Harry Bateman. *Higher transcendental functions [volumes i-iii]*, volume 1. McGraw-Hill Book Company, 1953.

Nicole Bäuerle and Sascha Desmettre. Portfolio optimization in fractional and rough heston models. *SIAM Journal on Financial Mathematics*, 11(1):240–273, 2020.

Nicole Branger, Holger Kraft, and Christoph Meinerding. Partial information about contagion risk, self-exciting processes and portfolio optimization. *Journal of Economic Dynamics and Control*, 39:18–36, 2014.

Michael Brennan, Sahn-Wook Huh, and Avanidhar Subrahmanyam. An analysis of the amihud illiquidity premium. *The Review of Asset Pricing Studies*, 3(1):133–176, 2013.

Charles Cao, Yong Chen, Bing Liang, and Andrew W Lo. Can hedge funds time market liquidity? *Journal of Financial Economics*, 109(2):493–516, 2013.

Peter Carr and Dilip Madan. Option valuation using the fast fourier transform. *Journal of computational finance*, 2(4):61–73, 1999.

Jing Chen, AG Hawkes, and E Scalas. A fractional hawkes process. In *Nonlocal and Fractional Operators*, pages 121–131. Springer, 2021.

José Da Fonseca and Riadh Zaatour. Hawkes process: Fast calibration, application to trade clustering, and diffusive limit. *Journal of Futures Markets*, 34(6):548–579, 2014.

Jean-Loup Dupret and Donatien Hainaut. Portfolio insurance under rough volatility and volterra processes. *International Journal of Theoretical and Applied Finance*, 24(06n07):2150036, 2021.

Omar El Euch and Mathieu Rosenbaum. The characteristic function of rough heston models. *Mathematical Finance*, 29(1):3–38, 2019.

Paul Embrechts, Thomas Liniger, and Lu Lin. Multivariate hawkes processes: an application to financial data. *Journal of Applied Probability*, 48(A):367–378, 2011.

Eymen Errais, Kay Giesecke, and Lisa R Goldberg. Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, 1(1):642–665, 2010.

A Ronald Gallant, Peter E Rossi, and George Tauchen. Stock prices and volume. *The Review of Financial Studies*, 5(2):199–242, 1992.

Jim Gatheral, Thibault Jaisson, and Mathieu Rosenbaum. Volatility is rough. *Quantitative finance*, 18(6):933–949, 2018.

Ruslan Y Goyenko, Craig W Holden, and Charles A Trzcinka. Do liquidity measures measure liquidity? *Journal of financial Economics*, 92(2):153–181, 2009.

B Gross. L. note on the inversion of the laplace transform. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(317):543–544, 1950.

Christian Hafner, Oliver B Linton, and Linqi Wang. Dynamic autoregressive liquidity (darliq). *Available at SSRN 4042069*, 2022.

Donatien Hainaut. A bivariate hawkes process for interest rate modeling. *Economic Modelling*, 57:180–196, 2016a.

Donatien Hainaut. A model for interest rates with clustering effects. *Quantitative Finance*, 16(8): 1203–1218, 2016b.

Donatien Hainaut. Fractional hawkes processes. *Physica A: Statistical Mechanics and its Applications*, 549:124330, 2020.

Donatien Hainaut. *Continuous time processes for finance. Switching, self-exciting, fractional and other recent dynamics.* Bocconi & Springer Series, 2022.

Donatien Hainaut and Franck Moraux. Hedging of options in the presence of jump clustering. *Journal of Computational Finance, Forthcoming*, 2018.

Stephen J Hardiman, Nicolas Bercot, and Jean-Philippe Bouchaud. Critical reflexivity in financial markets: a hawkes process analysis. *The European Physical Journal B*, 86(10):1–9, 2013.

Alan G Hawkes. Hawkes jump-diffusions and finance: a brief history and review. *The European Journal of Finance*, 28(7):627–641, 2022.

Isidore Isaac Hirschman and David V Widder. *The convolution transform*. Courier Corporation, 2012.

Jonathan M Karpoff. The relation between price changes and trading volume: A survey. *Journal of Financial and quantitative Analysis*, 22(1):109–126, 1987.

Andreas E Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Springer Science & Business Media, 2006.

Kyungsub Lee and Byoung Ki Seo. Modeling microstructure price dynamics with symmetric hawkes and diffusion model using ultra-high-frequency stock data. *Journal of Economic Dynamics and Control*, 79:154–183, 2017.

Guanghua Lian, Carl Chiarella, and Petko S Kalev. Volatility swaps and volatility options on discretely sampled realized variance. *Journal of Economic Dynamics and Control*, 47:239–262, 2014.

Gwo Dong Lin. On the mittag–leffler distributions. *Journal of Statistical Planning and Inference*, 74(1):1–9, 1998.

Ignacio N Lobato and Carlos Velasco. Long memory in stock-market trading volume. *Journal of Business & Economic Statistics*, 18(4):410–427, 2000.

Xiaoxia Lou and Tao Shu. Price impact or trading volume: Why is the amihud (2002) measure priced? *The Review of Financial Studies*, 30(12):4481–4520, 2017.

Francesco Mainardi. Why the mittag-leffler function can be considered the queen function of the fractional calculus? *Entropy*, 22(12):1359, 2020.

Charles G Njike Leunga. Long memory self-exciting jump diffusion for asset prices modeling. Technical report, LIDAM Discussion Paper ISBA ; 2022/03, 2022.

Xiao-Wei Zhang, Peter W Glynn, Kay Giesecke, and Jose Blanchet. Rare event simulation for a generalized hawkes process. In *Proceedings of the 2009 Winter Simulation Conference (WSC)*, pages 1291–1298. IEEE, 2009.