A FRACTIONAL HAWKES PROCESS FOR ILLIQUIDITY MODELING

Jean-Loup Dupret, Donatien Hainaut

REPRINT | 2024 / 41

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 economic and financial literature for measuring the illiquidity process of stocks and indices. None of the existing discrete-time illiquidity models in the literature are however adapted for reproducing peaks of illiquidity with long memory and for the management of the liquidity risk associated with these securities. This paper therefore proposes a new continuous-time paradigm for modeling illiquidity via a novel fractional Hawkes process in which the intensity process is ruled by a modified Mittag-Leffler excitation function. By considering a mean-reverting jump-diffusion model for the (log-)Amihud measure where jumps follow this modified fractional Hawkes process, we then manage to effectively reproduce the observed peaks of illiquidity in financial markets while introducing long-term effects and tractability in the model. We can therefore use this model to directly perform risk management on the Amihud illiquidity measure. This paper hence provides new tools for a better management of the illiquidity risk in financial markets.

Keywords: Illiquidity modeling, Amihud measure, Hawkes process, Mittag-Leffler function, 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 in Bacry and Muzy [1], Bacry et al. [2], Hardiman et al. [3] and Cayé and Muhle-Karbe [4]; the firm defaults in a credit risk portfolio in Da Fonseca and Zaatour [5] and Errais et al. [6]; or the jumps of asset prices in financial markets in Aït-Sahalia et al. [7], Aït-Sahalia and Hurd [8], Hawkes [9] and Hainaut and Goutte [10]. 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 enables a tractable and analytical treatment of Hawkes processes and hence provides explicit formulae for the moments, autocorrelation function and characteristic function of the number of jumps, see Aït-Sahalia et al. [7], Da Fonseca and Zaatour [5] or Errais et al. [6]. However, this exponential kernel cannot take into account long-term effects of past events on the number of jumps, which is typical in many financial applications such as in Bacry et al. [2] for

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

high-frequency trading, see also Hainaut [11] for a review. Therefore, another widely-used excitation kernel proposed in the literature is the power-law one, as it allows for a slower decay compared to the exponential kernel and can thus better reproduce these 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. [12], Bacry et al. [2], Hardiman et al. [3]. 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. [6], option pricing for self-exciting jump diffusion models as in Hawkes [9] or bond pricing under a self-exciting Hawkes dynamic for the short rate process as in Hainaut [13]. Hence, several authors have tried to fill this gap by introducing Hawkes models based on fractional dynamics. Hainaut [14] proposes a fractional Hawkes process defined by a subdiffusive dynamic for the intensity process while Njike Leunga [15] and Chen et al. [16] 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 perfectly suited for the modeling of financial and economic events with long-range dependence. We then introduce a new fractional Hawkes process based on a modified Mittag-Leffler kernel that yields a tractable characteristic function while exhibiting long-term effects. We study its properties and show that this new fractional Hawkes process is the best suited for financial applications with long memory, especially when the characteristic function of the number of jumps is required. Indeed, following Bäuerle and Desmettre [17], we show how to 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 some practical applications of this modified fractional Hawkes process in the context of illiquidity modeling and show that it outperforms existing self-exciting models in terms of goodness of fit with observed illiquidity data in the market. Several econometric indicators have indeed been proposed for measuring the illiquidity process of stocks and indices using daily (or weekly) frequency data, see Goyenko et al. [18] for a review. We here focus on the Amihud illiquidity measure as introduced in Amihud [19], which has proved 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 [20]. This premium arises due to investors taking into consideration illiquidity costs and pricing them into their expected returns so as to compensate for this market impact. Numerous econometric models in discrete time have then been proposed for modeling this Amihud measure, as in Amihud [20], Brennan et al. [21] or more recently Hafner et al. [22]. 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 where the jumps follow a fractional Hawkes process ruled by the modified Mittag-Leffler excitation function introduced in the first part of the paper. This continuous-time approach helps us to obtain the characteristic function satisfied by the Amihud measure in quasi-closed form, leading to a time-efficient and accurate numerical computation of the moments, density and risk measures of the Amihud illiquidity measure. Therefore, this model enables time-efficient and exact risk management on illiquidity as well as to 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, incorporating fractional Hawkes jumps in the model allows us to reproduce the observed peaks of illiquidity in financial markets and to show that these peaks effectively exhibit long-term effects. 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.

The goal of this paper is therefore fourfold. 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 economic and financial events with long-range dependence, compared with the existing Hawkes models. We then propose a new paradigm based on a continuous-time process for studying the illiquidity of financial assets via the Amihud illiquidity measure. More precisely, working with a mean-reverting jump model for the log-Amihud measure with jumps following this modified fractional Hawkes process allows to easily reproduce the observed peaks of illiquidity in financial markets while introducing long-term effects and tractability in the model. In a third step, we show via a Peaks Over Threshold procedure with likelihood maximization how to estimate the parameters of our mean-reverting fractional Hawkes model. By deriving the characteristic function satisfied by the Amihud measure in this model, we finally 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. We hence provide with this paper new tools for a better understanding, management and assessment of the liquidity risk in financial markets.

2. Hawkes processes

The *Hawkes process* is a self-exciting point process introduced in Hawkes [23]. More precisely, it 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.

We fix a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with right-continuous and complete filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$, on which are defined a nonexplosive counting process N given by $N_t := \sum_{j\geq 1} \mathbb{1}_{\{\tau_j \leq t\}}$ and a point process L given by

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

where the jump sizes $O_j \in \mathcal{F}_{\tau_j}$ are i.i.d. random variables and where $(\tau_j)_{j\geq 1}$ is a sequence of positive stopping times corresponding to the arrivals of jumps. These jumps represent market events such as peaks of illiquidity, defaults in a portfolio of firms or the occurrence of a trade. In a Hawkes setting, the processes N and L are directly specified through a conditional arrival rate or intensity λ as well as a distribution ν on $(0,\infty)$ for the jump size O (assumed to be strictly positive in this paper). We suppose that the jump transform $\int_0^\infty e^{\xi o} \, \mathrm{d}\nu(o)$ exists, is finite for complex ξ and admits a finite derivative $\int_0^\infty o \, e^{\xi o} \, \mathrm{d}\nu(o)$. The intensity then describes the conditional mean arrival rate of the Hawkes process and is assumed to follow a strictly positive stochastic process such that

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

where $\eta > 0$ and where $\lambda_0(t)$ is the positive 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_t 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[1]. 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. Note that in the self-excitation setting considered in this paper, jumps of the Hawkes process always increase its intensity. From Bacry et al. [2], we then have that the two-dimensional process $J = (L, N)^\top$ is a Hawkes process provided that the excitation kernel ϕ satisfies the following definition.

Definition 2.1. I. **Positive** : $\phi(t) \geq 0 \;\forall t$.

II. **Causal** : $\phi(t) = 0$ for $t < 0$.

III. $\mathbf{L^1}$**-integrable** : $\int_0^\infty \phi(s) \, \mathrm{d}s < \infty$.

From equation (2.2), we clearly have that the higher the jump size O, the higher the effect of the jump 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 is shown in Errais et al. [6] that $N - \int_0^\cdot \lambda_s \, \mathrm{d}s$ is a local martingale relative to $\mathbb{F}$ and $\mathbb{P}$.

[1] This notation in terms of η and f will allow to emphasize the separate role from the clustering parameter η and the kernel f on the Hawkes process' properties, see Proposition 3.1.

As explained in Bacry et al. [2], the L^1 norm of ϕ (denoted $||\phi||_1 := \int_0^\infty \phi(s)\,\mathrm{d}s = \eta\,||f||_1$) 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 a Hawkes process ruled by (2.2) is well-defined for any choice of kernel ϕ satisfying the three conditions of Definition 2.1, 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 $\mathbb{E}(O)||\phi||_1 < 1$. In the multivariate extension of the Hawkes process (see Appendix B.7), the L^1 norm required for this stationary condition is replaced by the spectral radius $\varrho(\cdot)$, as explained in Bacry et al. [2].

This setting is largely used in the financial literature 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), see the review of Bacry et al. [2]. 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, as in Errais et al. [6] or Da Fonseca and Zaatour [5]. Finally, jumps in asset prices can also be studied using these Hawkes processes, see Aït-Sahalia et al. [7], Hawkes [9], Njike Leunga and Hainaut [24]. We now introduce a new fractional Hawkes process based on a modified Mittag-Leffler kernel. We will study this kernel throughout the paper and compare it with standard choices of excitation kernel f proposed in the literature.

2.1. The modified Mittag-Leffler fractional Hawkes process

This new fractional Hawkes process is based on the following excitation kernel, defined as the difference of a Mittag-Leffler (ML) function and its derivative, namely

$$f_\gamma(t) = E_\gamma(t^\gamma) - \frac{\mathrm{d}E_\gamma(t^\gamma)}{\mathrm{d}t} = E_\gamma(t^\gamma) - t^{\gamma-1}E_{\gamma,\gamma}(t^\gamma)\,, \quad \gamma \in [1,2],\ t > 0, \tag{2.3}$$

where $E_\gamma(z) := \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(n\gamma+1)}$ is the one-parameter ML function and $E_{\gamma,\delta}(z) := \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(n\gamma+\delta)}$ the two-parameter ML function, both evaluated at $z = t^\gamma$. This kernel (2.3) with $\gamma \in [1,2]$ is called the modified Mittag-Leffler (mML) kernel. From the expansion 18.1(22) in Bateman [25], we can also easily derive the following asymptotics for our mML kernel (2.3) :

$$f_\gamma(t) \sim \begin{cases} 1 - \frac{t^{\gamma-1}}{\Gamma(\gamma)} \sim \exp\left[-\frac{t^{\gamma-1}}{\Gamma(\gamma)}\right], & t \to 0^+, \\ -\frac{t^{-\gamma}}{\Gamma(1-\gamma)}, & t \to \infty. \end{cases} \tag{2.4}$$

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. This kernel thus handles in an unified way the properties of the exponential and power-law kernels (see Section 2.2 and Fig. 1 below), providing both tractability and long memory to the Hawkes process. This will result in a more accurate modeling and fine tuning of the short- and long-term properties of the Hawkes process, as will be demonstrated in the following sections. The intensity of the mML fractional Hawkes process is then defined in the following way

$$\lambda_t = \lambda_0 + \int_0^t f_\gamma(t-u)\,\alpha\,(\lambda_\infty - \lambda_u)\,\mathrm{d}u + \int_0^t f_\gamma(t-u)\,\eta\,\mathrm{d}L_u\,, \tag{2.5}$$

where the background intensity function in (2.2) is chosen to be $\lambda_0(t) = \lambda_0 + \int_0^t f_\gamma(t-u)\,\alpha\,(\lambda_\infty - \lambda_u)\,\mathrm{d}u$, with $\lambda_0 > 0$ the intensity at initial time 0 and $\lambda_\infty > 0$ a fixed parameter controlling the long-run average intensity. Proposition 3.3 will show the existence and non-negativity of an intensity process satisfying (2.5) and Proposition 3.4 will provide some condition under which the mML Hawkes process effectively exhibits long-range dependence. The remainder of this section presents the most popular choices of decay function f discussed in the literature and compares them with the mML kernel (2.3). In particular, a numerical comparison is provided in Fig. 1, which highlights the benefits of the mML kernel for the financial and economic modeling of events with long-term effects. The exact mathematical and statistical properties of the mML kernel (2.3) and of the mML Hawkes process (2.5) are then thoroughly studied in Section 3. In Section 4 and 4.4, we apply this new fractional Hawkes process for modeling (il)liquidity and we show that the short- and long-term properties derived in Section 3 are perfectly suited for reproducing and studying the peaks of the (log-)Amihud illiquidity measure, compared with other excitation kernels.

2.2. Standard Hawkes processes

We first review the two most frequent choices of decay function in the literature :

- **Exponential kernel :**

$$f_\alpha(t) = e^{-\alpha t},\ t \geq 0\,,\ \alpha > 0\,, \tag{2.6}$$

where α is the rate at which the impact of an event decays over time and where the L^1 norm $||\phi||_1$ is equal to η/α. More precisely, a widely-used intensity process similar as our choice (2.5) with $\lambda_0, \lambda_\infty > 0$ is

$$\begin{aligned}\lambda_t &= \lambda_\infty + e^{-\alpha t}(\lambda_0 - \lambda_\infty) + \eta \int_0^t e^{-\alpha(t-u)}\,\mathrm{d}L_u &(2.7)\\ &= \lambda_0 + \int_0^t f_\alpha(t-u)\,\alpha\,(\lambda_\infty - \lambda_0)\,\mathrm{d}u + \eta\int_0^t f_\alpha(t-u)\,\mathrm{d}L_u\,, &(2.8)\end{aligned}$$

which can be rewritten by differentiating the previous equation as

$$\mathrm{d}\lambda_t = \alpha\,(\lambda_\infty - \lambda_t)\,\mathrm{d}t + \eta\,\mathrm{d}L_t\,. \tag{2.9}$$

As for the mML Hawkes intensity (2.5), the process λ does not blow up because the drift becomes negative whenever the intensity is above $\lambda_\infty > 0$ and prevents any explosion. The point process $J = (L, N)^\top$ 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 function $f : D \to \mathbb{R}$ with continuous partial derivative $\partial f / \partial \lambda$ such that

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

where $\mathbb{E}_t[\,\cdot\,] = \mathbb{E}[\,\cdot \mid \mathcal{F}_t]$ and $X_t = x = (\lambda_t, (L_t, N_t))^\top$. For the Hawkes process (2.7), we can write

$$\begin{aligned}\mathcal{L} f(x) &= \alpha(\lambda_\infty - \lambda)\frac{\partial f}{\partial \lambda}(x) + \lambda \mathbb{E}\left[f\left(\lambda + \eta O,\, L + O,\, N+1\right) - f(x)\right] \\ &= \alpha(\lambda_\infty - \lambda)\frac{\partial f}{\partial \lambda}(x) + \lambda \int_0^\infty \left(f\left(\lambda + \eta o,\, L + o,\, N+1\right) - f(x)\right)\nu(\mathrm{d}o)\,.\end{aligned}$$

Hence, Da Fonseca and Zaatour [5] obtain the following Dynkin's formula for this Hawkes process with $h > 0$

$$\mathbb{E}_t\left[f(X_{t+h})\right] = f(X_t) + \mathbb{E}_t\left[\int_t^{t+h} \mathcal{L} f(X_u)\,\mathrm{d}u\right]. \tag{2.10}$$

Furthermore, since the process $X = (\lambda, J)^\top$ is a 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 following PIDE

$$0 = \frac{\partial f}{\partial t}(t, T, x) + \mathcal{L} f(t, T, x)\,, \tag{2.11}$$

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. [6] that the point process $X = (\lambda, J)^\top = (\lambda, L, N)^\top$ is affine and that its conditional transform 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 Riccati ODEs

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

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

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

This exponential kernel is used *e.g.* in Errais et al. [6] in the context of portfolio credit risk, where the correlation of firm defaults is the main concern. Such specification of the Hawkes process allows the valuation of 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 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. [3] and in Bacry et al. [12] that the exponential kernel is not the best suited for reproducing the occurrence of trades at high-frequency, suggesting a long-memory nature of self-excitation phenomena at the microstructure level of price dynamics.

- **Power-law kernel :**

$$f_{\beta,\gamma}(t) = \frac{\beta}{(1+\beta t)^{1+\gamma}},\ \beta > 0,\ \gamma > 0,\ t \geq 0, \tag{2.12}$$

and $\lambda_t = \lambda_0 + \eta \int_0^t f_{\beta,\gamma}(t-u)\,\mathrm{d}L_u$. This kernel makes the pair (λ, J) non-Markov and we hence cannot rely on the above theory of infinitesimal generators to write the characteristic function of (λ, J), see Bacry et al. [2]. In the case of portfolio credit derivatives, this prevents from valuating efficiently these contracts as done in Errais et al. [6] with the exponential kernel (2.6) above. Similarly, this characteristic function is required for efficient risk management and estimation in the context of liquidity modeling, as we will see in Section 4.2-4.4. Other applications requiring this characteristic function include the pricing of options for self-exciting jump-diffusion models (see Aït-Sahalia et al. [7]) or the pricing of bonds under Hawkes dynamics (see Hainaut [13]). Nevertheless, the main benefit of this power-law kernel is to allow for a slower (power-law) decay than the exponential kernel so that it can be used for a better modeling of long-term self-exciting effects. More precisely, it is shown in Jaisson [26] that if $\gamma \in (0, 1/2)$, the related Hawkes process exhibits long-range dependence. This will be confirmed qualitatively in Fig. 1 below with a substantial contribution of past jumps in the intensity, even for large maturities. This long-term power-law decay is *e.g.* observed at high-frequency in Hardiman et al. [3] with $\gamma \approx 0$ (and thus a power-law exponent close to $1+\gamma \approx 1$), which hence enables a better modeling of the long-range nature of offer and demand. We therefore want to build a kernel such that the associated Hawkes process includes these long-term effects while having a characteristic function available in (semi-)closed form so that we can better manage the illiquidity risk in financial markets. Fractional Hawkes processes offer ways to achieve this, but first we show in the next section that none of the current fractional Hawkes in the literature fully satisfy the empirical properties observed in financial markets.

2.3. Fractional Hawkes processes

- **Mittag-Leffler kernel :**

$$f_\gamma(t) = E_\gamma(-t^\gamma)\,, \qquad t > 0,\ \gamma \in (0,1]\,, \tag{2.13}$$

where $E_\gamma(z)$ is again the one-parameter ML function but this time with $\gamma \in (0,1]$ and evaluated at $z = -t^\gamma$ (instead of $z = t^\gamma$ in (2.3)). For $t > 0$ and $\gamma \in (0,1]$, the one-parameter ML function $E_\gamma(-t^\gamma)$ has the meaning of survival function for some positive random variable T with infinite mean, called the ML random variable. The choice of this kernel is motivated in a more general setting in Njike Leunga [15] 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 with index $\gamma \in (0,1]$, see Mainardi [27],

$$E_\gamma(-t^\gamma) \sim \begin{cases} \exp\left[-\frac{t^\gamma}{\Gamma(1+\gamma)}\right], & t \to 0^+, \\ \\ \frac{t^{-\gamma}}{\Gamma(1-\gamma)} = \frac{\sin(\gamma\pi)}{\pi}\frac{\Gamma(\gamma)}{t^\gamma}, & t \to \infty. \end{cases} \tag{2.14}$$

As a consequence of the long-term power-law asymptotic, the process turns to be no longer Markov. 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 function for the positive ML random variable T, we find

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

since from Theorem 3 (ii) in Lin [28], the expected value of the ML random variable 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 kernel does not even define a Hawkes process since the third condition of Definition 2.1 is not satisfied.

- **Mittag-Leffler's derivative kernel (dML) :**

$$f_\gamma(t) = -\frac{\mathrm{d}E_\gamma(-t^\gamma)}{\mathrm{d}t} = t^{\gamma-1}E_{\gamma,\gamma}(-t^\gamma)\,, \quad t > 0,\ \gamma \in (0,1]. \tag{2.15}$$

where $E_{\gamma,\delta}(z)$ is the two-parameter ML function introduced in (2.3) but now evaluated at $z = -t^\gamma$. This kernel (2.15) is in fact the sign-changed first derivative of the ML kernel (2.13) and hence the pdf of the ML random variable T. It has been introduced in Chen et al. [16] as a non-Markov fractional Hawkes process with long-range dependence and with $||f||_1 = 1$, such that

$$t^{\gamma-1}E_{\gamma,\gamma}(-t^\gamma) \sim \begin{cases} +\infty & t \to 0^+, \\ \\ \frac{-t^{-\gamma-1}}{\Gamma(-\gamma)} = \frac{\sin(\gamma\pi)}{\pi}\frac{\Gamma(\gamma+1)}{t^{\gamma+1}}, & t \to \infty. \end{cases} \tag{2.16}$$

Although the three conditions of Definition 2.1 are satisfied (making this fractional Hawkes process (2.15) well defined), we directly see in equation (2.15) with $\gamma < 1$ and in (2.16) 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 3.2 will not apply to this dML kernel (due to the divergent quantity (B.8) with this kernel). As this characteristic function is required to perform risk management on illiquidity in Section 4.4, this prevents from using the dML kernel in our setting.

2.4. Comparison

We now compare numerically in Fig. 1 the different choices of kernels above. We first observe at short times

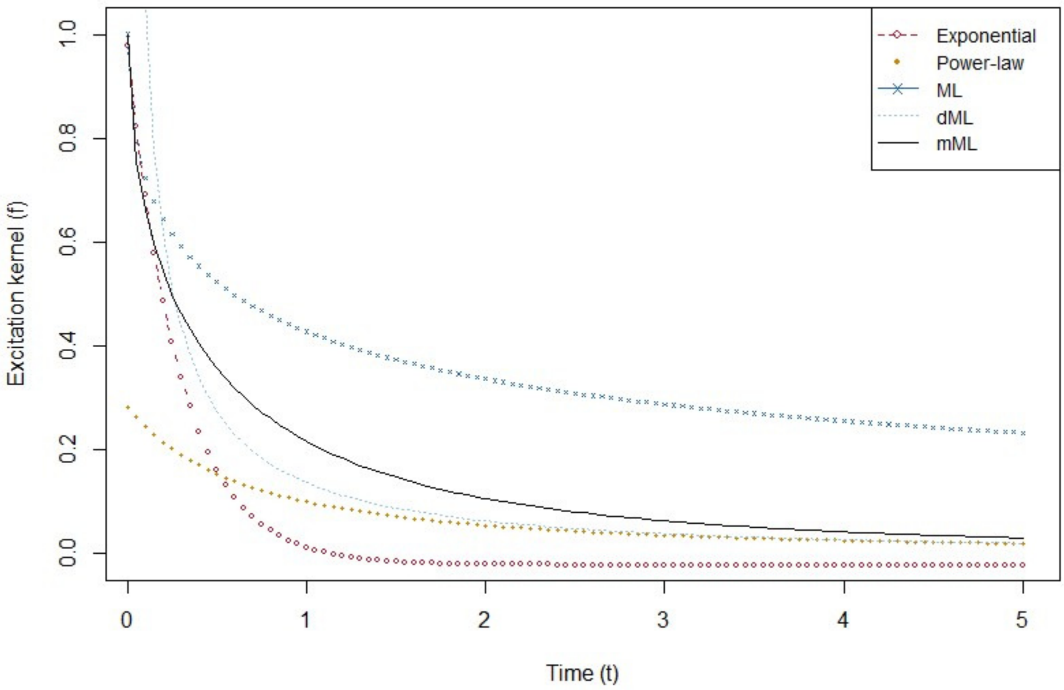

Fig. 1 Comparison of the excitation kernels proposed in the literature with the mML kernel (2.3). In red and orange dotted lines, we have the exponential kernel with $\alpha = 3.12$ and the power-law kernel with $\beta = 0.28, \gamma = 0.5$. In dark and light blue dotted lines, we have the ML and dML kernels with $\gamma = 0.5$. In black line, we have the mML kernel with $\gamma = 1.5$.

that the exponential, ML and mML kernels exhibit the same behavior. Indeed, both the ML kernel (2.13) and the mML kernel (2.3) behave as a stretched exponential for short times, and we in fact have the same asymptotics for $t \to 0^+$ in (2.4) and (2.14) given that $\gamma \in [1, 2]$ for the mML kernel whereas $\gamma \in (0, 1]$ in the ML kernel. We also clearly see the explosion of $f_\gamma(t)$ for $t \to 0^+$ in the dML kernel. At large times, we have instead that the power-law, dML and mML kernels exhibit the same decay. Indeed, the dML and mML kernels have the same long-term power-law decay since the asymptotics (2.4) and (2.16) for $t \to +\infty$ are equivalent (again from the condition $\gamma \in [1, 2]$ in the mML kernel and $\gamma \in (0, 1]$ in the dML kernel). We also see that the ML kernel in dark blue decays as a power-law function but much slower (at a rate γ instead of $1 + \gamma$ for $\gamma \in (0, 1]$), compared with the other fractional kernels. This explains that the L^1 norm explodes for this process making it not conceivable for financial modeling. Finally, the exponential kernel decreases very rapidly to zero, preventing from using this kernel for modeling long-term effects since there is no contribution anymore of past jumps in the intensity for maturities $t > 1$. Therefore, thanks to its power-law behavior at large times and since it keeps tractable formulae due to its fast decay at short times (see Section 3 and Proposition 3.4), our mML Hawkes process appears the best suited for modeling financial events with long-range dependence. This will be confirmed empirically with Fig. 6 in the context of the Amihud illiquidity measure. We now derive the exact mathematical and statistical properties of the mML kernel (2.3), required in Section 4 for the modeling of (il)liquidity in financial markets.

3. The modified ML fractional Hawkes process

3.1. Properties and spectral representation

We first recall that the mML kernel is defined in (2.3) as the difference between the ML kernel and its derivative, evaluated at t^γ with $\gamma \in [1, 2]$ and which admits the asymptotics (2.4) at short and large times. Moreover, we can directly expand the continuity of (2.3) 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. As depicted in Fig. 1, subtracting the derivative of the ML kernel in the mML kernel (2.3) allows a faster rate of decay at large times, which is necessary to properly define a fractional Hawkes process with finite average (as we now show in Proposition 3.1) and with long-range dependence (see Proposition 3.4).

Proposition 3.1. *(i) The L^1 norm of the modified Mittag-Leffler kernel* (2.3) *is given by*

$$||f_\gamma||_1 = \int_0^\infty f_\gamma(t)\,\mathrm{d}t = \int_0^\infty \left(E_\gamma(t^\gamma) - t^{\gamma-1}E_{\gamma,\gamma}(t^\gamma)\right)\mathrm{d}t = 1\,, \tag{3.1}$$

which thus leads to the L^1 norm $||\phi_\gamma||_1 = \eta||f_\gamma||_1 = \eta < \infty$.

(ii) The mML fractional Hawkes process ruled by the excitation function (2.3) *is well-defined according to Definition 2.1.*

The proof of this proposition is relegated in Appendix B.1. From this proposition, we see that the parameter η controls the clustering property of the financial events modeled by this mML fractional Hawkes process. Secondly, the following Proposition 3.2 is crucial for deriving the conditional transform of (λ, J) under the mML fractional Hawkes process. This proposition will indeed be fundamental in Section 4 in order to perform model estimation and risk management on illiquidity.

Proposition 3.2. *The modified Mittag-Leffler kernel* (2.3) *is completely monotone*[2]*. This is equivalent to the existence of a representation of the kernel* (2.3) *in terms of a Laplace-Stieltjes integral with non-decreasing density and non-negative measure given by*

$$f_\gamma(t) = \int_0^\infty e^{-\omega t}\mu(\mathrm{d}\omega)\,, \tag{3.2}$$

where

$$\mu(\mathrm{d}\omega) = \frac{-\sin(\gamma\pi)\,\omega^{\gamma-1}(1+\omega)}{\pi\left(\omega^{2\gamma} - 2\omega^\gamma\cos(\gamma\pi) + 1\right)}\mathrm{d}\omega\,. \tag{3.3}$$

We refer to Appendix B.2 for the proof. We then recall that the intensity of the mML fractional Hawkes process is obtained from (2.5) in the following way

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

$$\lambda_t = \lambda_0(t) + \eta \int_0^t f_\gamma(t-u)\mathrm{d}L_u = \lambda_0 + \int_0^t f_\gamma(t-u)\left(\alpha\left(\lambda_\infty - \lambda_u\right)\mathrm{d}u + \eta\,\mathrm{d}L_u\right), \quad t \geq 0. \tag{3.4}$$

This form for the intensity is in fact inspired by the affine Volterra structure of Abi Jaber et al. [29], Abi Jaber [30] with Ornstein-Uhlenbeck (OU) drift coefficient, and justified by Proposition 3.3 below. Such mML kernel with affine Volterra structure for λ will lead in the sequel to a semi-closed form solution for the conditional transform of N and L. Moreover, from the spectral representation (3.2) and using Fubini's theorem, we can rewrite the mML intensity (3.4) as

$$\lambda_t = \lambda_0 + \int_0^\infty \left(\int_0^t e^{-\omega(t-u)}\left(\alpha\left(\lambda_\infty - \lambda_u\right)\mathrm{d}u + \eta O \mathrm{d}N_u\right)\right)\mu(\mathrm{d}\omega) = \lambda_0 + \int_0^\infty Y_t^\omega \mu(\mathrm{d}\omega)\,. \tag{3.5}$$

In the last line, we denote by Y_t^ω a Markov process satisfying for each index $\omega \in \mathbb{R}^+$ the SDE

$$\mathrm{d}Y_t^\omega = \left(\alpha\left(\lambda_\infty - \lambda_t\right) - \omega Y_t^\omega\right)\mathrm{d}t + \eta O \mathrm{d}N_t\,, \quad Y_0^\omega = 0\,, \tag{3.6}$$

with solution

$$Y_t^\omega = \int_0^t e^{-\omega(t-u)}\left(\alpha(\lambda_\infty - \lambda_u)\,\mathrm{d}u + \eta O \mathrm{d}N_u\right). \tag{3.7}$$

From (3.5), 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 μ. We see this way that the mML fractional Hawkes intensity (3.4) 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 towards the level $\alpha(\lambda_\infty - \lambda_t)/\omega$. Hence, the mML 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 us to use the standard tools of stochastic calculus to derive its conditional transform from the affine property of equation (3.6).

Proposition 3.3. *There exists a unique non-negative (weak) solution for the intensity process λ satisfying (3.4).*

Proof. This directly comes from Theorem 2.13 of Abi Jaber [30] since the kernel f_γ in (3.2) is completely monotone and by taking $K = f_\gamma$, $g_0(t) = \lambda_0 + \int_0^t f_\gamma(t-u)\alpha\lambda_\infty \mathrm{d}u$ along with the semimartingale characteristics $b = 1-\eta$, $c = 0$, $\nu = \nu$. See also Lemma 9 of Bondi et al. [31] and Bondi et al. [32]. □

Finally, we confirm that the mML fractional Hawkes process can effectively exhibit long-range dependence under some condition. We first define the resolvent R as the unique solution to the linear Volterra equation $R(t) = f_\gamma(t) + \int_0^t R(t-s)f_\gamma(s)\mathrm{d}s$, $t \geq 0$, as well as the two functions $\mathcal{I}_R(t) := \int_0^t R(s)\mathrm{d}s$ and $\mathcal{I}_R^2(t) := \int_0^t \mathcal{I}_R(s)\mathrm{d}s$. Then, denoting W a one-dimensional standard Brownian motion and $\mathbb{D}$ the Skorokhod topology, we obtain the following proposition.

Proposition 3.4. *If $\eta\mathbb{E}[O] = 1$, then the following holds as $n \to \infty$,*

(1) $\sup_{t\in[0,T]}\left|\frac{N_{nt}}{\mathcal{I}_R^2(n)} - \lambda_0(t)t^\gamma\right| \to 0$ *in $\mathbb{P}$ and $L^2(\mathbb{P})$ for any $T \geq 0$.*

(2) $n|\mathcal{I}_R^2(n)|^{-3/2}\left(N_{nt} - \mathbb{E}[N_{nt}]\right) \to \sqrt{\lambda_0(t)\gamma}\int_0^t (t-s)^{\gamma-1}s^{(\gamma-1)/2}\,\mathrm{d}W_s$ *weakly in $\mathbb{D}(\mathbb{R}_+;\mathbb{R})$.*

Proof. This follows directly from Theorem 2.14 of Horst and Xu [33]. Indeed, it is straightforward to show from the Laplace transform (B.3) that $\Psi_1(t) := \int_0^t s\, f_\gamma(s)\mathrm{d}s = -\frac{\partial}{\partial_\lambda}\mathcal{L}_{f_\gamma}(\lambda)\Big|_{\lambda=0} = +\infty$ and that the mML fractional Hawkes process is strongly critical/quasi-stationary when $\eta\mathbb{E}[O] = 1$. It is also clear from the asymptotics (2.4) that $f_\gamma(t)$ is regularly varying at infinity of index $-\gamma$, and then from Karamata's theorem (see *e.g.* Proposition B.4 of [33]) that $\Phi(t) := \int_t^\infty f_\gamma(s)\mathrm{d}s$ is regularly varying at infinity of index $(-\gamma+1) \in [-1, 0]$. Therefore, we can apply Theorem 2.14 of Horst and Xu [33] with $\lambda_0(t)$ given as in (3.4), using Abi Jaber [30] and the completely monotone kernel f_γ. □

This proposition shows that when the mML fractional Hawkes process is critical, it displays long-range dependence and, more precisely, behaves asymptotically like a long-range dependent Gaussian process when suitably normalized. When $\eta\mathbb{E}[O] < 1$, the Hawkes process is subcritical/stationary and it is shown in Horst and Xu [33] that it does not display long-range dependencies. Finally, when $\eta\mathbb{E}[O] > 1$, the mML Hawkes process is a supercritical/non-stationary process for which long-range dependence cannot be defined. However, as seen in Fig. 1, it can still reproduce long-term effects thanks to the long-term properties/asymptotics of the kernel f_γ. Hence, we only consider from now on mML fractional Hawkes with $\eta\mathbb{E}[O] \geq 1$.

3.2. Discretization and conditional transform

Following Bäuerle and Desmettre [17], we now approximate the measure μ derived in (3.3) 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 $(\xi_k^{(n)}, \xi_{k+1}^{(n)})$ for $k = 0, \ldots, n-1$ is given by

$$b_{k+1}^{(n)} = \left(\int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} \omega\, \mu(\mathrm{d}\omega)\right) \Big/ \left(\int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} \mu(\mathrm{d}\omega)\right). \tag{3.8}$$

For $k = 0, \ldots, n-1$, the mass of any atom on each interval $\left(\xi_k^{(n)}, \xi_{k+1}^{(n)}\right)$ is hence

$$m_{k+1}^{(n)} = \int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} \mu(\mathrm{d}\omega) = \int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} \frac{-\sin(\gamma\pi)\,\omega^{\gamma-1}(1+\omega)}{\pi\left(\omega^{2\gamma} - 2\omega^\gamma\cos(\gamma\pi) + 1\right)}\, \mathrm{d}\omega\, . \tag{3.9}$$

We cannot obtain analytical expressions for $m_{k+1}^{(n)}$ and $b_{k+1}^{(n)}$. However, the exact expressions for the barycenter and the mass of each interval are not useful in what follows. The discrete measure for a partition of size n is then defined as follows

$$\mu^{(n)}(\omega) := \sum_{k=1}^{n} m_k^{(n)} \delta_{b_k^{(n)}}(\omega)\, , \tag{3.10}$$

where $\delta_{b_k^{(n)}}(\omega)$ 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 $\mathcal{E}^{(n)}$:

(i) $\xi_0^{(n)} \to 0$ and $\xi_n^{(n)} \to \infty$ when $n \to \infty$.

(ii) $\max\left|\xi^{(n)}_{k+1} - \xi^{(n)}_{k}\right| \to 0$ when $n \to \infty$.

(iii) $\mathcal{E}^{(n)} \subset \mathcal{E}^{(n+1)}$.

Considering the previous discretization setting, the following proposition connects the discretized measure $\mu^{(n)}$ to the true measure μ for any measurable function in the set $L^1(\mu)$ of integrable functions with respect to this measure μ.

Proposition 3.5. *For $n \in \mathbb{N}$, if a partition $\mathcal{E}^{(n)}$ satisfies (i)-(iii) and if a function $f : \mathbb{R} \to \mathbb{C}$ is in $L^1(\mu)$, then*

$$\lim_{n\to\infty} \int_0^\infty f(\omega)\mu^{(n)}(\mathrm{d}\omega) = \int_0^\infty f(\omega)\mu(\mathrm{d}\omega)\,. \tag{3.11}$$

We refer to Appendix B.3 for the proof, see also Bäuerle and Desmettre [17]. We then fix $n \in \mathbb{N}$ and adopt the following notation when discretizing the mean-reverting factor Y^ω defined in (3.7) :

$$Y^{(k)}_t := Y^{b^{(n)}_k}_t,\ \forall k = 1, \ldots, n,$$

where

$$\mathrm{d}Y^{(k)}_t = \left(\alpha\left(\lambda_\infty - \lambda^{(n)}_t\right) - b^{(n)}_k Y^{(k)}_t\right) + \eta\, O\, \mathrm{d}N^{(n)}_t\,,$$

and with $N^{(n)}$ a counting process of intensity $\lambda^{(n)}$. From equation (3.5), this discretized intensity $\lambda^{(n)}$ based on a partition of size n is given by

$$\lambda^{(n)}_t = \lambda_0 + \sum_{k=1}^{n} m^{(n)}_k Y^{(k)}_t\,. \tag{3.12}$$

From Proposition 3.5, we have that $\lim_{n\to\infty} \lambda^{(n)}_t = \lambda_t$ almost surely since we can rewrite equation (3.12) as[3] $\lambda^{(n)}_t = \lambda_0 + \int_0^\infty Y^\omega_t \mu^{(n)}(\mathrm{d}\omega)$ and since the function $\omega \in \mathbb{R}^+ \mapsto Y^\omega_t(w')$ is integrable with respect to the measure μ for fixed $w' \in \Omega$ as a sum of negative exponential functions, see equation (3.7). Finally, we obtain

$$\mathrm{d}\lambda^{(n)}_t = \sum_{k=1}^{n} m^{(n)}_k \mathrm{d}Y^{(k)}_t = \sum_{k=1}^{n} m^{(n)}_k \left(\alpha\left(\lambda_\infty - \lambda^{(n)}_t\right) - b^{(n)}_k Y^{(k)}_t\right) + \eta O \sum_{k=1}^{n} m^{(n)}_k \mathrm{d}N^{(n)}_t.$$

Now that we have discretized the mML fractional Hawkes process into a finite number of related Markov processes, we are able to obtain the main mathematical result of the paper about its conditional transform with Proposition 3.6 and Corollary 3.7. These tools are indeed required for modeling illiquidity based on the mML Hawkes process in Section 4.2 (see equations (4.6)-(4.7)), in Section 4.3 and in Section 4.4.

Proposition 3.6. *The transform of $X^{(n)}_T = \left(\lambda^{(n)}_T, J^{(n)}_T\right)^\top = \left(\lambda^{(n)}_T, \left(L^{(n)}_T, N^{(n)}_T\right)\right)^\top$ conditionally to $\mathcal{F}_t$ at time $t \le T$ with $u = (u_\lambda, u_L, u_N) \in \mathbb{C}^3$ is given by*

$$\mathbb{E}_t\left[e^{u\cdot X^{(n)}_T}\right] = \mathbb{E}_t\left[e^{u_\lambda \lambda^{(n)}_T + u_L L^{(n)}_T + u_N N^{(n)}_T}\right] = \exp\left(a(t,T) + b(t,T)\lambda^{(n)}_t + u'\cdot{}^\top J^{(n)}_t + \sum_{k=1}^{n} c_k(t,T)\, Y^{(k)}_t\right), \tag{3.13}$$

[3]Given that we have $\int_0^\infty Y^\omega_t \mu^{(n)}(\mathrm{d}\omega) = \sum_{k=1}^n m^{(n)}_k \int_0^\infty Y^\omega_t \delta_{b^{(n)}_k}(\mathrm{d}\omega) = \sum_{k=1}^n m^{(n)}_k Y^{b^{(n)}_k}_t = \sum_{k=1}^n m^{(n)}_k Y^{(k)}_t$.

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 $c_k(t,T) := c_k(u',t,T)$ satisfy the following system of equations

$$
\begin{aligned}
a(t,T) &= \alpha\,\lambda_\infty \sum_{k=1}^{n} m_k^{(n)} \left(\int_t^T e^{-b_k^{(n)}(v-t)} b(v,T)\mathrm{d}v \right), \\
\frac{\partial b(t,T)}{\partial t} &= \alpha \left(\sum_{k=1}^{n} m_k^{(n)} b(t,T) + \sum_{k=1}^{n} c_k(t,T) \right) - \theta \left(b(t,T)\,\eta \sum_{k=1}^{n} m_k^{(n)} + u' \cdot (1,0)^\top + \sum_{k=1}^{n} c_k(t,T)\,\eta \right) e^{u'\cdot(0,1)^\top} + 1\,, \\
c_k(t,T) &= -b_k^{(n)}\, m_k^{(n)} \int_t^T e^{-b_k^{(n)}(v-t)}\, b(v,T)\,\mathrm{d}v\,, \quad \forall k = 1, \dots n.
\end{aligned}
$$

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

The proof of this proposition is given in Appendix B.4. Passing to the limit $n \to \infty$ leads to the following corollary.

Corollary 3.7. *The transform of $X_T = (\lambda_T, J_T)^\top = (\lambda_T, (L_T, N_T))^\top$ conditionally to $\mathcal{F}_t$ at time $t \le T$ with $u = (u_\lambda, u_L, u_N) \in \mathbb{C}^3$ is given by*

$$
\begin{aligned}
\mathbb{E}_t\left[e^{u\cdot X_T}\right] &= \mathbb{E}_t\left[e^{u_\lambda \lambda_T + u_L L_T + u_N N_T}\right] \qquad (3.14) \\
&= \exp\left(\alpha\,\lambda_\infty \int_t^T f_\gamma(v-t)\, b(v,T)\,\mathrm{d}v + b(t,T)\lambda_t + u' \cdot J_t^\top - \int_0^\infty \left(\int_t^T e^{-\omega(v-t)} b(v,T)\,\mathrm{d}v \right) \omega\, Y_t^\omega \mu(\mathrm{d}\omega) \right),
\end{aligned}
$$

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

$$
\frac{\partial b(t,T)}{\partial t} = -\,\alpha\,\frac{\partial}{\partial t} \int_t^T f_\gamma(v-t)\, b(v,T)\,\mathrm{d}v - \theta \left(u' \cdot (1,0)^\top - \eta \frac{\partial}{\partial t} \int_t^T f_\gamma(v-t)\, b(v,T)\,\mathrm{d}v \right) e^{u'\cdot(0,1)^\top} + 1\,, \qquad (3.15)
$$

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

The proof of this corollary is again relegated in Appendix B.5. Based on the conditional transform (3.14), we can now obtain the moments of the process $X = (\lambda, J)^\top$ given in the following corollary.

Corollary 3.8. *The conditional expectation of X_T to $\mathcal{F}_t$ at time $t \le T$ is given by*

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

for $w = (w_\lambda, w_L, w_N) \in \mathbb{R}^3$. We denote $w' = (w_L, w_N) \in \mathbb{R}^2$, $u' = (u_L, u_N) \in \mathbb{C}^2$ and $\chi = \mathbb{E}[O] = \int_0^\infty o\,\nu(\mathrm{d}o)$. The coefficient function

$$
b'(0,t,T) := \left(w_L \frac{\partial}{\partial_{u_L}} + w_N \frac{\partial}{\partial_{u_N}} \right) b(u',t,T) \Big|_{u'=(0,0)} \qquad (3.16)
$$

satisfies

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

See Appendix B.6 for the proof. Note that Corollaries 3.7 and 3.8 drastically simplify at initial time $t = 0$ since $Y_0^\omega = 0$, $\forall \omega \in \mathbb{R}^+$. Similarly, higher order moments can be obtained by successive conditioning and differentiating. Moreover, the ACF of the mML Hawkes process N with $\eta\mathbb{E}[O] = 1$ and $\alpha = 0$ can be obtained using the second order properties derived in Proposition 4 of Bacry et al. [2] since the Laplace transform of f_γ is known in closed form via (B.3). This allows a comparison with the ACF of the number of jumps N based on the exponential, power-law and dML kernels, see Fig. 6. Finally, it is straightforward to extend the previous univariate setting by considering a multivariate mML fractional Hawkes process where the intensity $\lambda^{(i)}$ of each component process $i = 1, \dots, d$ may also depend on several of the other Hawkes components. We refer to Appendix B.7 for more details. From the mathematical properties (and particularly the conditional transforms) derived in this section, we are now ready to apply the mML fractional Hawkes process for modeling the illiquidity associated with low-frequency observations of asset prices.

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 market 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. [34] 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 initially introduced in Amihud [19]. This measure has proved to be very popular in the empirical literature, 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 *e.g.* Amihud [20]. This illiquidity premium indeed results from the fact that investors care about illiquidity costs and therefore price them in the expected returns for compensating this price/market impact.

Many econometric models in discrete time have then been studied in the literature for modeling illiquidity based on this Amihud measure, see Amihud [20], Brennan et al. [21], Lou and Shu [35] or more recently Hafner et al. [22]. We propose in this work a new approach by introducing instead a continuous-time process for modeling the log-Amihud measure. More precisely, this log-Amihud measure is modeled as a mean-reverting jump diffusion process with jumps following a fractional Hawkes process driven by a mML intensity with the multidimensional affine structure (B.11). 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 is also suitable for simultaneously studying the illiquidity process of multiple stocks and indices in financial markets. We will finally show in Section 4.4 that employing such a continuous-time process facilitates effective and straightforward risk management on illiquidity.

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

$$A_t = \frac{|R_t|}{V_t}\,, \tag{4.1}$$

where R_t is the stock return and V_t is the trading volume (in currency) at time t. We therefore see that the Amihud measure captures the fact that a stock is less liquid (more illiquid) if a given trading volume generates a larger move in its price. In practice, the measure is typically computed over low-frequency periods ranging from a day to a year by averaging the daily illiquidity ratio over the corresponding period p using $A_j := 1/p\sum_{k=0}^{p-1}|R_{t_{j\cdot p-k}}|/V_{t_{j\cdot p-k}}$ for $j = 1,\ldots,N$. We consider for the remaining of this work a daily sampling period with $p = 1$ (hence $A_j = A_{t_j}$) 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 as this standardizes its values, see Amihud [19], Brennan et al. [21], Lou and Shu [35]. We then denote $a_t := \log A_t$ in the sequel of the paper. Some descriptive statistics of the stock returns, traded volumes, Amihud and log-Amihud measures are given for the FTSE 100 in the data section of Appendix A, Table 3. Moreover, 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 also pictured in Fig. 5 for the FTSE 100 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 with some shocks that do not display any clear trend: negative abrupt movements followed by abrupt movements of either sign. We also observe some clustering with serial dependence, which is confirmed empirically with the autocorrelogram (ACF) of the log-Amihud measure a_t in Fig. 6. As defined in Chapter 2.1 of Pipiras and Taqqu [36], the sub-exponential decay of the ACF (and more precisely, its power-law decay at large lags) clearly indicates long-range dependence/memory in a_t. Moreover, we directly see that the ACF of the mML critical Hawkes process N is the only one able to reproduce both the steep exponential decay at short lags and the power-law one at larger lags, as investigated in Section 2[4]. This confirms the ability of this mML kernel to better reproduce the memory properties of the observed log-Amihud illiquidity process. Finally, the Histogram 7 of the FTSE100 log-Amihud increments Δa_t shows that they are close to Gaussian. However, standard Gaussian processes fail to reproduce the high kurtosis and ACF of the (log-)Amihud measure

[4] We also see that the power-law critical Hawkes N with $\gamma \approx 0$ provides a good fit of this ACF, inducing long-range dependence as shown in Jaisson [26] (even if it prevents from having tractable formulae, see Section 2, similarly as for the dML kernel).

(see Table 3 and Fig. 6) and need unrealistic parameter values for reproducing the abrupt movements of Fig. 5

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 [35] 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 Andersen et al. [37] or Gatheral et al. [38]. Finally, by incorporating jumps with a mML intensity, the model can effectively reproduce both the observed illiquidity shocks as well as the long-range dependence of the Amihud illiquidity measure, as highlighted above in the ACF Fig. 6 and in Lobato and Velasco [39]. We therefore consider the modified Mittag-Leffler Ornstein-Uhlenbeck (mML OU) process defined by the following dynamic of the log-Amihud measure $a_t = \log A_t$:

$$\begin{aligned} \mathrm{d}a_t &= \theta_1\,(\theta_2 - a_t)\,\mathrm{d}t + \theta_3\,\mathrm{d}W_t^{(a)} + \mathrm{d}L_t^{(1)} - \mathrm{d}L_t^{(2)} \\ &= \theta_1\,(\theta_2 - a_t)\,\mathrm{d}t + \theta_3\,\mathrm{d}W_t^{(a)} + O^{(1)}\mathrm{d}N_t^{(1)} - O^{(2)}\mathrm{d}N_t^{(2)}\,, \end{aligned} \tag{4.2}$$

where $L^{(1)}$ and $L^{(2)}$ are the loss processes defined by equation (2.1) of the up jump and down jump components, where $W^{(a)}$ is a standard Brownian motion independent from the two jump components and where $O^{(i)}$ represents the positive jump size/illiquidity shocks of each type $i = 1, 2$ with distribution $\nu^{(i)}$ on $(0, \infty)$. For simplicity, an exponential distribution with rate $\rho^{(i)}$ is chosen for $\nu^{(i)}$ in the sequel. The conditional intensity $\lambda^{(i)}$ of the counting process $N^{(i)}$ is given[5] for $i = 1, 2$ by

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

Note first that modeling the log-Amihud measure as a mML OU model guarantees the positivity of the Amihud measure and provides a semi-closed form formula for its conditional transform, as we will show in Section 4.3. Secondly, this model (4.2) allows to study the peaks of (il)liquidity of various assets while having a dependence and contagion between the up and down shocks of the log-Amihud measure. Finally, from Proposition 3.4, the mML fractional Hawkes process (4.3) can capture the long-range dependence (if $\varrho\big(\big(\widehat{\rho}^{(j)}\widehat{\eta}^{(i,j)}\big)_{i,j=1,2}\big) = 1$) or the long-term effects (if $\varrho\big(\big(\widehat{\rho}^{(j)}\widehat{\eta}^{(i,j)}\big)_{i,j=1,2}\big) > 1$) of these (il)liquidity shocks while ensuring good properties of the intensity process.

4.2. Estimation procedure

We now aim at estimating the log-Amihud dynamic of the FTSE 100 index based on equation (4.2). For estimating this mML OU process, we use a modified Peaks Over Threshold (POT) procedure, as described in Hainaut and Moraux [40]. This method is divided into three steps. In the first step, from the time series

[5]See equation (B.11) but without the Brownian component $W^{(i)}$, *i.e.* $\sigma^{(i)} = 0$ for $i = 1, 2$ in the intensities (4.3).

of daily log-Amihud measures a_{t_j} ($j = 1, \ldots, N$) of the FTSE 100 index (cfr Appendix A), we estimate the parameters $(\theta_1, \theta_2, \theta_3)$ of a classical OU process via maximum likelihood without considering the jump components $L^{(1)}$ and $L^{(2)}$. 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

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

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

where $\phi^{-1}(\alpha_1)$ and $\phi^{-1}(\alpha_2)$ are the Normal quantiles at probability levels α_1 and α_2. The levels α_1 and α_2 are chosen such that the sample without jumps $\left\{ a_{t_j} \,\middle|\, a_{t_j} \in \left[\mathrm{Tr}_j^{(1)}(\alpha_1) \,,\; \mathrm{Tr}_j^{(2)}(\alpha_2)\right] \right\}$ 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, *i.e.*

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

The set of the $n_T^{(1)}$ up jumps and $n_T^{(2)}$ down jumps over $[0, T]$ is then given by

$$\left\{ \tau_1^{(1)}, \ldots, \tau_{n_T^{(1)}}^{(1)} \right\} = \left\{ t_j \,\middle|\, a_{t_j} > \mathrm{Tr}_j^{(1)}(\alpha_1^*) \right\} \text{ and } \left\{ \tau_1^{(2)}, \ldots, \tau_{n_T^{(2)}}^{(2)} \right\} = \left\{ t_j \,\middle|\, a_{t_j} < \mathrm{Tr}_j^{(2)}(\alpha_2^*) \right\},$$

respectively. In the case of the FTSE 100, we find the following Fig. 2 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 $\alpha_2^* = 0.0191$). Considering these thresholds, we re-estimate the parameter θ_3 such that the sample without

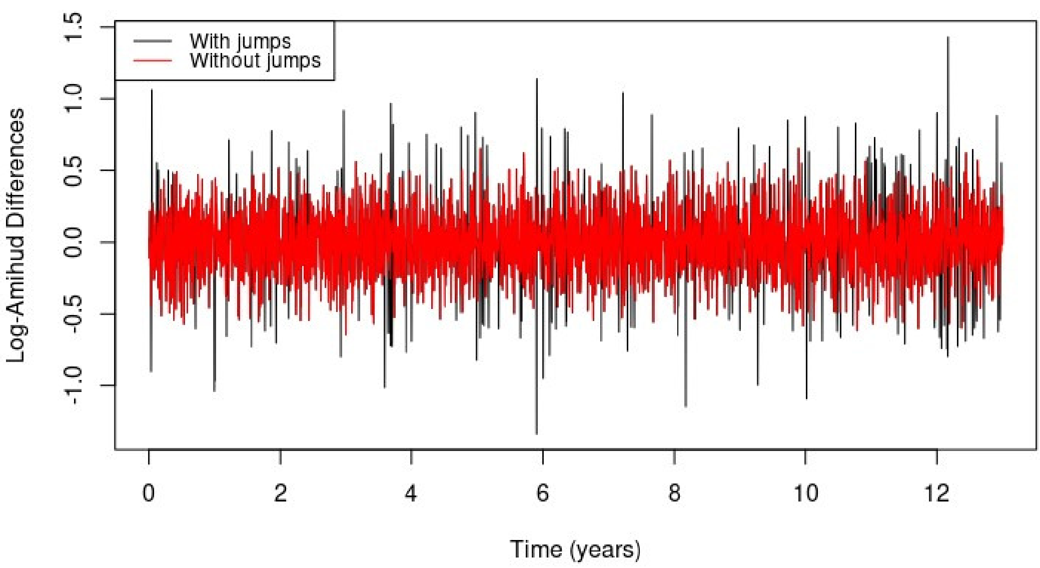

Fig. 2 Daily log-Amihud measure in difference Δa_{t_j} ($j = 1, \ldots, 3266$) for the FTSE 100 index over a period T of 13 years (see Appendix A). In red, sample path without jumps and in black, sample path with jumps.

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 $a_{t_j} \in \left[\mathrm{Tr}_j^{(1)}(\alpha_1^*), \mathrm{Tr}_j^{(2)}(\alpha_2^*)\right]$ with

$\widehat{\theta}_1$ and $\widehat{\theta}_2$ the estimators found at the first calibration step. When a jump is detected at time $\tau_l^{(i)}$ $(i=1,2)$, the jump size $\Delta L^{(i)}_{\tau_l^{(i)}}$ is then equal to the absolute variation of the log-Amihud measure, *i.e.*

$$\Delta L^{(i)}_{\tau_l^{(i)}} = \left| \Delta a_{\tau_l^{(i)}} - \widehat{\theta}_1 \left(\widehat{\theta}_2 - a_{\tau_l^{(i)}-1} \right) \Delta t \right| .$$

Since the jump sizes $O^{(i)}$ $(i=1,2)$ of the up and down jumps are independent and identically distributed exponential random variables with parameters $\rho^{(i)} > 0$ and density

$$\nu^{(i)}(z \,|\, \rho^{(i)}) = \rho^{(i)} \exp\left(-\rho^{(i)} z\right), \quad z > 0 ,$$

we can estimate the jump parameter $\rho^{(i)}$ by log-likelihood maximization

$$\rho^{(i)} = \arg\max \sum_{l=1}^{n_T^{(i)}} \log \nu^{(i)} \left(\Delta L^{(i)}_{\tau_l^{(i)}} \,\middle|\, \rho^{(i)} \right) .$$

Based on the jump times $\left\{ \tau_1^{(i)}, \ldots, \tau^{(i)}_{n_T^{(i)}} \right\}$ of each type i, we can finally obtain via Embrechts et al. [41] the log-likelihood of the mML Hawkes process with intensity (4.3), which is given by

$$\log L \left(\sum_{i=1}^{d} \sum_{l=1}^{n_T^{(i)}} \tau_l^{(i)} \,\middle|\, \theta \right) = -\sum_{i=1}^{d} \int_0^T \lambda_s^{(i)} \,\mathrm{d}s + \sum_{i=1}^{d} \sum_{l=1}^{n_T^{(i)}} \log \left(\lambda^{(i)}_{\tau_l^{(i)}} \right), \tag{4.6}$$

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 (2.6), the expression (4.6) can be drastically simplified, see Errais et al. [6], which speeds up computation. In our mML case, knowing the up and down jumps $\left\{ \tau_1^{(i)}, \ldots, \tau^{(i)}_{n_T^{(i)}} \right\}_{i=1,2}$ with their jump size $\Delta L^{(i)}_{\tau_l^{(i)}}$ allows us to use the spectral representation (B.12) and (B.13) along with their discretized counterparts (B.15) and (B.16) for $i=1,2$ based on a partition of size n (see Section 3.2). This leads to an efficient computation of the discretized intensity $\lambda^{(n,i)}$ and of the log-likelihood (4.6), based on

$$\lambda^{(n,i)}_{t_j} = \lambda_0^{(i)} + \sum_{k=1}^{n} m_k^{(n,i)} Y^{(k,i)}_{t_j} , \tag{4.7}$$

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

$$Y^{(k,i)}_{t_{j+1}} = Y^{(k,i)}_{t_j} + \left(-b_k^{(n)} Y^{(k,i)}_{t_j} + \alpha^{(i)} \left(\lambda_\infty^{(i)} - \lambda^{(n,i)}_{t_j} \right) \right) \Delta t + \eta^{(i,1)} \Delta L^{(n,1)}_{t_j} + \eta^{(i,2)} \Delta L^{(n,2)}_{t_j} .$$

Restricting the number of factors n to 50 000 with $\Delta t = 0.001$ gives very accurate results and drastically reduces the computing time. A similar application of such technique with detailed convergence results can be found in Abi Jaber [42]. Note that since this discretized representation (4.7) cannot be written for the power-law and dML kernels (2.12)-(2.15), the computation of the log-likelihood (4.6) is much more slow in these cases. 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 1. We first observe from representation

Table 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 the POT procedure on the FTSE 100 index based on a 13-year period of daily log-Amihud measures a_{t_j} $(j = 1, \ldots, 3266)$, see data section in Appendix A.

	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	
$\gamma^{(1)}$			1.029	0.064
$\lambda_0^{(1)}$			3.840	0.68
$\gamma^{(2)}$			1.024	0.021
$\lambda_0^{(2)}$			3.022	0.64
$\alpha^{(1)}$			6.048	0.264
$\alpha^{(2)}$			0.451	0.422
$\lambda_\infty^{(1)}$			4.721	0.812
$\lambda_\infty^{(2)}$			3.386	0.774
$\eta^{(1,1)}$			0.168	0.221
$\eta^{(1,2)}$			9.062	3.694
$\eta^{(2,1)}$			3.359	1.012
$\eta^{(2,2)}$			0.195	1.029
Log-likelihood			140.188	

(B.12)-(B.13) that the intensities $\lambda_t^{(1)}$ and $\lambda_t^{(2)}$ mean revert towards a value $\widehat{\lambda}_\infty^{(1)} = 4.721$ and $\widehat{\lambda}_\infty^{(2)} = 3.386$, respectively (at a rate depending on $\widehat{\alpha}^{(1)} = 6.048$ and $\widehat{\alpha}^{(2)} = 0.451$). We also see that the parameters $\widehat{\gamma}^{(1)}$, $\widehat{\gamma}^{(2)}$ capturing the long-term power-law decay of the illiquidity peaks in the log-Amihud measure are both very close to 1. Hence, this confirms the strong long-term effects of these illiquidity peaks. Moreover, based on equation (2.4) and Section 2.4, 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 for mid-price changes at high frequency, which are based on the power-law kernel (2.12) with $\gamma \approx 0$ (and hence also with a power-law exponent close to $\gamma + 1 \approx 1$), cfr Bacry and Muzy [1]. This is also in line with the observed ACF of a_t in Fig. 6. Concerning the self-exciting parameters $\widehat{\eta}^{(1,1)}$ and $\widehat{\eta}^{(2,2)}$, we see that they are positive which confirms the presence of clustering effects in the intensity of the jumps. Moreover, these parameters are lower than the cross-excitation parameters $\widehat{\eta}^{(1,2)}$ and $\widehat{\eta}^{(2,1)}$ (which are also positive), leading to a predominant cross-excitation effect. Testing for stationarity in our bivariate setting (cfr Section 2), we obtain that the spectral radius $\varrho\big(\big(\widehat{\rho}^{(j)}\widehat{\eta}^{(i,j)}\big)_{i,j=1,2}\big)$ is equal to 3.79. 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 $\widehat{\theta}_1$ and of the volatility $\widehat{\theta}_3$. Computing the log-Amihud measure based on a weekly ($p = 5$) or monthly ($p = 21$) time window would lead to less extreme values. Another alternative frequently used in the literature consists in averaging this measure between different stocks/indices, as done in Amihud [20]. The large standard errors of the self and cross-exciting parameters in Table 1 are explained by the limited number of up and down jumps used for estimating the model, leading to numerical instability. Finally, other distributions $\nu^{(i)}$ with positive support could have been chosen for the jump sizes $O^{(i)}$ (Log-normal, Gamma, etc). Determining the one leading to the best goodness of fit is let to further study.

We now consider other kernels as excitation function $f_{\gamma^{(i)}}(\cdot)$ in the jumps' intensity (4.3) and compare the results to Table 1. With the exponential kernel (2.6), we obtain the following estimated parameters $\widehat{\alpha}^{(1)} = 75.103$, $\widehat{\lambda}_0^{(1)} = 4.755$, $\widehat{\lambda}_\infty^{(1)} = 4.055$, $\widehat{\eta}^{(1,1)} = 0.051$, $\widehat{\eta}^{(1,2)} = 19.901$, $\widehat{\alpha}^{(2)} = 0.366$, $\widehat{\lambda}_0^{(2)} = 2.619$, $\widehat{\lambda}_\infty^{(2)} = 3.212$, $\widehat{\eta}^{(2,1)} = 0.454$, $\widehat{\eta}^{(2,2)} = 0.002$, leading to a log-likelihood (4.6) of 130.9. With the power-law kernel (2.12), we find the estimated parameters $\widehat{\beta}^{(1)} = 5.992$, $\widehat{\gamma}^{(1)} = 12.528$, $\widehat{\lambda}_0^{(1)} = 4.647$, $\widehat{\eta}^{(1,1)} = 0.011$, $\widehat{\eta}^{(1,2)} = 3.615$, $\widehat{\beta}^{(2)} = 0.161$, $\widehat{\gamma}^{(2)} = 15.67$, $\widehat{\lambda}_0^{(2)} = 2.750$, $\widehat{\eta}^{(2,1)} = 2.467$, $\widehat{\eta}^{(2,2)} = 0$ and a log-likelihood equal to 131.885. Using the AIC criterion to compare these models, we find that the mML Hawkes process gives the best results and hence better fits the time series of up and down jumps. The ML and dML kernels can only be estimated if $\alpha^{(i)} = 0$ (no mean-reversion in the intensity (4.3)) and also lead to a lower log-likelihood and a worse AIC criterion.

4.3. Characteristic function of the log-Amihud measure

Before performing risk management on the Amihud measure, we first need to derive the characteristic function $\varphi(t,T,\xi) := \mathbb{E}\left[\exp\left(\xi\, a_T\right) \mid \mathcal{F}_t\right]$ of the log-Amihud random variable a_T for $\xi \in \mathbb{C}$. We consider here the log-Amihud dynamic (4.2) with intensity (4.3), as proposed in the previous section where $i = 1, 2$ for the up and down jumps, respectively. We can then apply exactly the same proof as for Proposition B.7 in Appendix B.7 in order to obtain the following corollary.

Corollary 4.1. *The characteristic function of the log-Amihud measure a_T satisfying (4.2) conditionally to $\mathcal{F}_t$ at time $t \leq T$ with $\xi \in \mathbb{C}$ is given by*

$$\begin{aligned}\varphi(t,T,\xi) = \exp\Bigg[& A(t,T) + B(t,T)\cdot {}^{\top}\lambda_t + D(t,T)\, a_t \\ & + \sum_{i=1}^{2}\Bigg(\alpha^{(i)}\lambda_\infty^{(i)}\int_t^T f_{\gamma^{(i)}}(v-t)\, B^{(i)}(v,T)\mathrm{d}v - \int_0^\infty\left(\int_t^T e^{-\omega(v-t)}B^{(i)}(v,T)\mathrm{d}v\right)\omega Y_t^{\omega,(i)}\mu^i(\mathrm{d}\omega)\Bigg)\Bigg].\end{aligned} \tag{4.8}$$

where the dynamic of $Y_t^{\omega,(i)}$ is given by equation (B.13) *with $\sigma^{(i)} = 0$, $i = 1, 2$. The boundary conditions are given by $A(T,T) = 0$, $D(T,T) = \xi$ and $B(T,T) = (0,0)$. Each component $B^{(i)}(t,T)$ of the coefficient vector $B(t,T)$ satisfies the following integro-differential equation for $i = 1, 2$:*

$$\begin{aligned}\frac{\partial B^{(i)}(t,T)}{\partial t} &= -\,\alpha^{(i)} \frac{\partial}{\partial t} \int_t^T f_{\gamma^{(i)}}(v-t)\, B^{(i)}(v,T)\,\mathrm{d}v \\ &\quad -\,\theta^{(i)} \left(-\sum_{j=1}^{2} \eta^{(j,i)} \frac{\partial}{\partial t} \int_t^T f_{\gamma^{(j)}}(v-t)\, B^{(j)}(v,T)\,\mathrm{d}v + (-1)^{i-1}\, D(t,T) \right) + 1.\end{aligned} \tag{4.9}$$

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

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

Finally, the coefficient $A(t,T) = \int_t^T \left(\theta_1 \theta_2\, D(v,T) + \frac{\theta_3^2}{2}\, D(v,T)^2\right) \mathrm{d}v$ is equal to

$$A(t,T) = \xi\, \theta_2 \left[1 - e^{\theta_1(t-T)}\right] + \frac{\theta_3^2\, \xi^2}{4\, \theta_1} \left[1 - e^{2\theta_1(t-T)}\right].$$

We do not give a dedicated proof for this corollary since it is directly derived from the proof of Proposition B.7 in Appendix B.7 and from the PIDE (2.11).

4.4. ***Risk management on illiquidity***

We first use Corollary 4.1 with the obtained transform (4.8) 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 of the Amihud measure 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 $k \in \mathbb{N}$. The transform $\varphi(t,T,\xi)$ is obtained thanks to equation (4.8) of the previous corollary with the estimated parameters from Table 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 inverting the characteristic function

$$\mathbb{E}_t\left[e^{iz a_T}\right] = \int_0^\infty e^{iz a_1}\, p(T, a_1 \,|\, t, a_2)\,\mathrm{d}a_1 ,$$

it is indeed possible to determine the conditional probability density $p(T, \cdot \,|\, t, a_2)$ of a_T, which is defined for $a_1, a_2 \in \mathbb{R}$ by

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

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

$$VaR_\varepsilon(A_T) = \inf\left\{q \in \mathbb{R} \,|\, \mathbb{P}(A_T \le q) \ge \varepsilon\right\} = \inf\left\{ q \in \mathbb{R} \,\middle|\, \int_{-\infty}^{\log q} p(T, a_1 \mid 0, a_2)\,\mathrm{d}a_1 \ge \varepsilon \right\}.$$

We now use the following definition for the TVaR of the Amihud measure at level ε, $TVaR_\varepsilon(A_T) := \mathbb{E}\left[A_T|A_T > VaR_\varepsilon(A_T)\right]$, since we are interested in the right tail of the distribution. It can be rewritten in the following way

$$TVaR_\varepsilon(A_T) = \frac{1}{1-\varepsilon}\int_{\log(VaR_\varepsilon(A_T))}^{+\infty} e^{a_1}\, p(T, a_1 \mid 0, a_2)\, \mathrm{d}a_1$$

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 [43] or Albanese et al. [44] for a review. With the model parameters estimated from the POT procedure above in Table 1, we obtain in Table 2 the moments and risk measures of the Amihud measure A_T for $T = 1$ year. A comparison with the classical OU model based again on the parameters $(\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3)$ of Table 1 is also performed and the density of the Amihud and log-Amihud measures is depicted for each model in Fig. 4 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 2 The first columns of the table give the first four central moments of the Amihud measure A_T at $T = 1$ under the classical OU model and the mML OU model with estimated parameters of Table 1, obtained from the daily log-Amihud measures of the FTSE 100 index (01/2009 - 01/2022), see Appendix A. The two last columns give the respective $VaR_\varepsilon(A_T)$ and $TVaR_\varepsilon(A_T)$ at level $\varepsilon = 99.5\%$.

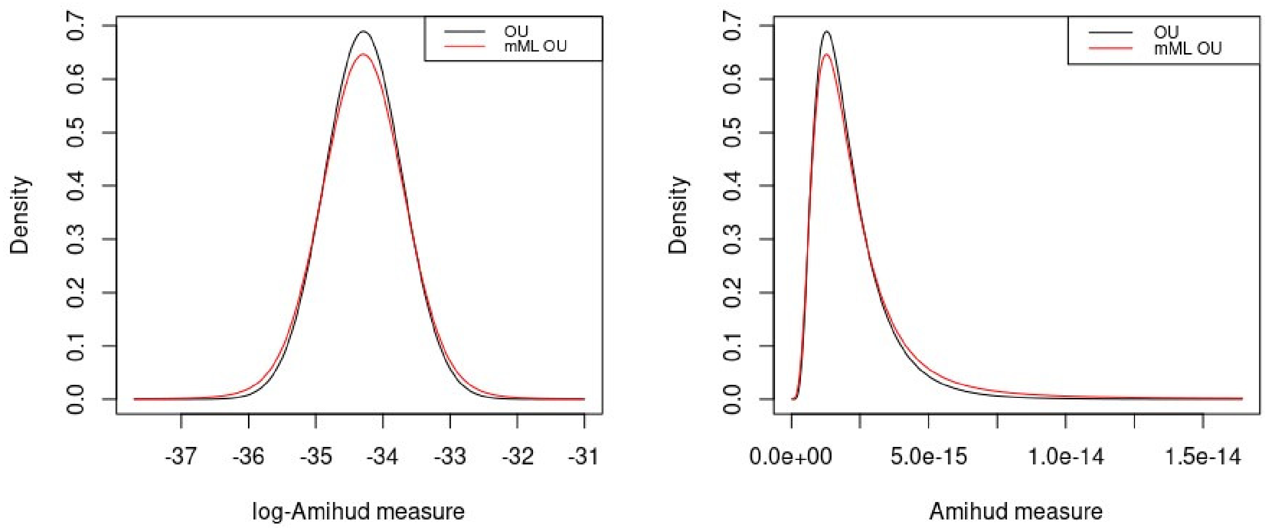

Fig. 4 Density of the log-Amihud measure a_T (left) and of the Amihud measure A_T (right) for $T = 1$ year under the classical OU model (black line) and the mML OU model (red line), again based on the estimates of Table 1 for the FTSE 100 index.

We clearly see from Table 2 and Fig. 4 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 parameters θ and ρ between the up and down jumps are very close, the effect of the mML OU model is relatively small on the skewness (and mean) 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. Finally, if we compare Table 2 with the true descriptive statistics of the Amihud measure for the FTSE 100 index in Table 3 (third column), we see that the mML OU model indeed provides much more accurate results than the OU model, confirming the ability of this model to better reproduce the observed properties of illiquidity data.

However, limited insight can be provided regarding risk measures for OU models with ML or dML Hawkes jumps. The ML kernel (2.13) proves unfeasible as it is not properly defined (see Section 2.3) with an exploding average intensity and hence exploding moments for the (log-)Amihud measures a_t and A_t. The dML kernel (2.15) is much more suitable for the comparison but requires Monte-Carlo simulations (see *e.g.* the thinning algorithm of Ogata [45]), which are highly unstable and time-consuming due to the self-exciting and non-Markovian nature of N in these models. Indeed, these simulations are not sufficiently accurate for computing precisely higher order moments (skewness or kurtosis) or tail measures (VaR or TVaRs). This in fact outlines the nice properties of the mML kernel which enables tractable formula for the characteristic function and hence facilitates risk management through the numerical computation of the associated pdf.

Finally, as mentioned in Section 4, fund managers have the ability to time market liquidity through adjusting their portfolio's market exposure to the observed and anticipated liquidity conditions, highlighting the importance of understanding and incorporating liquidity in investment decision making, see Cao et al. [34]. Therefore, an interesting application could be to perform optimal stopping on the mML OU model (4.2)-(4.3) for the log-Amihud measure such as to find the right buying time to maximize the liquidity of the related asset (and hence minimize trading costs/price impacts). No interesting theoretical result could be found but this problem can still be solved numerically using the discretized setting proposed in Section 3.2, see Bäuerle and Desmettre [17] and p.3773-3775 of Floudas and Pardalos [46]. Moreover, this section shows how to compute several risk measures based on the estimated mML OU model for a specified maturity T. By varying this maturity T, financial institutions can instead build evolution curves of these risk measures across time and hence better understand the future evolution of the market liquidity risk. This could be used to perform stress testing and scenario analysis as well as to model the impact of liquidity shocks on asset prices and portfolio values. It is also interesting to note that the main advantage of such risk measures built directly on the Amihud ratio (4.1) is that they do not require bid-ask spread/high-frequency data, which is the case for existing liquidity-adjusted VaRs, see *e.g* Weiß and Supper [47] or Weber et al. [48]. These finer measures of illiquidity risk indeed require for their calculation microstructure data on transactions and quotes that are unavailable in most markets around the world for long time periods of time. In contrast, the Amihud illiquidity measure used in this paper is calculated from daily data on returns and volume that are readily available over long periods of time for most markets. Therefore, while it is more coarse and less accurate, it is readily available for the study of the time series effects of liquidity and its associated risk. These risk measures on the Amihud ratio are also relatively simple to calculate and interpret, making it easy for practitioners to understand and to use.

5. Conclusion

This paper proposes a new procedure for modeling illiquidity in financial markets. We indeed 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 L^1 norm and of short and long-term asymptotics make this fractional Hawkes process better suited for the modeling of financial/economic events with long-term effects and dependence, compared with existing self-exciting models (exponential, power-law, ML or dML 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 its components. Therefore, this new mML fractional 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 with jumps following this mML fractional Hawkes process. This so-called mML OU model simplifies many computations and leads to new tools for managing and reducing the illiquidity risk, while reproducing the observed peaks of illiquidity in financial markets and their long-term effects. In particular, 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 of liquidity. Finally, after estimating our mML OU model via a POT procedure on FTSE 100's historical data, we demonstrate how this model results in a more risky distribution of the Amihud measure in comparison with the classical OU model, which is more aligned with empirical data.

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 or for the modeling of trade orders at high-frequency (and more generally for any financial/economic events with long memory). Similarly, the results of Section 3.2 could be used to study the stability of optimal control/stopping problems based on mML Hawkes processes, in particular for portfolio selection problems where the volatility/variance process would exhibit mML fractional Hawkes jumps with long memory.

Appendix A. Data section

	Stock return R_t	Volume V_t	Amihud measure A_t	Log-Amihud measure a_t
Mean	8.98×10^{-5}	8.31×10^{8}	1.55×10^{-15}	-34.66
Variance	1.16×10^{-4}	8.28×10^{16}	2.95×10^{-30}	1.44
Skewness	-0.74	1.85	5.85	-0.91
Kurtosis	9.31	7.62	86.25	1.37
Range	0.21	3.75×10^{9}	3.72×10^{-14}	10.13

Table 3 Data statistics for the FTSE100's daily returns and trading volume between 01/2009 and 01/2022 (Yahoo Finance). The daily Amihud and log-Amihud measures are then computed from these data via equation (4.1) with $p = 1$.

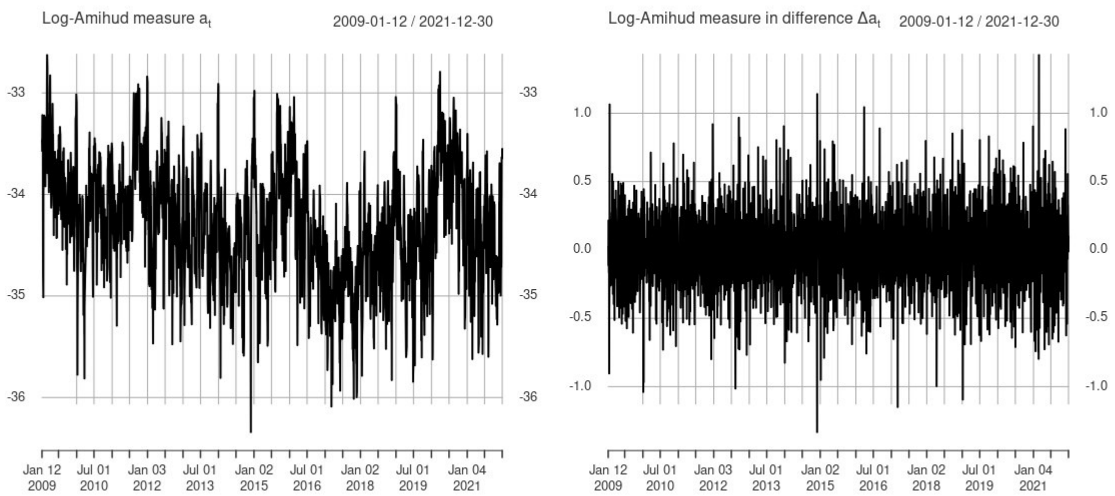

Fig. 5 Daily log-Amihud measure in level a_{t_j} (left) and in difference Δa_{t_j} (right) with $j = 1, \ldots, 3266$ for the FTSE 100 index over a period T of 13 years (01/2009 – 01/2022).

Appendix B. Proofs

Appendix B.1. Proof of Proposition 3.1.

(i) In order to derive the L^1 norm of the mML kernel (2.3), we first recall some probabilistic results derived from Kyprianou [49] (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 an α-stable process without positive jumps $\widetilde{U} := -U$. We can now define the supremum $\overline{U}_t := \sup\{U_s,\, s \leq t\} = -\inf\{\widetilde{U}_s,\, s \leq t\}$ and $\mathbf{e_1}$ an exponential random variable with parameter 1, independent from U. Therefore, we have from Kyprianou [49] that the Wiener-Hopf factorization leads for $\lambda \geq 0$ to

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

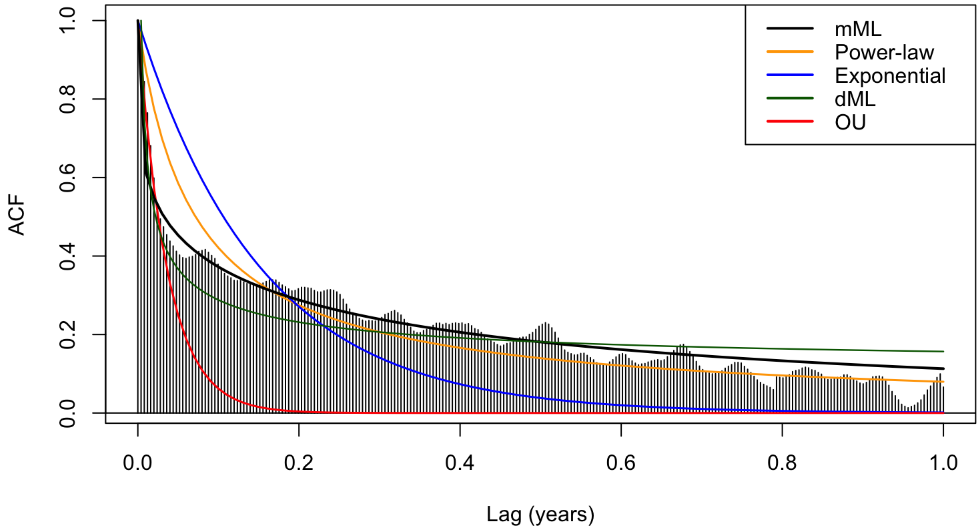

Fig. 6 Autocorrelogram (ACF) of the daily log-Amihud measure a_{t_j} $(j = 1, \ldots, 3266)$ for the FTSE 100 based on a 13-year period (2009–2022) with lag up to 1 year. In red, ACF of the classical OU model (without jumps) with the estimated parameters of Table 1. The four remaining curves depict the ACF of critical Hawkes process N with jump measure $\nu = \delta_1$ and $\lambda_0 = 1$: in black, ACF of N with the mML kernel (3.4) and $\gamma = 1.07, \eta = 1, \alpha = 0$; in blue, ACF of N with the exponential kernel (2.8) and $\alpha = 6.53, \eta = \alpha, \lambda_\infty = \lambda_0$; in orange, ACF of N with the power-law kernel (2.12) and $\beta = 16.72$, $\gamma = 0.06$, $\eta = \gamma$; in green, ACF of N with the dML kernel (2.15) and $\gamma = 0.56, \eta = 1$ (with $\alpha = 0$).

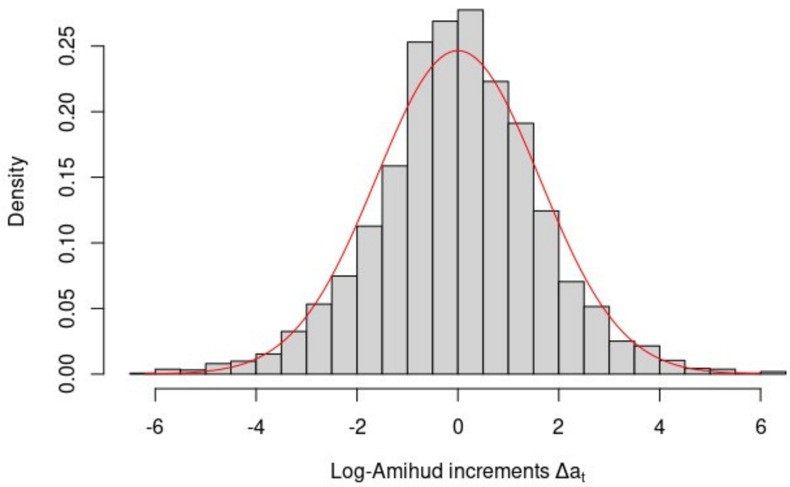

Fig. 7 Histogram of the daily log-Amihud increments Δa_{t_j} $(j = 1, \ldots, 3266)$ for the FTSE 100 index based on a 13-year period (01/2009 – 01/2022) with the Gaussian fit on these data superimposed in red.

and hence, integrating by parts, to

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

Moreover, it is well-known from Mainardi [27] that the function $E_\gamma(t^\gamma)$ has the following explicit Laplace transform

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

and integrating by parts, it comes for the modified Mittag-Leffler kernel (2.3) :

$$\mathcal{L}\left[f_\gamma(t)\right](\lambda) = \int_0^\infty e^{-\lambda t}\left(E_\gamma(t^\gamma) - t^{\gamma-1}E_{\gamma,\gamma}(t^\gamma)\right)\mathrm{d}t = \frac{\lambda^{\gamma-1}}{\lambda^\gamma - 1} - \frac{1}{\lambda^\gamma - 1} = \int_0^\infty e^{-\lambda t}\mathbb{P}\left(\overline{U}_{\mathbf{e_1}} \geq t\right)\mathrm{d}t\,, \quad \text{(B.2)}$$

when identifying γ and α. Hence, we finally find by inverting the Laplace transform (B.2) :

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

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

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

The expectation is obtained by differentiating the moment generating function (mgf) (B.1) with respect to λ and evaluating the derivative at $\lambda = 0$. This thus leads to $||\phi_\gamma||_1 = \eta||f_\gamma||_1 = \eta < \infty$.

(ii) Combined with the fact that ϕ_γ is also positive and causal, the three conditions of Definition 2.1 are verified and we finally have that the mML fractional Hawkes process ruled by the excitation function (2.3) is well-defined. □

Appendix B.2. Proof of Proposition 3.2.

The proof of Proposition 3.2 is direct from the Titchmarsh formula for the inversion of Laplace transforms, see Gross [50]. We first have that the Laplace transform of the kernel (2.3) is given by equation (B.2) :

$$\mathcal{L}_{f_\gamma}(\lambda) = \int_0^\infty e^{-\lambda t} f_\gamma(t)\mathrm{d}t = \frac{\lambda^{\gamma-1} - 1}{\lambda^\gamma - 1}\,. \quad \text{(B.3)}$$

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 $\theta \in (-\pi, \pi)$:

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

with $\alpha = \gamma - 1/2 \in [1/2, 3/2]$. Therefore, the assumptions of Hirschman and Widder [51] 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

$$f_\gamma(t) = \int_0^\infty e^{-\omega t}\mu(\omega)\,\mathrm{d}\omega\,. \quad \text{(B.4)}$$

Following Gross [50], we have from (B.3) and (B.4) that

$$\mathcal{L}_{f_\gamma}(\lambda) = \int_0^\infty \frac{\mu(\omega)\,\mathrm{d}\omega}{\omega + \lambda}\,. \quad \text{(B.5)}$$

This is a Stieltjes equation, which is inverted by

$$\begin{aligned}\mu(\omega) = \frac{1}{\pi}\mathrm{Im}\left[\mathcal{L}_{f_\gamma}(\omega e^{-i\pi})\right] &= \frac{\omega^{2\gamma-1}\sin(-\pi(\gamma-1) + \pi\gamma) - \omega^{\gamma-1}\sin(-\pi\gamma + \pi) + \omega^\gamma\sin(-\pi\gamma)}{\pi\left(\omega^{2\gamma} - 2\omega^\gamma\cos(\gamma\pi) + 1\right)} \\ &= \frac{-\sin(\gamma\pi)\,\omega^{\gamma-1}(1+\omega)}{\pi\left(\omega^{2\gamma} - 2\omega^\gamma\cos(\gamma\pi) + 1\right)}\,. \quad \text{(B.6)}\end{aligned}$$

□

Appendix B.3. Proof of Proposition 3.5

Using (3.10) in the first line and (3.9) to pass to the second line, along with assumptions $(i)-(iii)$ on the partition in Section 3.2, we find

$$\begin{aligned}\int_0^\infty f(\omega)\mu^{(n)}(\mathrm{d}\omega) = \sum_{k=1}^n m_k^{(n)} \int_0^\infty f(\omega)\,\delta_{b_k^{(n)}}(\mathrm{d}\omega) &= \sum_{k=1}^n m_k^{(n)} f\left(b_k^{(n)}\right) \\ &= \sum_{k=0}^{n-1} \int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} f\left(b_{k+1}^{(n)}\right)\mu(\mathrm{d}\omega) \\ &= \int_{\xi_0^{(n)}}^{\xi_n^{(n)}} f\left((t^{(n)}(\omega)\right)\mu(\mathrm{d}\omega)\,,\end{aligned}$$

where $t^{(n)}(\omega) = \sum_{k=0}^{n-1} b_{k+1}^{(n)}\,\mathbb{1}_{\left[\xi_k^{(n)},\xi_{k+1}^{(n)}\right]}(\omega)$. Given that f is continuous and that $\lim_{n\to\infty} t^{(n)}(\omega) = \omega$ under assumptions $(i)-(iii)$, passing to the limit leads to the announced result. □

Appendix B.4. Proof of Proposition 3.6.

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 $f\left(t, T, X_t^{(n)}, \left(Y_t^{(k)}\right)_{k=1,\ldots,n}\right)$. Moreover, the infinitesimal generator of $X^{(n)}$ is given by

$$\begin{aligned}\mathcal{L}f(x) &= \frac{\partial f}{\partial \lambda^{(n)}}(x) \sum_{k=1}^n m_k^{(n)} \left(\alpha\left(\lambda_\infty - \lambda^{(n)}\right) - b_k^{(n)} Y^{(k)}\right) + \sum_{k=1}^n \frac{\partial f}{\partial Y^{(k)}}(x)\left(\alpha\left(\lambda_\infty - \lambda^{(n)}\right) - b_k^{(n)} Y^{(k)}\right) \\ &+ \lambda^{(n)} \int_0^\infty \left(f\left(\lambda^{(n)} + \eta\, o \sum_{k=1}^n m_k^{(n)}, L^{(n)} + o, N^{(n)} + 1, \left(Y^{(k)} + \eta o\right)_{k=1,\ldots,n}\right) - f(x)\right)\nu(\mathrm{d}o)\,, \quad \text{(B.7)}\end{aligned}$$

where $X_t^{(n)} = x = (\lambda_t^{(n)}, (L_t^{(n)}, N_t^{(n)}))^\top$. Since f must again satisfy the PIDE (2.11) 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

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

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 $c_k(t,T) := c_k(u',t,T)$ is a function that depends again on time but also on the position $k = 1,\ldots,n$, within the partition $\mathcal{E}^{(n)}$, and which satisfies the boundary condition $c_k(T,T) = 0$, $\forall k = 1,\ldots,n$. Differentiating f with respect to t, $\lambda^{(n)}$ and $Y^{(k)}$ leads to

$$\begin{aligned}\frac{\partial f}{\partial t} &= f\left(\frac{\partial a(t,T)}{\partial t} + \frac{\partial b(t,T)}{\partial t}\lambda^{(n)} + \sum_{k=1}^n \frac{\partial c_k(t,T)}{\partial t}\, Y^{(k)}\right), \\ \frac{\partial f}{\partial \lambda^{(n)}} &= f b(t,T)\,, \\ \frac{\partial f}{\partial Y^{(k)}} &= f c_k(t,T), \quad \forall k = 1,\ldots,n,\end{aligned}$$

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

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

Injecting these equations in (B.7) and in the PIDE (2.11), we have

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

From which we find the following system of PDE

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

where $\theta(\cdot)$ is again the transform of the jump distribution.

The previous system of equation finally gives with $a(T,T) = 0$ and $c_k(T,T) = 0$:

$$
\begin{aligned}
a(t,T) &= \alpha\,\lambda_\infty \int_t^T \left(\sum_{k=1}^{n} m_k^{(n)}\,b(v,T) + \sum_{k=1}^{n} c_k(v,T)\right)\mathrm{d}v\,, \\
\frac{\partial b(t,T)}{\partial t} &= \alpha\left(\sum_{k=1}^{n} m_k^{(n)} b(t,T) + \sum_{k=1}^{n} c_k(t,T)\right) - \theta\left(b(t,T)\,\eta\sum_{k=1}^{n} m_k^{(n)} + u'\cdot(1,0)^\top + \sum_{k=1}^{n} c_k(t,T)\,\eta\right) e^{u'\cdot(0,1)^\top} + 1\,, \\
c_k(t,T) &= -b_k^{(n)}\,m_k^{(n)} \int_t^T e^{-b_k^{(n)}(v-t)}\,b(v,T)\,\mathrm{d}v\,, \quad \forall k = 1,\dots n,
\end{aligned}
$$

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

$$
\begin{aligned}
a(t,T) &= \alpha\,\lambda_\infty \sum_{k=1}^{n} m_k^{(n)}\left(\int_t^T b(v,T)\,\mathrm{d}v - b_k^{(n)}\int_t^T\int_v^T e^{-b_k^{(n)}(s-v)}\,b(s,T)\,\mathrm{d}s\,\mathrm{d}v\right) \\
&= \alpha\,\lambda_\infty \sum_{k=1}^{n} m_k^{(n)}\left(\int_t^T b(v,T)\mathrm{d}v - b_k^{(n)}\int_t^T \frac{1}{b_k^{(n)}}\left(1 - e^{-b_k^{(n)}(s-t)}\right) b(s,T)\,\mathrm{d}s\right) \\
&= \alpha\,\lambda_\infty \sum_{k=1}^{n} m_k^{(n)}\left(\int_t^T e^{-b_k^{(n)}(v-t)} b(v,T)\,\mathrm{d}v\right).
\end{aligned}
$$

□

Appendix B.5. Proof of Corollary 3.7.

Part 1. If a partition $\mathcal{E}^{(n)}$ satisfies assumptions (i)-(iii), we first show that $\left(\lambda_t^{(n)}, N_t^{(n)}, L_t^{(n)}, \left(Y_t^{(k)}\right)_{k=1,\dots,n}\right)$ converges to $\left(\lambda_t, N_t, L_t, \left(Y_t^\omega\right)_{\omega\in\mathbb{R}^+}\right)$ almost surely when $n\to\infty$ for $t \geq 0$. The almost sure convergence $\lim_{n\to\infty}\lambda_t^{(n)} = \lambda_t$ has been shown in p.150 (3.12) since we have $\lambda_t^{(n)} = \lambda_0 + \int_0^\infty Y_t^\omega \mu^{(n)}(\mathrm{d}\omega)$ and since the function $\omega\in\mathbb{R}^+ \mapsto Y_t^\omega(w')$ is integrable with respect to the measure μ for fixed $w'\in\Omega$ as a sum of negative exponential functions. It follows that $N_t^{(n)} = \int_0^t \lambda_u^{(n)}\mathrm{d}u$ converges almost surely to N_t and $L_t^{(n)}$ to L_t, when $n\to\infty$. Finally, from (3.7), we have that $Y_t^{(k)} := Y_t^{b_k^{(n)}} = \int_0^t e^{-b_k^{(n)}(t-u)}\left(\alpha(\lambda_\infty - \lambda_u^{(n)})\mathrm{d}u + \eta\mathrm{d}L_u^{(n)}\right)$ for

$k \in \{1, \ldots, n\}$. If the partition $\mathcal{E}^n$ satisfies assumption (*i*)-(*iii*), the barycenter $b_k^{(n)}$ tends to be an element of $\mathbb{R}_+$, when $n \to \infty$. Hence, the sequence $\big(Y_t^{(k)}\big)_{k=1,\ldots,n}$ converges almost surely to $(Y_t^{\omega})_{\omega \in \mathbb{R}_+}$.

Part 2. The dominated convergence theorem then provides

$$\mathbb{E}_t\left[e^{u_\lambda \lambda_T + u_L L_T + u_N N_T}\right] = \lim_{n\to\infty} \mathbb{E}_t\left[e^{u_\lambda \lambda_T^{(n)} + u_L L_T^{(n)} + u_N N_T^{(n)}}\right].$$

Moreover, for deriving the coefficient function $b(t,T)$, we first have from equation (3.2) that

$$\lim_{n\to\infty} \sum_{k=1}^{n} m_k^{(n)} = \int_0^\infty \mu(\mathrm{d}\omega) = f_\gamma(0+) = 1\,. \tag{B.8}$$

We note that this quantity diverges toward infinity for the power-law and dML kernels, preventing from using this discretization scheme 3.2 in such cases. As in Part 1 of the proof, we then have almost surely in the limit $n \to \infty$:

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

Then, from equation (3.2), we have

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

and finally from Leibniz integral rule to

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

□

Appendix B.6. Proof of Corollary 3.8.

By differentiating the conditional transform (3.14) with respect to u_λ, u_L, u_N and evaluating the sum of derivatives at $u = (0,0,0)$, we find the following conditional expectation

$$\begin{aligned}\mathbb{E}_t\left[w\cdot X_T\right] = \alpha\lambda_\infty \int_t^T f_\gamma(v-t)\,b'(0,v,T)\,\mathrm{d}v + b'(0,t,T)\lambda_t + w'\cdot J_t&\\ - \int_0^\infty \left(\int_t^T e^{-\omega(v-t)} b'(0,v,T)\,\mathrm{d}v\right)\omega\, Y_t^{\omega}\mu(\mathrm{d}\omega)\,,&\end{aligned} \tag{B.9}$$

where $b'(0,t,T)$ is the sum of derivatives of the coefficient function $b(u',t,T)$ in (3.15) with respect to (u_L, u_N) evaluated at $(0,0)$, and which satisfies

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

with boundary condition $b'(0,T,T) = w_\lambda$ and where $\theta'(\xi) = \int o\, e^{\xi o} \mathrm{d}\nu(o)$ is the derivative of the jump transform. Since $u = (0,0,0)$, we can choose $b(0,t,T) = 0$, $a(0,t,T) = 0$ and $c_k(0,t,T) = 0$ thanks to equation (3.13) and hence, the previous equation simplifies to

$$\frac{\partial b'(0,t,T)}{\partial t} = -\alpha \frac{\partial}{\partial t} \int_t^T f_\gamma(v-t)\, b'(0,v,T)\, \mathrm{d}v - \theta(0)\, w' \cdot (0,1)^\top - \theta'(0) \left(w' \cdot (1,0)^\top - \eta \frac{\partial}{\partial t} \int_t^T f_\gamma(v-t)\, b'(0,v,T)\, \mathrm{d}v \right).$$

We then have

$$\begin{aligned} \frac{\partial b'(0,t,T)}{\partial t} &= -\alpha \frac{\partial}{\partial t} \int_t^T f_\gamma(v-t) b'(0,v,T) \mathrm{d}v + \chi\eta \frac{\partial}{\partial t} \int_t^T f_\gamma(v-t) b'(0,v,T) \mathrm{d}v - w' \cdot (\chi,1)^\top \\ &= (\chi\eta - \alpha) \frac{\partial}{\partial t} \int_t^T f_\gamma(v-t)\, b'(0,v,T)\, \mathrm{d}v - w' \cdot (\chi,1)^\top, \end{aligned} \tag{B.10}$$

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

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

□

Appendix B.7. A multivariate extension

We can extend the univariate fractional Hawkes setting of Section 3 by considering a multivariate mML fractional Hawkes process where the intensity $\lambda^{(i)}$ of each component process $i = 1, \ldots, d$ may also depend on several of the other Hawkes components. For completeness, we also let each component's intensity $\lambda^{(i)}$ be influenced by exogenous risk factors in the form of a diffusive stochastic process. These intensities $\lambda^{(i)}$ can therefore replicate the diffusive fluctuations of the observed time series of market prices and market measures. We thus consider the $\mathbb{N}^d$-valued counting process N and the $\mathbb{R}_+^d$-valued point process L where the conditional intensity $\lambda_t^{(i)}$ 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 following affine structure,

$$\lambda_t^{(i)} = \lambda_0^{(i)} + \int_0^t f_{\gamma^{(i)}}(t-u) \left(\alpha^{(i)} (\lambda_\infty^{(i)} - \lambda_u^{(i)})\, \mathrm{d}u + \eta^{(i)} \cdot \mathrm{d}L_u^\top + \sigma^{(i)} \sqrt{\lambda_u^{(i)}}\, \mathrm{d}W_u^{(i)} \right) \tag{B.11}$$

where[6] $\eta^{(i)} = (\eta^{(i,1)}, \eta^{(i,2)} \ldots, \eta^{(i,d)}) \in \mathbb{R}_+^d$ and where $W^{(i)}$ is a one-dimensional standard Brownian motion independent from the jump processes $L^{(i)}$, $\forall i = 1, \ldots, d$, and from the other Brownian motions $W^{(j)}, \forall\, j \neq i$. The jump size $O^{(i)}$ of the i^{th} component $L^{(i)}$ 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 $L^{(i)}$ 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 (B.11) also allows the modeling of (il)liquidity peaks through the Amihud measure, as we will explore in the Section 4, and to simultaneous study the illiquidity process of multiple stocks and indices as well as their contagion. Finally, this multivariate extension is required at

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

high-frequency in order to model both the up and down mid-changes in prices, see Bacry and Muzy [1] and Bacry et al. [2]. As in Section 3.2, we can rewrite the conditional intensity (B.11) as

$$\lambda_t^{(i)} = \lambda_0^{(i)} + \int_0^\infty Y_t^{\omega,(i)} \mu^{(i)}(\mathrm{d}\omega)\,, \tag{B.12}$$

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

$$\mathrm{d}Y_t^{\omega,(i)} = \left(-\omega Y_t^{\omega,(i)} + \alpha^{(i)}\left(\lambda_\infty^{(i)} - \lambda_t^{(i)}\right)\right)\mathrm{d}t + \eta^{(i)} \cdot \mathrm{d}L_t^\top + \sigma^{(i)}\sqrt{\lambda_t^{(i)}}\mathrm{d}W_t^{(i)}, \tag{B.13}$$

with $Y_0^{\omega,(i)} = 0$. From the affine structure of equation (B.11) and from the representation (B.12) of each conditional intensity in terms of infinitely many mean-reverting factors $Y_t^{\omega,(i)}$, we can apply directly the same reasoning as in the univariate case (3.14) so as to obtain the conditional transform of the process $Z = (\lambda, J)^\top = (\lambda, (L, N))^\top$. The following Proposition B.7 gives the exact expression of this conditional transform $\mathbb{E}_t\left[e^{u\cdot Z_T}\right]$ in the multivariate setting and is obtained as for Proposition 3.6 from the discretization scheme 3.2 based on the infinitesimal generator of Z. The conditional transform of the process $Z = (\lambda, J)^\top = (\lambda, (L, N))^\top$ is then defined on a state space $D \subseteq \mathbb{R}_+^d \times (\mathbb{R}_+^d \times \mathbb{N}^d)$ with $u_\lambda, u_L, u_N \in \mathbb{C}^d$ and $u = (u_\lambda, u_L, u_N) \in \mathbb{C}^{3d}$ and is given as follows.

Proposition B.7. *The transform of $Z_T = (\lambda_T, (L_T, N_T))^\top$ conditionally to $\mathcal{F}_t$ at time $t \le T$ with $u = (u_\lambda, u_L, u_N) \in \mathbb{C}^{3d}$ is given by*

$$\begin{aligned}\mathbb{E}_t\left[e^{u\cdot Z_T}\right] &= \mathbb{E}_t\left[e^{u_\lambda\cdot\lambda_T^\top + u_L\cdot L_T^\top + u_N\cdot N_T^\top}\right] = \exp\Bigg[\beta(t,T)\cdot\lambda_t^\top + u_L\cdot L_t^\top + u_N\cdot N_t^\top \\ &+ \sum_{i=1}^d\left(\alpha^{(i)}\lambda_\infty^{(i)}\int_t^T f_{\gamma^{(i)}}(v-t)\beta^{(i)}(v,T)\mathrm{d}v - \int_0^\infty\left(\int_t^T e^{-\omega(v-t)}\beta^{(i)}(v,T)\mathrm{d}v\right)\omega Y_t^{\omega,(i)}\mu^{(i)}(\mathrm{d}\omega)\right)\Bigg]\end{aligned}$$

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

$$\begin{aligned}\frac{\partial\beta^{(i)}(t,T)}{\partial t} =& -\alpha^{(i)}\frac{\partial}{\partial t}\int_t^T f_{\gamma^{(i)}}(v-t)\beta^{(i)}(v,T)\mathrm{d}v - \frac{1}{2}\sigma^{(i)^2}\left(\frac{\partial}{\partial t}\int_t^T f_{\gamma^{(i)}}(v-t)\beta^{(i)}(v,T)\mathrm{d}v\right)^2 \\ & - \theta^{(i)}\left(u_{L,i} - \sum_{j=1}^d \eta^{(j,i)}\frac{\partial}{\partial t}\int_t^T f_{\gamma^{(j)}}(v-t)\,\beta^{(j)}(v,T)\,\mathrm{d}v\right)e^{u_{N,i}} + 1\,,\end{aligned} \tag{B.14}$$

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

Proof. As for Proposition 3.6, we start by studying the discretized version of the conditional intensity (B.12) based on a partition $\mathcal{E}^n$ of size n, and hence obtain the following dynamics

$$\mathrm{d}Y_t^{(k,i)} = \left(-b_k^{(n)}Y_t^{(k,i)} + \alpha^{(i)}\left(\lambda_\infty^{(i)} - \lambda_t^{(n,i)}\right)\right)\mathrm{d}t + \sum_{j=1}^d \eta^{(i,j)}O^{(j)}\mathrm{d}N_t^{(n,j)} + \sigma^{(i)}\sqrt{\lambda_t^{(n,i)}}\mathrm{d}W_t^{(i)}\,, \tag{B.15}$$

where $Y^{(k,i)} := Y^{b_k^{(n)},(i)}$ for $i = 1, \ldots, d$, $k = 1, \ldots, n$ using (B.13), and

$$\mathrm{d}\lambda_t^{(n,i)} = \sum_{k=1}^n m_k^{(n,i)}\mathrm{d}Y_t^{(k,i)}\,, \tag{B.16}$$

where $m_k^{(n,i)} = \int_{\xi_{k-1}^{(n)}}^{\xi_k^{(n)}} \mu^{(i)}(\mathrm{d}\omega)$, where $N^{(n,i)}$ is a counting process of intensity $\lambda^{(n,i)}$, and where $L_t^{(n,i)} = \sum_{j=1}^{N_t^{(n)}} O_j$. The infinitesimal generator of $Z^{(n)} = \left(\lambda^{(n)}, J^{(n)}\right)^\top$ is given by

$$\begin{aligned}\mathcal{L}f(z) = &\sum_{i=1}^{d}\Bigg\{\frac{\partial f}{\partial \lambda^{(n,i)}}(z)\sum_{k=1}^{n} m_k^{(n,i)}\left(\alpha^{(i)}\left(\lambda_\infty^{(i)}-\lambda^{(n,i)}\right)-b_k^{(n)}Y^{(k,i)}\right)+\sum_{k=1}^{n}\frac{\partial f}{\partial Y^{(k,i)}}(z)\Big[\alpha^{(i)}\left(\lambda_\infty^{(i)}-\lambda^{(n,i)}\right)\\
&-b_k^{(n)}Y^{(k,i)}\Big]+\frac{1}{2}\lambda^{(n,i)}\sigma^{(i)^2}\left(\sum_{k=1}^{n} m_k^{(n,i)}\right)^2\frac{\partial^2 f}{\partial \lambda^{(n,i)^2}}(z)+\frac{1}{2}\lambda^{(n,i)}\sigma^{(i)^2}\sum_{k_1=1,k_2=1}^{n}\frac{\partial^2 f}{\partial Y^{(k_1,i)}\partial Y^{(k_2,i)}}(z)\\
&+\lambda^{(n,i)}\int_0^\infty\Bigg[f\Bigg(\left(\lambda^{(n,j)}+\eta^{(j,i)}o\sum_{k=1}^{n} m_k^{(n,j)}\right)_{j=1,\ldots,d},\left(Y^{(k,j)}+\eta^{(j,i)}o\right)_{\substack{k=1,\ldots,n\\ j=1,\ldots,d}},L^{(n,i)}+o,N^{(n,i)}+1\Bigg)\\
&-f(z)\Bigg]\nu^{(i)}(\mathrm{d}o)+\lambda^{(n,i)}\sigma^{(i)^2}\left(\sum_{k=1}^{n} m_k^{(n,i)}\right)\sum_{k=1}^{n}\frac{\partial^2 f}{\partial \lambda^{(n,i)}\partial Y^{(k,i)}}(z)\Bigg\},\end{aligned}$$

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

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

Since f must satisfy the PIDE (2.11) 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 :

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

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

$$\begin{aligned}0 = &\frac{\partial\alpha(t,T)}{\partial t}+\sum_{i=1}^{d}\Bigg\{\frac{\partial\beta^{(i)}(t,T)}{\partial t}\lambda^{(n,i)}+\sum_{k=1}^{n}\frac{\partial c_k^{(i)}(t,T)}{\partial t}Y^{(k,i)}\\
&+\beta^{(i)}(t,T)\sum_{k=1}^{n} m_k^{(n,i)}\left(\alpha^{(i)}\left(\lambda_\infty^{(i)}-\lambda^{(n,i)}\right)-b_k^{(n)}Y^{(k,i)}\right)+\sum_{k=1}^{n}c_k^{(i)}(t,T)\left(\alpha^{(i)}\left(\lambda_\infty^{(i)}-\lambda^{(n,i)}\right)-b_k^{(n)}Y^{(k,i)}\right)\\
&+\lambda^{(n,i)}\left[\left(\int\exp\left(\sum_{j=1}^{d}\beta^{(j)}(t,T)\eta^{(j,i)}o\sum_{k=1}^{n}m_k^{(n,j)}+u_{L,i}\,o+\sum_{j=1}^{d}\sum_{k=1}^{n}c_k^{(j)}(t,T)\eta^{(j,i)}o\right)\nu^{(i)}(\mathrm{d}o)\right)e^{u_{N,i}}-1\right]\\
&+\frac{1}{2}\lambda^{(n,i)}\sigma^{(i)^2}\left(\sum_{k=1}^{n}\left(m_k^{(n,i)}\beta^{(i)}(t,T)+c_k^{(i)}(t,T)\right)\right)^2\Bigg\}.\end{aligned}$$

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

$$\begin{aligned}\frac{\partial\alpha(t,T)}{\partial t} &= -\sum_{i=1}^{d}\left(\alpha^{(i)}\lambda_\infty^{(i)}\sum_{k=1}^{n}\left(\beta^{(i)}(t,T)\,m_k^{(n,i)}+c_k^{(i)}(t,T)\right)\right),\\
\frac{\partial c_k^{(i)}(t,T)}{\partial t} &= \beta^{(i)}(t,T)\,m_k^{(n,i)}\,b_k^{(n)}+c_k^{(i)}(t,T)\,b_k^{(n)}\,,\quad \forall i=1,\ldots,d\,,\ \forall k=1,\ldots,n,\\
\frac{\partial\beta^{(i)}(t,T)}{\partial t} &= \alpha^{(i)}\sum_{k=1}^{n}\left(m_k^{(n,i)}\beta^{(i)}(t,T)+c_k^{(i)}(t,T)\right)-\frac{1}{2}\sigma^{(i)^2}\left(\sum_{k=1}^{n}\left(m_k^{(n,i)}\beta^{(i)}(t,T)+c_k^{(i)}(t,T)\right)\right)^2\\
&\quad-\theta^{(i)}\left(\sum_{j=1}^{d}\eta^{(j,i)}\sum_{k=1}^{n}\left(\beta^{(j)}(t,T)\,m_k^{(n,j)}+c_k^{(j)}(t,T)\right)+u_{L,i}\right)e^{u_{N,i}}+1\,,\quad \forall i=1,\ldots,d.\end{aligned}$$

This finally leads for the coefficients $\beta^{(i)}(\cdot)$ and $c_k^{(i)}(\cdot)$ to

$$c_k^{(i)}(t,T) = -b_k^{(n)}\, m_k^{(n,i)} \int_t^T e^{-b_k^{(n)}(v-t)}\, \beta^{(i)}(v,T)\,\mathrm{d}v\,, \quad \forall i = 1,\dots,d,\ \forall k = 1,\dots,n,$$

and

$$\begin{aligned}\alpha(t,T) &= \sum_{i=1}^d \left(\alpha^{(i)}\lambda_\infty^{(i)} \sum_{k=1}^n m_k^{(n,i)} \left(\int_t^T \beta^{(i)}(v,T)\mathrm{d}v - b_k^{(n)} \int_t^T \int_v^T e^{-b_k^{(n)}(s-v)}\beta^{(i)}(s,T)\,\mathrm{d}s\mathrm{d}v\right)\right)\\ &= \sum_{i=1}^d \left(\alpha^{(i)}\lambda_\infty^{(i)} \sum_{k=1}^n m_k^{(n,i)} \left(\int_t^T e^{-b_k^{(n)}(v-t)}\beta^{(i)}(v,T)\,\mathrm{d}v\right)\right).\end{aligned}$$

From Proposition 3.5 and Corollary 3.7 along with dominated convergence, passing to the limit $n \to \infty$ leads to the announced result. In particular, we have almost surely that

$$\lim_{n\to\infty} \sum_{k=1}^n \left(m_k^{(n,i)}\beta^{(i)}(t,T) + c_k^{(i)}(t,T)\right) = -\frac{\partial}{\partial t}\int_t^T f_{\gamma^{(i)}}(v-t)\beta^{(i)}(v,T)\mathrm{d}v\,, \quad \forall i \in 1,\dots,d.$$

□

Declaration

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 financial or non-financial interests to disclose.

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

References

[1] Bacry, E., Muzy, J.-F.: Hawkes model for price and trades high-frequency dynamics. Quantitative Finance **14**(7), 1147–1166 (2014)

[2] Bacry, E., Mastromatteo, I., Muzy, J.-F.: Hawkes processes in finance. Market Microstructure and Liquidity **1**(01), 1550005 (2015)

[3] Hardiman, S.J., Bercot, N., Bouchaud, J.-P.: Critical reflexivity in financial markets: a hawkes process analysis. The European Physical Journal B **86**(10), 1–9 (2013)

[4] Cayé, T., Muhle-Karbe, J.: Liquidation with self-exciting price impact. Mathematics and Financial Economics **10**, 15–28 (2016)

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

[6] Errais, E., Giesecke, K., Goldberg, L.R.: Affine point processes and portfolio credit risk. SIAM Journal on Financial Mathematics **1**(1), 642–665 (2010)

[7] Aït-Sahalia, Y., Cacho-Diaz, J., Laeven, R.J.: Modeling financial contagion using mutually exciting jump processes. Journal of Financial Economics **117**(3), 585–606 (2015)

[8] Aït-Sahalia, Y., Hurd, T.R.: Portfolio choice in markets with contagion. Journal of Financial Econometrics **14**(1), 1–28 (2016)

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

[10] Hainaut, D., Goutte, S.: A switching microstructure model for stock prices. Mathematics and Financial Economics **13**, 459–490 (2019)

[11] Hainaut, D.: Continuous Time Processes for Finance. Switching, Self-exciting, Fractional and Other Recent Dynamics. Bocconi & Springer Series, London (2022)

[12] Bacry, E., Dayri, K., Muzy, J.-F.: Non-parametric kernel estimation for symmetric hawkes processes. application to high frequency financial data. The European Physical Journal B **85**(5), 1–12 (2012)

[13] Hainaut, D.: A model for interest rates with clustering effects. Quantitative Finance **16**(8), 1203–1218 (2016)

[14] Hainaut, D.: Fractional hawkes processes. Physica A: Statistical Mechanics and its Applications **549**, 124330 (2020)

[15] Njike Leunga, C.G.: Long memory self-exciting jump diffusion for asset prices modeling. Technical report, LIDAM Discussion Paper ISBA ; 2022/03 (2022)

[16] Chen, J., Hawkes, A., Scalas, E.: A fractional hawkes process. In: Nonlocal and Fractional Operators, pp. 121–131. Springer, New-York (2021)

[17] Bäuerle, N., Desmettre, S.: Portfolio optimization in fractional and rough heston models. SIAM Journal on Financial Mathematics **11**(1), 240–273 (2020)

[18] Goyenko, R.Y., Holden, C.W., Trzcinka, C.A.: Do liquidity measures measure liquidity? Journal of financial Economics **92**(2), 153–181 (2009)

[19] Amihud, Y.: Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets **5**(1), 31–56 (2002)

[20] Amihud, Y.: The pricing of illiquidity as a characteristic and as risk. Multinational Finance Journal **19**(3), 149–168 (2015)

[21] Brennan, M., Huh, S.-W., Subrahmanyam, A.: An analysis of the amihud illiquidity premium. The Review of Asset Pricing Studies **3**(1), 133–176 (2013)

[22] Hafner, C.M., Linton, O.B., Wang, L.: Dynamic autoregressive liquidity (darliq). Journal of Business & Economic Statistics **42**(2), 774–785 (2024)

[23] Hawkes, A.G.: Spectra of some self-exciting and mutually exciting point processes. Biometrika **58**(1), 83–90 (1971)

[24] Njike Leunga, C.G., Hainaut, D.: Affine heston model style with self-exciting jumps and long memory. Annals of Finance, 1–43 (2024)

[25] Bateman, H.: Higher Transcendental Functions [volumes I-iii] vol. 1. McGraw-Hill Book Company, New-York (1953)

[26] Jaisson, T.: Market impact as anticipation of the order flow imbalance. Quantitative Finance **15**(7), 1123–1135 (2015)

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

[28] Lin, G.D.: On the mittag–leffler distributions. Journal of Statistical Planning and Inference **74**(1), 1–9 (1998)

[29] Abi Jaber, E., Larsson, M., Pulido, S.: Affine volterra processes. The Annals of Applied Probability **29**(5), 3155–3200 (2019)

[30] Abi Jaber, E.: Weak existence and uniqueness for affine stochastic Volterra equations with L1-kernels. Bernoulli **27**(3), 1583–1615 (2021)

[31] Bondi, A., Livieri, G., Pulido, S.: Affine Volterra processes with jumps. Stochastic Processes and their Applications **168**, 104264 (2024)

[32] Bondi, A., Pulido, S., Scotti, S.: The rough Hawkes Heston stochastic volatility model. arXiv preprint arXiv:2210.12393 (2022)

[33] Horst, U., Xu, W.: Functional limit theorems for hawkes processes. arXiv preprint arXiv:2401.11495 (2024)

[34] Cao, C., Chen, Y., Liang, B., Lo, A.W.: Can hedge funds time market liquidity? Journal of Financial Economics **109**(2), 493–516 (2013)

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

[36] Pipiras, V., Taqqu, M.S.: Long-range Dependence and Self-similarity vol. 45. Cambridge university press, Cambridge (2017)

[37] Andersen, T.G., Bollerslev, T., Diebold, F.X., Ebens, H.: The distribution of realized stock return volatility. Journal of Financial Economics **61**(1), 43–76 (2001)

[38] Gatheral, J., Jaisson, T., Rosenbaum, M.: Volatility is rough. Quantitative finance **18**(6), 933–949 (2018)

[39] Lobato, I.N., Velasco, C.: Long memory in stock-market trading volume. Journal of Business & Economic Statistics **18**(4), 410–427 (2000)

[40] Hainaut, D., Moraux, F.: Hedging of options in the presence of jump clustering. Journal of Computational Finance, Forthcoming (2018)

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

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

[43] Dupret, J.-L., Hainaut, D.: Portfolio insurance under rough volatility and volterra processes. International Journal of Theoretical and Applied Finance **24**(06n07), 2150036 (2021)

[44] Albanese, C., Jackson, K., Wiberg, P.: A new Fourier transform algorithm for value-at-risk. Quantitative Finance **4**(3), 328 (2004)

[45] Ogata, Y.: On lewis' simulation method for point processes. IEEE transactions on information theory **27**(1), 23–31 (1981)

[46] Floudas, C.A., Pardalos, P.M.: Encyclopedia of Optimization, 2nd edn. Springer, New-York (2008)

[47] Weiß, G.N., Supper, H.: Forecasting liquidity-adjusted intraday value-at-risk with vine copulas. Journal of Banking & Finance **37**(9), 3334–3350 (2013)

[48] Weber, S., Anderson, W., Hamm, A.-M., Knispel, T., Liese, M., Salfeld, T.: Liquidity-adjusted risk measures. Mathematics and Financial Economics **7**, 69–91 (2013)

[49] Kyprianou, A.E.: Introductory Lectures on Fluctuations of Lévy Processes with Applications. Springer, New-York (2006)

[50] Gross, B.: 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)

[51] Hirschman, I.I., Widder, D.V.: The Convolution Transform. Courier Corporation, Chelmsford, Massachusetts, US (2012)