Multivariate Quadratic Hawkes Processes -Part I: Theoretical Analysis

Cecilia Aubrun, Michael Benzaquen, Jean-Philippe Bouchaud

Multivariate Quadratic Hawkes Processes – Part I: Theoretical Analysis

Cécilia Aubrun,[1, 2] Michael Benzaquen,[1, 2, 3, *] and Jean-Philippe Bouchaud[1, 3, 4]

[1]*Chair of Econophysics and Complex Systems, École polytechnique, 91128 Palaiseau Cedex, France*
[2]*LadHyX UMR CNRS 7646, École polytechnique, 91128 Palaiseau Cedex, France*
[3]*Capital Fund Management, 23 Rue de l'Université, 75007 Paris, France*
[4]*Académie des Sciences, 23 Quai de Conti, 75006 Paris, France*
(Dated: June 28, 2022)

Quadratic Hawkes (QHawkes) processes have proved effective at reproducing the statistics of price changes, capturing many of the stylised facts of financial markets. Motivated by the recently reported strong occurrence of endogenous co-jumps (simultaneous price jumps of several assets) we extend QHawkes to a multivariate framework (MQHawkes), that is considering several financial assets and their interactions. Assuming that quadratic kernels write as the sum of a time-diagonal component and a rank one (trend) contribution, we investigate endogeneity ratios and the resulting stationarity conditions. We then derive the so-called Yule-Walker equations relating covariances and feedback kernels, which are essential to calibrate the MQHawkes process on empirical data. Finally, we investigate the volatility distribution of the process and find that, as in the univariate case, it exhibits power-law behavior, with an exponent that can be exactly computed in some limiting cases.

Keywords: Multivariate QHawkes

INTRODUCTION

Modelling the volatility of financial assets is a significant challenge for academics, market participants and regulators alike. In fact, models describing the statistics of price changes are widely used for, e.g., risk control and derivative pricing. When not in line with the behaviour of real markets, these models can lead to disappointing outcomes, or even major mishaps. Thus, it is important to account for the main stylized facts observed in real price time series, in particular: fat tails of the returns distribution, multi-timescale volatility clustering, price-volatility correlations and time reversal asymmetry.

Among the most widely used models, one can quote stochastic volatility models and GARCH models. However, these models often capture only poorly the above stylized facts. Indeed, the most popular stochastic volatility models lead to thin tails in the returns distribution, and time reversal symmetry. Whereas GARCH models do present time reversal asymmetry, such asymmetry is stronger than that observed in real markets. Furthermore, volatility clustering is governed by a single time scale. Some of these problems are elegantly solved by the family of rough volatility models [1], in particular the multifractal random walk which is but a special member of that family [2, 3].

Hawkes processes, on the other hand, provide an interesting alternative to more traditional models, while clearly highlighting the feedback loop at the origin of the stylized facts characterizing financial prices. However, Jaisson and Rosenbaum [4] have shown that the continuous time limit of a (near-critical) Hawkes process is a fractional CIR Heston process, which is characterized by tails that are too thin to reproduce empirical returns distribution. Furthermore, such a process is time reversal invariant.

To alleviate these problems, Donier, Blanc, Bouchaud in [5] introduced the family of "Quadratic" Hawkes (QHawkes) processes, which encodes a feedback between past price *trends* and future volatility, a clear empirical effect first discovered by Zumbach [6]. The quadratic feedback allows one to overcome the limitations of standard (linear) Hawkes processes. QHawkes processes naturally generate fat tail distributions and time reversal asymmetry even in the continuous time limit.

In this paper, we generalize the monovariate QHawkes process of Ref. [5] to the multivariate case, allowing one to capture how the trend on one asset can impact the volatility of another asset. In Section I, we define the multivariate QHawkes model in full generality. Section II establishes generalized Yule-Walker equations that relate 2- and 3-point correlation functions to the various feedback kernels appearing in the definition of the model. These relations are the starting point for the calibration of the model to empirical data, a topic we will explore in a follow-up, companion paper ([7]). Section III investigates the volatility distribution generated by the model and establishes the power-law nature of the tails, with an exponent that we analytically compute in several special cases. In Section IV we conclude.

Beside a notations table is provided after the bibliography to ease the reading (See Appendix 1).

---

* michael.benzaquen@polytechnique.edu

I. MODEL PRESENTATION

In this section, we introduce Multivariate Quadratic Hawkes processes (MQHawkes). We start by recalling the definition of Hawkes and Quadratic Hawkes processes in their univariate form. We then define their extension to the multivariate case, with a special focus on the bivariate case. Next we introduce a much simplified specification, where the quadratic feedback can be factorized into a product of *trend indicators* – a limit called ZHawkes in [5], after Zumbach's seminal work on the time reversal asymmetry of financial time series. We discuss endogeneity ratios and stationarity conditions for MQHawkes within the ZHawkes framework. Finally, we further generalize the MQHawkes model to take into account the possible intrinsic correlations between the underlying (subordinated) Poisson processes.

A. Hawkes and QHawkes Processes

Hawkes processes were first introduced to model earthquakes [8, 9]. Their ability to consider the past events in the probability of a future event was indeed very useful to reproduce the aftershock phenomenon: an earthquake is in fact more likely to occur after another just took place. This same property appeared to hold in a variety of fields, like biological neural networks [10, 11], financial markets [12, 13], and also crime rates or riot propagation [14, 15].

A Hawkes process $(N_t)_{t\geq 0}$ is an inhomogenous[1] Poisson process with intensity λ_t defined from the past realisations of the process itself, according to:

$$\lambda_t = \lambda_\infty + \int_{-\infty}^t \phi(t - u)\, \mathrm{d}N_u \qquad (1)$$

with

$$n = \int_{\mathbb{R}} \phi(u)\, \mathrm{d}u < 1, \qquad (2)$$

where λ_∞ is the baseline intensity, ϕ is the feedback kernel and n represents the "endogeneity ratio", which needs to be strictly less than unity in order to guarantee that the process reaches a stationary state.

Like Hawkes processes, QHawkes processes are inhomogenous Poisson processes with an intensity defined from past events. In the context of financial markets, the event process $(N_t)_{t>0}$ represents the sequence of

prices changing events, where the price change is defined by $\mathrm{d}P_t = \epsilon_t \psi \mathrm{d}N_t$ [5] where $\epsilon_t = \pm 1$ (random sign), and ψ is the tick size (the smallest unit of price change) which we set to one henceforth. The intensity self-exciting feedback loop is now defined on the returns $\mathrm{d}P_t$ rather than on the events $\mathrm{d}N_t$. Hence, the sign of the returns influences the process intensity [5]:

$$\begin{aligned} \lambda_t =& \lambda_\infty + \int_{-\infty}^t L(t - s)\, \mathrm{d}P_s \\ &+ \iint_{-\infty}^t K(t - s, t - u)\, \mathrm{d}P_s \mathrm{d}P_u, \end{aligned} \qquad (3)$$

where λ_∞ is again the baseline intensity, $L(.)$ is the so-called leverage kernel (capturing the increase of activity following a drop in prices) and $K(.,.)$ is the quadratic kernel.

Following [5] we simplify the quadratic kernel K assuming it can be written as a general time-diagonal and a rank one contribution:

$$K(s, u) = \phi(s)\delta(s - u) + k(s)k(u). \qquad (4)$$

This defines what we shall refer to as the ZHawkes model:

$$\begin{aligned} \lambda_t =& \lambda_\infty + \int_{-\infty}^t L(t - s)\, \mathrm{d}P_s + \int_{-\infty}^t \phi(t - s)\, \mathrm{d}N_s \\ &+ \left(\int_{-\infty}^t k(t - s)\, \mathrm{d}P_s \right)^2, \end{aligned} \qquad (5)$$

where we have used $(\mathrm{d}P_s)^2 = \mathrm{d}N_s$. Note that the diagonal term precisely reproduces the standard Hawkes feedback.

B. MQHawkes Processes

Let us now consider N financial assets with prices $(P_t^i)_{t\geq 0, i=1,\ldots,N}$. We associate such price processes with jump processes $(N_t^i)_{t\geq 0}$, with:

$$\mathrm{d}P_t^i = \epsilon_i \psi_i \mathrm{d}N_t^i, \qquad (6)$$

where we set henceforth $\psi_i \equiv 1$, $\forall i$, with loss of generality. Each jump process $(N_t^i)_{t\geq 0}$ is a conditionally independent[2] Quadratic Hawkes process with intensity

---

[1] Here, inhomogenous means that the intensity of the process is time dependent.

[2] This means that for given intensities $\lambda_{i,t}$, the subordinated Poisson processes $\mathrm{d}N_t^i$ are independent. See further down for the case of correlated jump processes.

$(\lambda_{i,t})_{t \geq 0}$

$$\lambda_{i,t} = \lambda_{i,\infty} + \sum_{j=1}^{N} \int_{-\infty}^{t} L_j^i(t-s)\, \mathrm{d}P_s^j$$
$$+ \sum_{j \leq k}^{N} \iint_{-\infty}^{t} K_{jk}^i(t-s, t-u)\, \mathrm{d}P_s^j\, \mathrm{d}P_u^k. \quad (7)$$

Note that the superscript in the kernels indicates which asset is affected by the feedback, whereas subscripts indicate which assets are responsible for it.

In the following we shall generically call $\mathbb{K}_\mathrm{d}$ the "diagonal" feedback terms, i.e. the kernels K_{jk}^i with $j = k$. These terms describe the quadratic feedback from two price changes of the same asset j onto the activity of asset i. Similarly, $\boldsymbol{K}_\times$ describes cross effects, i.e. K_{jk}^i with $j < k$, from two different assets $j \neq k$ onto the activity of asset i. In order to guarantee that intensities λ_t^i remain positive at all times, kernels $L, \mathbb{K}_\mathrm{d}$ and $\boldsymbol{K}_\times$ need to verify some conditions, which are detailed in Appendix 2.

Although all these terms could in principle play a role, in the present paper we shall restrict to cross terms K_{jk}^i such that either j or k are equal to i. The intuitive meaning of such terms will become clear below in the context of ZHawkes processes; in particular we shall see why terms with $j \neq i$ and $k \neq i$ are not expected to play a large role in practice.

For the sake of clarity we will mainly focus on the bivariate case $N = 2$, for which the leverage kernel and the diagonal quadratic kernel are 2×2 matrices:

$$\mathbb{L} := \begin{pmatrix} L_1^1 & L_2^1 \\ L_1^2 & L_2^2 \end{pmatrix}; \quad \mathbb{K}_\mathrm{d} := \begin{pmatrix} K_{11}^1 & K_{22}^1 \\ K_{11}^2 & K_{22}^2 \end{pmatrix},$$

whereas the cross quadratic kernel is a vector

$$\boldsymbol{K}_\times := \begin{pmatrix} K_{12}^1 \\ K_{12}^2 \end{pmatrix}.$$

In the following 2×2 matrices and 2D vectors are respectively noted $\mathbb{A}$ and $\boldsymbol{A}$, for example $\boldsymbol{\lambda} := (\lambda_1, \lambda_2)$. We will also need to distinguish the "time diagonal" of a matrix $\mathbb{A}(\tau_1, \tau_2)$, by which we mean $\mathbb{A}(\tau, \tau)$, from the "diagonal" of $\mathbb{A}$, which refers to the diagonal components in asset space $\mathbb{A}_{ii}$.

C. ZHawkes Model

The multivariate generalisation of the ZHawkes model amounts to assuming that the quadratic kernels $\mathbb{K}_\mathrm{d}$ and $\boldsymbol{K}_\times$ can be written as arbitrary singular time-diagonal and time-rank-one contributions, to wit:

$$K_{jj}^i(s, u) = (\phi_\mathrm{d})_{jj}^i(s)\, \delta(s-u) + k_{jj}^i(s) k_{jj}^i(u)$$
$$K_{ij}^i(s, u) = (\phi_\times)_{ij}^i(s)\, \delta(s-u) + k_{ji}^i(s) k_{ij}^i(u) \quad (i < j).$$

However, as long as we consider independent Poisson processes, we can set $\boldsymbol{\phi}_\times = \boldsymbol{0}$. This is because the two processes will almost never jump simultaneously, such that for all u, $\mathrm{d}P_u^i \mathrm{d}P_u^j = 0$ when $i \neq j$ (see however Eq. (11) below when co-jumps are taken into account). Actually in the so-called Thinning Algorithm [16] for multivariate processes, commonly used to simulate inhomogenous Poisson process, at most one process can be jumping at each time step.

D. Endogeneity Ratio and Stationarity Condition

The endogeneity ratios indicate by how much, on average, the feedback loop contributes to the future of the process, and thus, whether the process is stationary or not. To define them for the MQHawkes we use analogies with univariate QHawkes and multivariate linear Hawkes.

1. Mean Intensity

Since price changes are centred and processes are assumed to be independent, we have $\mathbb{E}(\mathrm{d}\boldsymbol{P}) = 0$ and $\mathbb{E}(\mathrm{d}P_s^i \mathrm{d}P_u^j) = \delta_{ij}\delta_{s,u}\bar{\lambda}^i \mathrm{d}s$. Using this, we find that the vector of mean intensities $\bar{\boldsymbol{\lambda}}$ writes:

$$\bar{\boldsymbol{\lambda}} = \left(\mathbb{I} - \int_0^\infty \mathbb{K}_\mathrm{d}(s, s)\mathrm{d}s \right)^{-1} \boldsymbol{\lambda}_\infty. \quad (8)$$

This expression must be interpreted with care when $\mathbb{K}_\mathrm{d}(s, u)$ contains a singular time-diagonal contribution $\phi_\mathrm{d}(s)\, \delta(s-u)$. In such a case, and throughout this paper, we will interpret $\mathbb{K}_\mathrm{d}(s, s)$ as $\mathbb{K}_{\mathrm{d,reg.}}(s, s) + \phi_\mathrm{d}(s)$, where $\mathbb{K}_{\mathrm{d,reg.}}$ is the regular part of the diagonal quadratic feedback. Equation (8) shows that the mean intensity diverges when the spectral radius[3] of the matrix $\int_0^\infty \mathbb{K}_\mathrm{d}(s, s)\mathrm{d}s$ tends to one from below.

2. Endogeneity Ratio

In the multidimensional case, the endogeneity ratio of a standard linear Hawkes process is defined by the spectral radius of the kernel matrix ϕ involved in the expression of $\bar{\boldsymbol{\lambda}}$ [12]. More precisely, introducing

$$\|f\| := \int_0^{+\infty} f(s)\mathrm{d}s$$

---

[3] The spectral radius of a square matrix is the maximum of the absolute values of its eigenvalues.

one constructs the matrix $\|(\phi)^i_j\|$ and determines its eigenvalue with largest modulus, which in turn defines the dominant endogeneity ratio n.

For a general univariate QHawkes model, one can always decompose the endogeneity ratio n as the sum of the Hawkes endogeneity ratio n_H associated with the singular time-diagonal contribution to $K(\cdot,\cdot)$, and a regular contribution n_Q:

$$n = \int_0^{+\infty} \phi(s)\mathrm{d}s + \int_0^{+\infty} K_{\text{reg.}}(s,s)\mathrm{d}s := n_H + n_Q. \quad (9)$$

In the special case of a ZHawkes quadratic kernel, the regular contribution reads:

$$n_Q = n_Z = \int_0^{+\infty} k^2(s)\mathrm{d}s. \quad (10)$$

For a general MQHawkes, the endogeneity ratio is then defined by the spectral radius of $\int_0^\infty \mathbb{K}_\mathrm{d}(s,s)\mathrm{d}s$. However, when decomposing the kernel as a singular time-diagonal contribution and a regular contribution, one must be careful about the fact that the two matrices $\mathbb{K}_{\mathrm{d,reg.}}(s,s)$ and ϕ_d do not commute in general. Hence the spectral radius defining the endogeneity ratio cannot be simply expressed as the sum of a Hawkes contribution n_H (i.e. the spectral radius of $\|\phi_\mathrm{d}\|$) and of a regular or Zumbach contribution (i.e. the spectral radius of $\|\mathbb{K}_{\mathrm{d,reg.}}\|$).

3. Stationarity Condition

For linear Hawkes processes to be stationary, the endogeneity ratio n needs to be strictly less than one. Hence, only processes with finite mean intensity $\bar{\boldsymbol{\lambda}}$ can be stationary. In the presence of a quadratic feedback, the situation is more intricate. In a previous communication [17] we found that for a univariate quadratic Hawkes process to be stationary one needs $n_H < 1$, but not necessarily $n = n_H + n_Q < 1$. When $n_H < 1$ and $n > 1$ one has a stationary process with an infinite mean. In the present multivariate framework, we conjecture that a similar situation holds: defining n_H as the spectral radius of $\|\phi_\mathrm{d}\|$, one has that if $n > 1$ and $n_H < 1$ the process is stationary with infinite mean intensity, and if $n < 1$ and $n_H < 1$, then the process is stationary with finite mean intensity.

E. Correlated Poisson Processes

One may wonder if the hypothesis of independence of the $\mathrm{d}N^i_t$ for different $i = 1,\ldots,N$ is not too strong to faithfully account for events happening in financial markets. In fact, as in [18], we find that *co-jumps* (i.e. simultaneous jumps in the price of different assets within

1 minute bins) occur fairly frequently, adding a new stylised fact to consider in this multivariate framework. In order to investigate co-jumps empirically, we use a jump detection method (see [19]) on 295 large cap. US stock prices from January 2015 to December 2020. We then count how many stocks display anomalous price jumps in a given minute, and represented such counts in Figure 1. Co-jumps are clearly seen to occur. On average, there are 4.5 co-jumps per day, and up to 68 co-jumps in one day. Using [19] to classify each jump as endogenous or exogenous, we compute the cumulative density function of number of stocks included in endogenous co-jumps which displays a power law of slope -1.25 (see insert in Fig. 1).
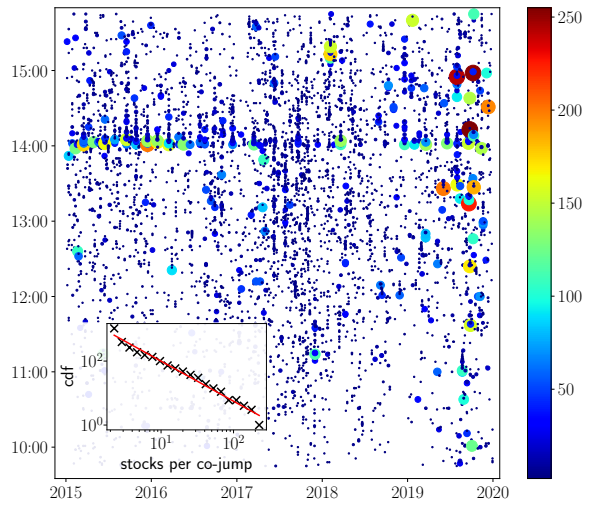

Figure 1: Time series of 1 minute bins co-jumps in a set of 295 US stocks over 5 years. The horizontal axis corresponds to the day in the sample and the vertical axis gives the time of day. The size and the color of the circle encodes the number of stocks simultaneously jumping in a given minute (see color bar). The insert shows the cumulative distribution function of the number of stocks in co-jumps for jumps that were classified as endogenous according to the methodology developed in [19], the slope is -1.25.

Co-jumps may be due to either a common exogenous shock (like an external piece of news affecting several stocks), or to some endogenous instability triggering a jump for one given stock, which propagates to other stocks. The very interesting question of the exogenous/endogenous nature of co-jumps clearly needs a more refined investigation, in the spirit of [19], and is left for future work.

In [18], the authors show that multivariate linear Hawkes models with independent realisations of the Poisson process do not satisfactorily reproduce co-

jumps. Here we propose a way to enforce correlations between Poisson processes, and allow one to account for co-jumps within bivariate QHawkes processes.

1. Bivariate Poisson Processes

We focus on the bivariate case $N = 2$. The extension to $N > 2$ is also possible, although beyond the scope of the present paper. In order to allow for the possibility of co-jumps, i.e. such that $\mathrm{d}P_t^1\mathrm{d}P_t^2 \neq 0$, we consider three independent QHawkes counting processes $(N_t^1, N_t^2, N_t^c)_{t\geq 0}$ with intensities $(\mu_{1,t}, \mu_{2,t}, \mu_{c,t})$ defined from past returns:

$$
\begin{aligned}
\mu_{a,t} =\mu_{a,\infty} &+ \sum_{j\in\{1,2\}} \int_{-\infty}^t L_j^a(t-s)\mathrm{d}P_s^j \\
&+ \sum_{j\leq k\in\{1,2\}} \iint_{-\infty}^t Q_{jk}^a(t-s,t-u)\mathrm{d}P_s^j\mathrm{d}P_u^k,
\end{aligned}
\tag{11}
$$

with $a = 1, 2, c$, and we model price moves as:

$$
\begin{aligned}
\mathrm{d}P_t^1 &= \epsilon_t^1(\mathrm{d}N_t^1 + \mathrm{d}N_t^c) \\
\mathrm{d}P_t^2 &= \epsilon_t^2(\mathrm{d}N_t^2 + \mathrm{d}N_t^c),
\end{aligned}
\tag{12}
$$

where $\epsilon_t^i = \pm 1$, with $\mathbb{E}(\epsilon_t^i) = 0$ and $\mathbb{E}(\epsilon_t^1\epsilon_t^2) = \rho_t$. While the sign correlation ρ_t could indeed be time dependent, here we will assume that $\rho_t = \rho$ is independent of time. Thus, price moves have both an idiosyncratic part, represented by $\mathrm{d}N^i$, $i \in \{1,2\}$, and a common part $\mathrm{d}N^c$, which make co-jumps possible (hence the subscript c). In fact, one now has:

$$
\mathrm{d}P_t^1\mathrm{d}P_t^2 = \epsilon_t^1\epsilon_t^2\mathrm{d}N_t^c. \tag{13}
$$

We can then define the total intensities $(\lambda_{1,t}, \lambda_{2,t})_{t\geq 0}$ of the price jumps $(P_t^1, P_t^2)_{t\geq 0}$, according to the definition of Poisson processes:

$$
\begin{cases}
\lambda_{1,t} := \mu_{1,t} + \mu_{c,t} \\
\lambda_{2,t} := \mu_{2,t} + \mu_{c,t},
\end{cases}
\tag{14}
$$

and notice that the structure of the equations governing the dynamics of intensities is conveniently the same as in the independent case, with the following identifications:

$$
\begin{aligned}
L_1^i &\leftarrow L_1^i + L_1^c & L_2^i &\leftarrow L_2^i + L_2^c \\
K_{11}^i &\leftarrow Q_{11}^i + Q_{11}^c & K_{22}^i &\leftarrow Q_{22}^i + Q_{22}^c \\
K_{12}^i &\leftarrow Q_{12}^i + Q_{12}^c & \lambda_\infty^i &\leftarrow \mu_\infty^i + \mu_\infty^c.
\end{aligned}
$$

2. Mean Intensity

Now correlations between instantanous price changes are non-zero, i.e. $\mathbb{E}(\mathrm{d}P_s^1\mathrm{d}P_s^2) = \rho\overline{\mu^c}\mathrm{d}s \neq 0$ (see Ap-

pendix 3), the expression for the mean intensity $\bar{\boldsymbol{\lambda}}$ is modified and reads, for regular kernels:

$$
\begin{aligned}
\bar{\boldsymbol{\lambda}} =&\left(\mathbb{I} - \int_0^\infty \mathbb{K}_\mathrm{d}(s,s)\mathrm{d}s\right)^{-1} \\
&\left(\boldsymbol{\lambda}_\infty + \rho\overline{\mu^c}\int_0^\infty \mathbf{K}_\times(s,s)\mathrm{d}s\right),
\end{aligned}
\tag{15}
$$

with:

$$
\overline{\mu^c} = \frac{\mu_\infty^c + \bar{\boldsymbol{\lambda}}^\top \cdot \int_0^\infty \boldsymbol{Q}^c(s,s)\mathrm{d}s}{1 - \rho\int_0^\infty Q_{12}^c(s,s)\mathrm{d}s},
$$

and

$$
\boldsymbol{Q}^c = \begin{pmatrix} Q_{11}^c \\ Q_{22}^c \end{pmatrix}.
$$

In the presence of a singular contribution to the K's, one should again take it into account by substituting all $K(s,s)$ by the corresponding $K_\mathrm{reg}(s,s) + \phi(s)$, and similarly for Q's. The stationarity conditions are now that:

1. The spectral radius of the Hawkes kernel $||\phi_\mathrm{d}||$ is strictly less than 1

2. The time-diagonal of the co-jump kernel must be such that:

$$
\left|\rho \int_0^\infty Q_{12}^c(s,s)\mathrm{d}s\right| < 1.
$$

II. COVARIANCE STRUCTURE & YULE-WALKER EQUATIONS

Naturally, the feedback kernels can not be directly observed in data. However, they can be computed from inverting covariance functions, which can easily be estimated from empirical data. Here we introduce the covariance structures of a multivariate QHawkes process, thereby establishing the matrix Yule-Walker equations (which link covariance structures and QHawkes kernels).

A. Covariance Definition

The first quantity of interest is the covariance of the activities of the process. For all $\tau \neq 0$:

$$
\mathbb{C}_{ij}(\tau) := \mathbb{E}\left(\frac{\mathrm{d}N_t^i}{\mathrm{d}t}\frac{\mathrm{d}N_{t-\tau}^j}{\mathrm{d}t}\right) - \bar{\lambda}_i\bar{\lambda}_j. \tag{16}
$$

As for the univariate QHawkes, $\mathbb{C}$ is even, and its extension to $\tau = 0$ can be worked out following [20]. Thus, for $i = j$, the extension will be the same as in

[5], with $\mathbb{C}_{ii}^\star(\tau) := \mathbb{C}_{ii}(\tau) + \delta_{\tau,0}\bar{\lambda}_i$. (Note indeed that $\mathbb{E}((\mathrm{d}N^i)^2) = \mathbb{E}(\mathrm{d}N^i) = \bar{\lambda}_i \mathrm{d}t$, if we consider that events cannot overlap). For $i \neq j$, the extension must account for co-jumps and now reads $\mathbb{C}_{ij}^\star(\tau) := \mathbb{C}_{ij}(\tau) + \delta_{\tau,0}\bar{\mu}_\mathrm{c}$. Without co-jumps, one has $\mathbb{C}_{ij}^\star(\tau) := \mathbb{C}_{ij}(\tau)$ for $i \neq j$.

We also define a relevant three-point correlation structure $\mathcal{D}$ as a tensor following:

$$\mathcal{D}_{ijk}(\tau_1, \tau_2) = \mathbb{E}\left(\frac{\mathrm{d}N_t^i}{\mathrm{d}t} \frac{\mathrm{d}P_{t-\tau_1}^j}{\mathrm{d}t} \frac{\mathrm{d}P_{t-\tau_2}^k}{\mathrm{d}t}\right). \qquad (17)$$

Note that within our hypothesis that the return signs are of mean zero and uncorrelated, this quantity is only nonzero when $\tau_1 > 0$ and $\tau_2 > 0$, or when $\tau_1 = \tau_2$. In the latter case, $\mathcal{D}_{ijj}(\tau, \tau) = \mathbb{C}_{ij}(\tau)$.

In the bivariate case, $\mathcal{D}$ defines again two types of 2×2 matrices: "diagonal" ($j = k$) and "cross" ($j \neq k$). We shall consistently use the following notations to distinguish them:

$$\mathbb{D}_\mathrm{d}(\tau_1, \tau_2) := \begin{pmatrix} \mathcal{D}_{111}(\tau_1, \tau_2) & \mathcal{D}_{122}(\tau_1, \tau_2) \\ \mathcal{D}_{211}(\tau_1, \tau_2) & \mathcal{D}_{222}(\tau_1, \tau_2) \end{pmatrix}, \qquad (18)$$

and

$$\mathbb{D}_\times(\tau_1, \tau_2) = \begin{pmatrix} \mathcal{D}_{112}(\tau_1, \tau_2) & \mathcal{D}_{121}(\tau_1, \tau_2) \\ \mathcal{D}_{212}(\tau_1, \tau_2) & \mathcal{D}_{221}(\tau_1, \tau_2) \end{pmatrix}. \qquad (19)$$

B. Two-point Yule-Walker Equations

In order to deduce kernels from empirical correlations, direct relations must be determined. The method we use to find such relations is very similar to that used in Appendix 1 of [5]. Assuming that prices are martingales,[4] and without considering co-jumps, we find the following matrix equation for $\mathbb{C}$:

$$\mathbb{C}(\tau) = \mathbb{K}_\mathrm{d}(\tau)\bar{\mathbb{\lambda}} + \int_0^\infty \mathbb{K}_\mathrm{d}(u,u)\mathbb{C}(\tau - u)\mathrm{d}u$$
$$+ 2\int_0^{+\infty} \int_{u^+}^{+\infty} \mathbb{K}_\mathrm{d}(\tau + u, \tau + r)\mathbb{D}_\mathrm{d}(u,r)\mathrm{d}r\mathrm{d}u$$
$$+ \int_0^{+\infty} \int_{u^+}^{+\infty} \mathbf{K}_\times(\tau + u, \tau + v)(\mathbf{D}_\times)^\top(u,v)\mathrm{d}v\mathrm{d}u$$
$$+ \int_0^{+\infty} \int_{v^+}^{+\infty} \mathbf{K}_\times(\tau + u, \tau + v)(\mathbf{D}_\times)^\top(u,v)\mathrm{d}u\mathrm{d}v, \qquad (20)$$

where $\bar{\mathbb{\lambda}}$ is a 2×2 matrix defined as

$$\bar{\mathbb{\lambda}} := \begin{pmatrix} \bar{\lambda}_1 & 0 \\ 0 & \bar{\lambda}_2 \end{pmatrix}. \qquad (21)$$

and where

$$\mathbf{D}_\times := \begin{pmatrix} \mathcal{D}_{112} \\ \mathcal{D}_{212} \end{pmatrix} \qquad (22)$$

This Yule-Walker equation boils down to Eq. (8) of Ref. [5] in the univariate case.

The Yule-Walker equation for $\mathbb{C}$ accounting for co-jumps can be found in Appendix 4.

C. Three-point Yule-Walker Equations

The full three-point Yule-Walker equations for the tensor $\mathcal{D}$ are quite intricate. In order to give a simplified version of these equations we restrict here to the case with no co-jumps, i.e. $\mu^\mathrm{c} \equiv 0$. We find that, for $\tau_1 > \tau_2 > 0$:

$$\mathbb{D}_\mathrm{d}(\tau_1, \tau_2) = 2\mathbb{K}_\mathrm{d}(\tau_1, \tau_2)\overline{\mathbb{C}}_\mathrm{d}(\tau_2 - \tau_1)$$
$$+ \int_{\tau_1^+}^{+\infty} \mathbb{K}_\mathrm{d}(u)\mathbb{D}_\mathrm{d}(\tau_1 - u, \tau_2 - u)\,\mathrm{d}u$$
$$+ 2\int_{\tau_1^+}^{+\infty} \mathbb{K}_\mathrm{d}(u, \tau_1)\mathbb{D}_\mathrm{d}^0(u - \tau_1, \tau_2 - \tau_1)\,\mathrm{d}u$$
$$+ \int_{\tau_1^+}^{+\infty} \mathbf{K}_\times(\tau_1, u)(\mathbf{D}_\times^1)^\top(u - \tau_1, \tau_2 - \tau_1)\mathrm{d}u$$
$$+ \int_{\tau_1^+}^{+\infty} \mathbf{K}_\times(u, \tau_1)(\mathbf{D}_\times^2)^\top(u - \tau_1, \tau_2 - \tau_1)\mathrm{d}u, \qquad (23)$$

where we have introduced the two following diagonal matrices:

$$\overline{\mathbb{C}}_\mathrm{d}(\tau) := \begin{pmatrix} \mathbb{C}_{11}(\tau) + (\bar{\lambda}_1)^2 & 0 \\ 0 & \mathbb{C}_{22}(\tau) + (\bar{\lambda}_2)^2 \end{pmatrix}, \qquad (24)$$

$$\mathbb{D}_\mathrm{d}^0(\tau_1, \tau_2) := \begin{pmatrix} \mathcal{D}_{111}(\tau_1, \tau_2) & 0 \\ 0 & \mathcal{D}_{222}(\tau_1, \tau_2) \end{pmatrix}, \qquad (25)$$

as well as the notation $\mathbf{D}_\times$ for the 2-vectors:

$$\mathbf{D}_\times^1 := \begin{pmatrix} \mathcal{D}_{121} \\ 0 \end{pmatrix}, \quad \mathbf{D}_\times^2 := \begin{pmatrix} 0 \\ \mathcal{D}_{212} \end{pmatrix}.$$

For the corresponding Yule-Walker equation for $\mathbb{D}_\times$ see Appendix 4.

D. Asymptotic Behavior with Power Law Kernels

An interesting special case for which the Yule-Walker equations can be asymptotically solved is when kernels are decaying as power-laws, as considered in [5]. Of special interest is the relationships between exponents

---

[4] This assumption is often violated at high frequencies, and some amendments will need to be introduced when calibrating the model on actual HF data – see our Part II companion paper [7].

governing the correlation functions and the kernel exponents when $\tau \to +\infty$.

The general analysis is quite cumbersome and is relegated to Appendix 5 a. In the non-critical case ($n < 1$), the calculations are not particularly difficult and generalise the results of [5] to the multivariate case.

The critical case ($n = 1$), on the other hand, is much more subtle, in particular for the QHawkes processes.[5] In order to treat this case, we have generalised the method introduced by Hawkes in [23], which combines the Yule-Walker equations in the Fourier domain with Liouville's Theorem to find a direct relationship between exponents. This method is recalled in Appendix 5 b, from which the value of the exponents in the critical case can be derived, see Tables in Appendix 6. Some of the exponent values reported in [5] turn out to be incorrect; the correct values can be inferred from our new results. Note that in the critical case the decay of the correlation functions cannot be faster than τ^{-1}, as in the case of linear Hawkes models without ancestors (see the work of Brémaud & Massoulié [24]).

III. POWER-LAW TAILS OF THE VOLATILITY DISTRIBUTION

Here we investigate the distribution of volatility of a MQHawkes process. We adapt the methodology of Ref. [5] to the multivariate setting, restricting for simplicity to the two assets case. Our main goal is to establish that MQHawkes lead to fat (power-law) tails for the empirical intensity distribution, which translates into fat tails in the distribution of returns [5]. We limit our study to the ZHawkes specification with exponentially decaying kernels, that allow one to construct a tractable continuous time limit.

A. ZHawkes Model with Exponential Kernels

We neglect the leverage feedback and set $\mathbb{L} = 0$. We also neglect the possible presence of co-jumps, i.e. set $\mu_c \equiv 0$ hereafter. Within the ZHawkes specification, we can rewrite the intensities as follows:

$$\begin{cases} \lambda_1 & = H_1^1 + H_2^1 + (Z_1^1)^2 + (Z_2^1)^2 + Y_1^1 \\ \lambda_2 & = H_1^2 + H_2^2 + (Z_1^2)^2 + (Z_2^2)^2 + Y_1^2 \end{cases}$$

with

$$H_j^i := \int_{-\infty}^{t} (\phi_d)_{jj}^i(t-s)\mathrm{d}N_s^j$$

---

[5] For the linear version, one can refer to [21] or to [22] where Bacry, Jaisson and Muzy use convolution to link the Fourier transform of $\mathbb{C}$ with that of $\mathbb{K}$.

and

$$Z_j^i = \int_{-\infty}^{t} k_{jj}^i(t-s)\mathrm{d}P_s^j$$

$$Y_j^i = \left(\int_{-\infty}^{t} k_{ji}^i(t-s)\mathrm{d}P_s^i\right)\left(\int_{-\infty}^{t} k_{ij}^i(t-u)\mathrm{d}P_u^j\right).$$

To keep things tractable, we choose all kernels to be exponentials and consider that only four "features" are important to describe all feedback effects, namely:

$$\begin{cases} h_1(t) = \beta_1 \int_{-\infty}^{t} e^{-\beta_1(t-s)}\mathrm{d}N_s^1 \\ h_2(t) = \beta_2 \int_{-\infty}^{t} e^{-\beta_2(t-s)}\mathrm{d}N_s^2, \end{cases} \quad (26)$$

for activity feedback, and

$$\begin{cases} z_1(t) = \omega_1 \int_{-\infty}^{t} e^{-\omega_1(t-s)}\mathrm{d}P_s^1 \\ z_2(t) = \omega_2 \int_{-\infty}^{t} e^{-\omega_2(t-s)}\mathrm{d}P_s^2, \end{cases} \quad (27)$$

for trend feedback, with ω_i's and β_i's positive constants. From such features, we construct the quantities H, Z and Y as

$$H_1^i = n_{H,1}^i h_1(t) \qquad H_2^i = n_{H,2}^i h_2(t)$$
$$Z_1^i = a_{Z,1}^i z_1(t) \qquad Z_2^i = a_{Z,2}^i z_2(t)$$
$$Y_2^1 = a_{Z,\times}^1 z_1(t)z_2(t) \qquad Y_1^2 = a_{Z,\times}^2 z_1(t)z_2(t).$$

Under such hypotheses, the intensities write:

$$\lambda_{1,t} = \lambda_{1,\infty} + n_{H,1}^1 h_1(t) + n_{H,2}^1 h_2(t) + (a_{Z,1}^1 z_1(t))^2$$
$$\qquad + (a_{Z,2}^1 z_2(t))^2 + a_{Z,\times}^1 z_2(t)z_1(t)$$
$$\lambda_{2,t} = \lambda_{2,\infty} + n_{H,1}^2 h_1(t) + n_{H,2}^2 h_2(t) + (a_{Z,1}^2 z_1(t))^2$$
$$\qquad + (a_{Z,2}^2 z_2(t))^2 + a_{Z,\times}^2 z_2(t)z_1(t).$$
$$(28)$$

B. Endogeneity Ratios

As mentioned earlier, the Hawkes endogeneity ratio n_H is obtained as the spectral radius of the 2×2 matrix of Hawkes coefficients, namely

$$\mathbb{N}_H := \begin{pmatrix} n_{H,1}^1 & n_{H,2}^1 \\ n_{H,1}^2 & n_{H,2}^2 \end{pmatrix}. \quad (29)$$

The Zumbach endogeneity coefficient n_Z is the spectral radius of the trend feedback kernel. In the monovariate case, one finds $n_Z = \int_0^{+\infty} k^2(s)\mathrm{d}s$. From the features defined above, we now have $k_{jj}^i(t) = a_{Z,j}^i \omega_j e^{-\omega_j t}$. Noting that

$$\omega^2 \int_0^{\infty} e^{-2\omega s}\,\mathrm{d}s = \frac{\omega}{2},$$

we simply need, in the present case, to find the top eigenvalue of the following 2×2 matrix:

$$\mathbb{N}_Z := \frac{1}{2} \begin{pmatrix} (a_{Z,1}^1)^2 \omega_1 & (a_{Z,2}^1)^2 \omega_2 \\ (a_{Z,1}^2)^2 \omega_1 & (a_{Z,2}^2)^2 \omega_2 \end{pmatrix}. \tag{30}$$

In the simplest case when all coefficients and decay rates are equal, one finds that n_H is equal to any of the coefficients $n_{H,j}^i$ (which are all equal), and $n_Z = a_Z^2 \omega$, where $a_Z := a_{Z,j}^i$ (again all equal).

Note that in the general case, one needs to diagonalize the matrix $\mathbb{N}_H + \mathbb{N}_Z$, which leads to a total endogeneity ratio that is in general different from $n_H + n_Z$.

C. Fokker-Planck Equation

As in [5], we consider the process on a time scale $T > 0$, that shall eventually tend to $+\infty$, and simultaneously take all decay rates $\beta, \omega \to 0$, but in such a way that ωT and βT remain finite. For the Zumbach feedback not to disappear in this limit one needs to simultaneously scale up both $(a_{Z,j}^i)_{i,j \in \{1,2\}}$ as $\omega^{-1/2}$ and $(a_{Z,\times}^i)_{i \in \{1,2\}}$ as ω^{-1}. In this scaling regime, one can establish the following Fokker-Planck equation for the time dependent probability density p_t of (h_1, h_2, z_1, z_2) (see Appendix 7):

$$\begin{aligned}
\partial_t p_t = {} & \beta_1 \partial_{h_1}\big((h_1 - \lambda_{1,t})p_t\big) + \beta_2 \partial_{h_2}\big((h_2 - \lambda_{2,t})p_t\big) \\
& + \omega_1 \partial_{z_1}\big(z_1 p_t\big) + \omega_2 \partial_{z_2}\big(z_2 p_t\big) \\
& + \frac{\omega_1^2}{2} \partial_{z_1 z_1}^2\big(\lambda_{1,t} p_t\big) + \frac{\omega_2^2}{2} \partial_{z_2 z_2}^2\big(\lambda_{2,t} p_t\big),
\end{aligned} \tag{31}$$

where p_t is a shorthand for $p_t(h_1, h_2, z_1, z_2)$, $\lambda_{1,t}$ and $\lambda_{2,t}$ are given by Eq. (28), and where we have disregarded co-jumps and direct correlations between the returns of asset 1 and asset 2, meaning that $\mathbb{E}(\mathrm{d}P^1 \mathrm{d}P^2) = 0$. The inclusion of such correlations can be considered and adds further cross-derivative terms $\partial_{z_1 z_2}^2$ in the Fokker-Planck equation.

Solving for the stationary distribution p_∞ of the Fokker Planck equation allows one to determine the tail behaviour of the distribution of intensities (which translate into volatilities since $\mathbb{E}[(\mathrm{d}P_t)^2] = \lambda_t \mathrm{d}t$). In the monovariate case, Ref. [5] established that p_∞ decays as a power-law, with an exponent α that depends on both n_H and n_Z. The general expression for α is however not available in closed form, although asymptotic results in various regimes could be worked out, in particular when $n_H \to 0$. The most important conclusion is that $\alpha \to \infty$ when $n_Z \to 0$, i.e. power-law tails disappear in the absence of a quadratic Zumbach coupling. Interestingly, the coupling between the Hawkes feedback with $n_H \sim 1$ and even a small Zumbach effect ($n_Z \ll 1$) was shown to generate an exponent compatible with empirical data.

D. ZHawkes without Hawkes ($n_H = 0$)

To make further analytical progress, we now consider the case in which the Hawkes coupling is absent ($n_H = 0$), that is when h and z decouple, leading to a tractable two-dimensional Fokker-Planck for z_1 and z_2. In the stationary regime, the Fokker Planck equation (31) becomes:

$$\begin{aligned}
0 = {} & \omega_1 \partial_{z_1}\big(z_1 p_\infty(z_1, z_2)\big) + \omega_2 \partial_{z_2}\big(z_2 p_\infty(z_1, z_2)\big) \\
& + \frac{\omega_1^2}{2} \partial_{z_1 z_1}^2\big(\lambda_1 p_\infty(z_1, z_2)\big) + \frac{\omega_2^2}{2} \partial_{z_2 z_2}^2\big(\lambda_2 p_\infty(z_1, z_2)\big).
\end{aligned} \tag{32}$$

This equation describes the stationary measure of the stochastic path of the bivariate process $(Z_1, Z_2)_t$ defined as:

$$\begin{aligned}
\mathrm{d}Z_1 &= -\omega_1 Z_1 \mathrm{d}t + \omega_1 \sqrt{\lambda_{1,t}} \mathrm{d}W_t^1 \\
\mathrm{d}Z_2 &= -\omega_2 Z_2 \mathrm{d}t + \omega_2 \sqrt{\lambda_{2,t}} \mathrm{d}W_t^2,
\end{aligned}$$

with

$$\begin{aligned}
\lambda_{1,t} &= \lambda_{1,\infty} + a_{Z,\times}^1 Z_2 Z_1 + (a_{Z,1}^1 Z_1)^2 + (a_{Z,2}^1 Z_2)^2 \\
\lambda_{2,t} &= \lambda_{2,\infty} + a_{Z,\times}^2 Z_2 Z_1 + (a_{Z,2}^2 Z_2)^2 + (a_{Z,1}^2 Z_1)^2.
\end{aligned} \tag{33}$$

These equations allow one to simulate paths $(Z_1, Z_2)_t$ numerically, from which an empirical determination of $p_\infty(z_1, z_2)$ can be confronted to our analytical solution of Eq. (32). Note that the coefficients $(a_{Z,j}^i)_{i,j \in \{1,2\}}$ and $(a_{Z,\times}^i)_{i \in \{1,2\}}$ must satisfy the following inequalities for $\lambda_{1,t}$ and $\lambda_{2,t}$ to remain positive at all times:

$$4(a_{Z,1}^i a_{Z,2}^i)^2 \geq (a_{Z,\times}^i)^2, \qquad i = 1, 2.$$

How can one determine the tail exponent for such a two dimensional process? We first introduce polar coordinates (r, θ), such that $z_1 = r \cos \theta$ and $z_2 = r \sin \theta$. We then surmise that when $r^2 \gg \max(\lambda_{1,\infty}, \lambda_{2,\infty})$ the stationary distribution $p_\infty(r, \theta)$ factorizes into an angular component $F(\theta)$ and a power-law contribution, to wit: $p_\infty(r, \theta) \underset{r \to \infty}{\approx} F(\theta) r^{-\alpha}$, where α is the tail exponent. Note that α should be strictly larger than 2 for p_∞ to be normalisable, and strictly larger than 3 for the mean intensity to be non divergent. Injecting our factorized guess into Eq. (32) and taking the limit $r^2 \gg \lambda^\infty$, we find a second order ODE on the function F, where α appears as a parameter (see Appendix 8, Eq. (A.13)).

The value of α is selected by the analogue of energy quantification in quantum mechanics: only for some special values of α can one find a solution F of the above ODE that satisfies the correct boundary conditions compatible with symmetries of the problem. Clearly, F must be everywhere non-negative and, because of the symmetry $z_1 \to -z_1$ and $z_2 \to -z_2$, one must have $F(\theta + \pi) = F(\theta)$ (see below for explicit examples). In principle, there can be more than one value of

α that allows one to find an acceptable solution F. This is the analogue of the energy spectrum in quantum mechanics. The value of α that governs the tail behaviour of $p_\infty(r,\theta)$ is then the *smallest* of all such acceptable values. Once the asymptotic tail behaviour of $p_\infty(r,\theta)$ is known, it is easy to derive the tail behaviour of the marginals $p_\infty(z_1)$ and $p_\infty(z_2)$, which both behave as $z^{-1-\mu}$ with $\mu = \alpha - 2$. The volatility distribution then has the same tail behaviour.

In order to illustrate this general procedure on a simple example, we focus in the following on the case where $a_{Z,i}^j = \sqrt{2n_Z/\omega_i}$, $\forall i,j = 1,2$, such that the two eigenvalues of $\mathbb{N}_Z$ are equal to n_Z (the Zumbach endogeneity ratio) and 0. We also set $a_{Z,\times}^i = 2\gamma n_Z/\sqrt{\omega_1\omega_2}$, where γ is an arbitrary coefficient $\in (-2,2)$ (such that λ_1 and λ_2 are always positive). When $\omega_1 = \omega_2$, Eq. (A.13) considerably simplifies and reads:

$$\begin{aligned}&\left[(\alpha-2)(\alpha-\alpha_0) + (\alpha-2)^2\gamma\cos(\theta)\sin(\theta)\right]F(\theta) \\ &+ [(1+\gamma\cos(\theta)\sin(\theta))F(\theta)]'' = 0,\end{aligned} \quad (34)$$

where we have defined

$$\alpha_0 := 2 + \frac{1}{n_Z}.$$

Note that for a given value of α this equation is invariant under the simultaneous change $\gamma \to -\gamma$, $\theta \to -\theta$. Hence the value of α can only depend on $|\gamma|$.

1. The isotropic case $\gamma = 0$

When cross terms $a_{Z,\times}^i$ are absent and $\omega_1 = \omega_2$, the problem becomes isotropic in the sense that the dynamics of $r^2 = z_1^2 + z_2^2$ decouple from that of θ. The problem then boils down to the univariate QHawkes model without Hawkes coupling, for which the value of α is known, and given by α_0.

Furthermore, the evolution of θ is that of a free Brownian motion on the unit circle, leading to a uniform distribution $F(\theta) = F_0$, which is indeed a solution of Eq. (34) in this case. Note that other periodic solutions exist whenever

$$(\alpha-2)(\alpha-\alpha_0) = 4\ell^2, \qquad \ell = 0, 1, 2, \dots$$

but lead to larger values of α when $\ell > 0$.

2. The case $\gamma \neq 0$

In order to make progress, we posit that α and F can be expanded as power series of γ, namely

$$\begin{aligned}\alpha &= \alpha_0 + \alpha_1\gamma + \alpha_2\gamma^2 + \dots \\ F(\theta) &= F_0 + F_1(\theta)\gamma + F_2(\theta)\gamma^2 + \dots\end{aligned}$$
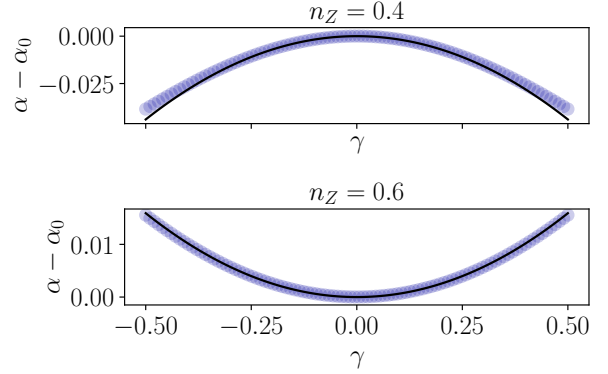
Figure 2: Plot of $\alpha - \alpha_0$ as function of γ for $n_Z = 0.4$ (above) and $n_Z = 0.6$ (below). The α are found numerically by looking for the smallest α for a π-periodic and positive solution F. The solid black line represents the theoretical prediction, Eq. (35), that is $\alpha \approx \alpha_0 - 0.18\gamma^2$ for $n_Z = 0.4$ and $\alpha \approx \alpha_0 + 0.06\gamma^2$ for $n_Z = 0.6$.

where $\alpha_0 = 2 + 1/n_Z$ is the solution for $\gamma = 0$ and F_0 is the constant solution found above. The coefficient α_1 must be zero for symmetry reasons.

Inserting this expansion in Eq. (34) and imposing that all $F_n(\cdot)$ remain π-periodic, the identification of terms of order γ^n finally leads to

$$\alpha = \alpha_0 + \frac{\gamma^2}{32}\left(\frac{4}{n_Z} - \frac{1}{n_Z^3}\right) + O(\gamma^4). \quad (35)$$

Figure 2 displays the numerical values of α as a function of γ for $n_Z = 0.4$ and $n_Z = 0.6$ in with the corresponding theoretical parabolas, see Eq. (35). Note that the γ^2 correction changes sign when $n_Z = 1/2$.

In this case ($n_Z = 1/2$), finding an exact solution of the associated Schrodinger solution is possible [25], and leads to $\alpha = 4$ for all values of γ.[6] The corresponding solution for F is also known in that case and is a constant independent of θ, as can be directly checked from Eq. (34). An expansion around the special point $n_Z = 1/2$ can in fact be performed and leads to first order to

$$\alpha = 4 + \frac{16}{4+\gamma^2}\left(\frac{1}{2} - n_Z\right) + o\left(\frac{1}{2} - n_Z\right)$$

.

---

[6] The other exact solutions found in [25] unfortunately correspond to sub-dominant, larger values of α.

3. The case $n_Z \to +\infty$

Equation (34) can be easily analysed when $n_Z \to +\infty$, in which case α tends to the smallest possible value, 2, corresponding to a maximally "fat" distribution. The only periodic solution to Eq. (34) in that limit is

$$F_\infty(\theta) = \frac{C}{2 + \gamma \sin(2\theta)}, \qquad (36)$$

where C is a constant. When n_Z is very large but not infinite, we posit that the solution writes

$$F(\theta) = F_\infty(\theta) + \frac{1}{n_Z^2} G(\theta) + O(\frac{1}{n_Z^4}), \qquad (37)$$

together with $\alpha = 2 + \zeta/n_Z$, with $G(\cdot)$ and ζ to be determined. Inserting Eq. (37) in the ODE on F Eq. (34), one obtains the ODE for G:

$$[(2 + \gamma \sin(2\theta))G(\theta)]'' = -C \frac{2\zeta(\zeta - 1) + \gamma \zeta^2 \sin(2\theta)}{2 + \gamma \sin(2\theta)}. \qquad (38)$$

Imposing $G(\cdot)$ to be π-periodic, that is imposing that the right hand side of Eq. (38) integrates to zero between $\theta = 0$ and $\theta = \pi$, sets the value of ζ:

$$\zeta = \left(1 - \frac{\gamma^2}{4}\right)^{-\frac{1}{2}}, \qquad (39)$$

recovering $\zeta = 1$ when $\gamma = 0$ and the small γ expansion result above, see Eq. (35). The function $G(\cdot)$ is plotted in Fig. 3 for $\gamma = 1$.

When comparing the solution above with the histogram of simulated θ for $n_Z = 10$ and $\gamma = 1$, we find an excellent match with F_∞, without need of any correction, see Fig. 4. This is expected since the correction term $G(\theta)/n_Z^2$ is of the order of 1% in that
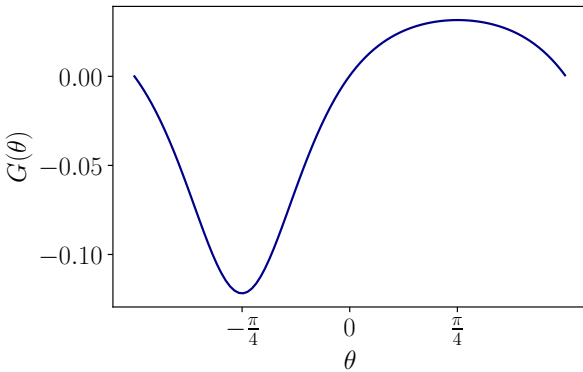
Figure 3: Plot of $G(\theta)$ for $\gamma = 1$, $C = 0.6$, see Eq. (38).
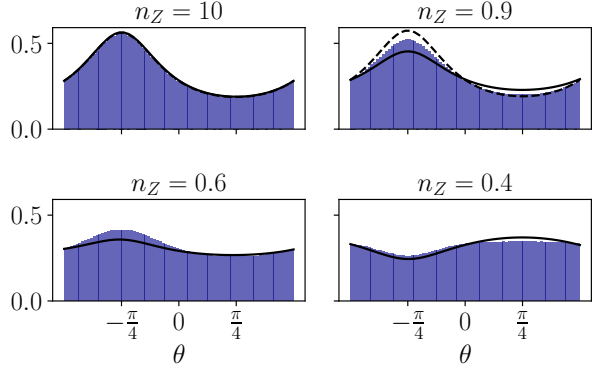
Figure 4: Comparison between the numerical solution F of Eq. (34) (black lines) and the results of simulating Eq. (33) (purple histogram) for several values of n_Z, and $\gamma = 1$. The solid line shows the numerical solution F found while looking for a solution couple of a positive π-periodic solution F of Eq. (34) and a value α_{num} (the value of α is found numerically). The the dashed line for $n_Z = 10$ and $n_Z = 0.9$ shows the solution F_∞ for $\alpha = 2$ ($n_Z \to +\infty$). The upper right figure, obtained for $n_Z = 10$, matches very well the solution obtained when $n_Z \to +\infty$ since corrections are $O(n_Z^{-2})$. Note that when $n_Z = 1/2$, $F(\theta)$ is a constant.

case. When n_Z decreases, corrections become more pronounced. The numerically computed F is in good agreement with the angular distribution obtained from a direct numerical simulation of the two dimensional stochastic process (Eq. (33)).

4. The case $n_Z \to 0$

When n_Z tends to zero, we expect that the exponent α of the power-law tail diverges. Looking again for α of the form $\alpha = 2 + \zeta/n_Z$, we find that Eq. (34) reads:

$$\begin{aligned}&\left(2\zeta(\zeta - 1) + \zeta^2 \gamma \sin(2\theta)\right) F(\theta) \\ &+ n_Z^2 \left[(2 + \gamma \sin(2\theta))F(\theta)\right]'' = 0.\end{aligned} \qquad (40)$$

When $n_Z \to 0$, this equation looks self-contradictory because the remaining term can only be zero if $F(\theta) = 0$. However, the second derivative term is a singular perturbation, so it must be treated with care. The idea is to look for a solution F which is zero nearly everywhere, except very close to some special values of θ where the second derivative diverges. It turns out that all the action takes place close to $\theta = \pi/4$ when $\gamma > 0$ and $\theta = -\pi/4$ when $\gamma < 0$. Choosing $\gamma > 0$, $\theta = \pi/4 + u\sqrt{n_Z}$, and $n_Z \to 0$, Eq. (40) becomes

the Schrodinger equation of a harmonic oscillator, up to terms $O(n_Z^2)$:

$$\left[-(2+\gamma)\frac{d^2\Psi}{du^2} + 2\zeta^2\gamma u^2\Psi\right] = \frac{2\zeta(\zeta-1)+\zeta^2\gamma}{n_Z}\Psi,$$

with

$$\Psi(u) := F\left(\frac{\pi}{4} + u\sqrt{n_Z}\right).$$

The smallest α solution (or "ground state") is

$$\Psi(u) = C'e^{-\varkappa u^2} + O(n_Z), \qquad \varkappa = \zeta\sqrt{\frac{\gamma}{4+2\gamma}}, \quad (41)$$

where C' is another constant, together with

$$\zeta = \frac{2}{2+\gamma} + \sqrt{\frac{2\gamma}{2+\gamma}}n_Z + O(n_Z^2). \qquad (42)$$

This solution is only accurate in a region of width $\sim \sqrt{n_Z}$ around $\pi/4$, beyond which it quickly goes to zero. This is in perfect agreement with the numerical solution of Eq. (34) for small n_Z, shown in Fig. 5. The parabolic shape in a semi-log plot around $\theta = \pi/4$ agrees with the predictions of Eq. (41). Moreover, the value of $\alpha = 2+\zeta/n_Z$, with ζ given in Eq. (42), also perfectly matches the numerical values reported in Fig. 6. Note that for $\gamma < 0$, the same results hold with $|\gamma|$ replacing γ in the above equations.

E. The General Case

In the previous section, we have shown how to compute the power-law tail exponent in the case where only
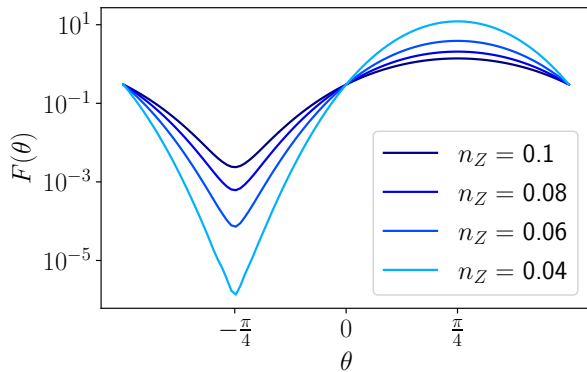
Figure 5: Solution F of Eq. (34) when $n_Z \leq 0.1$. Note the logarithm scale on the y-axis, such that the Gaussian solution for $\Psi(u)$, Eq. (41), is a parabola around $\theta = \pi/4$.
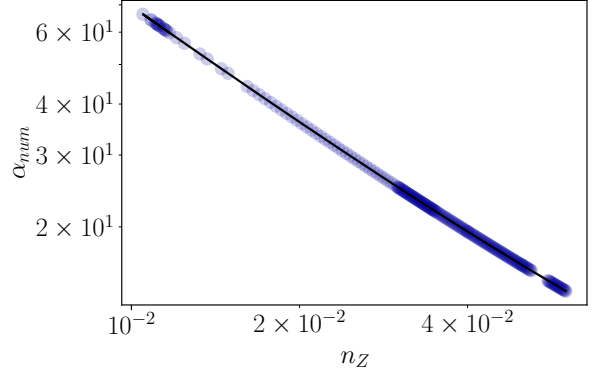
Figure 6: Tail exponent α as a function of n_Z for $n_Z \leq 0.06$. Circles: numerical determination of α such that F is a π-periodic solution of equation (34). Solid black line: theoretical value $\alpha = 2 + \zeta/n_Z$ with ζ given by Eq. (42)

the quadratic "ZHawkes" kernel is present. We also restricted to simple cases where the frequencies ω and coupling constants a_Z are symmetric. Although analytically more challenging, the method outlined above can be implemented more generally, and amounts to solving a problem akin to the quantification condition for the Schrödinger equation. Similarly to the monovariate case, the volatility distribution develops a power-law tail for all values of the Hawkes feedback n_H, as long as some amount of Zumbach feedback n_Z is present. When $n_H \neq 0$, the equation setting the tail in the bivariate case is a three dimensional partial derivative equation generalizing the equation written in the appendix of Ref. [5]. We expect that even a tiny amount of Zumbach feedback coupled to the standard Hawkes effect brings the exponent α into the empirical range, as found in the monovariate case [5].

IV. CONCLUSION

Let us summarise what we have achieved in this study. Building on the work of Blanc *et al.* [5], we extended the Quadratic Hawkes model to a multivariate framework (MQHawkes). We emphasized that in the multivariate case, both idiosyncratic and common jumps must be considered in the general case, leading to a quite complex framework that we only detailed in the bivariate case.

We defined the endogeneity ratio of MQHawkes, as well as the associated conditions for the process to be stationary. Within the ZHawkes approximation – where quadratic kernels write as a sum of a time diagonal component, reproducing a linear Hawkes feedback, and a

rank one component – we gave a deeper understanding of the roles of the different feedback terms in the stationarity condition. We found that, in particular, the spectral radius of the Hawkes component needs to be strictly inferior to 1, as for the 1D case. The rank one component contains both exciting and inhibiting realisations and is not involved in the condition, although such a contribution can lead to a divergence of the average intensity of the process, see [17].

We further defined the covariance structures for MQHawkes processed and established the associated Yule-Walker equations. The latter allow one to fully determine the quadratic kernels from data, and thereby pave the way for their empirical calibration.

Finally, we studied the volatility distribution of a 2D MQHawkes process. Restricting our study to ZHawkes without Hawkes ($n_H = 0$), with exponential kernels and in symmetric cases, we were able to characterize exactly the tail of the joint probability density function of the ZHawkes intensity terms. We found that it displays a power law behavior, inline with the observed fat tails of financial returns.

In a forthcoming companion paper [7], we shall calibrate the model on empirical data. On the theoretical side, note that, while expected to be quite heavy, it would be interesting to further develop the analysis in the presence of co-jumps and correlations. Also, in the present analysis of the volatility distribution, we restricted to symmetric coefficients and exponential kernels. Confirming that our conclusions are qualitatively robust against changes in symmetry and kernel functionals, and studying its quantitative implications would also be of interest. Finally, note that, for the sake of clarity, most of our expressions are given in the 2D case. Expanding them further in the N-dimensional case will be necessary in order to calibrate the model on a large number of assets, say a pool of stocks.

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

[2] Bacry E, Delour J, Muzy JF. Multifractal random walk. Physical Review E. 2001;64(2):026103.

[3] Wu P, Muzy JF, Bacry E. From Rough to Multifractal volatility: the log S-fBM model. arXiv preprint arXiv:220109516. 2022.

[4] Jaisson T, Rosenbaum M, et al. Limit theorems for nearly unstable Hawkes processes. Annals of Applied Probability. 2015;25(2):600–631.

[5] Blanc P, Donier J, Bouchaud JP. Quadratic Hawkes processes for financial prices. Quantitative Finance. 2017;17(2):171–188.

[6] Zumbach G. Time reversal invariance in finance. Quantitative Finance. 2009;9(5):505–515.

[7] Aubrun C, Hey N, Benzaquen M, Jean-Philippe B. Multivariate Quadratic Hawkes Processes - Part II: Empirical Analysis. In preparation.

[8] Koshelev M. Testing Earthquake Damage Prevention Methods with Simulated Fractal Earthquakes. physical Research. 1981;2853:2862.

[9] Zhuang J, Harte DS, Werner MJ, Hainzl S, Zhou S. Basic models of seismicity: Temporal models. 2012.

[10] Osorio I, Frei MG, Sornette D, Milton J, Lai YC. Epileptic seizures: quakes of the brain? Physical Review E. 2010;82(2):021919.

[11] Sornette D, Osorio I. Prediction. chapter in "Epilepsy: The Intersection of Neurosciences, Biology, Mathematics, Physics and Engineering", Editors: Osorio I, Zaveri HP, Frei MG, Arthurs S, CRC Press, Taylor & Francis Group, pp 203-237. 2010.

[12] Bacry E, Mastromatteo I, Muzy JF. Hawkes processes in finance. Market Microstructure and Liquidity. 2015;1(01):1550005.

[13] Fosset A, Bouchaud JP, Benzaquen M. Non-parametric estimation of quadratic Hawkes processes for order book events. The European Journal of Finance. 2021:1–16.

[14] Mohler GO, Short MB, Brantingham PJ, Schoenberg FP, Tita GE. Self-exciting point process modeling of crime. Journal of the American Statistical Association. 2011;106(493):100–108.

[15] Bonnasse-Gahot L, Berestycki H, Depuiset MA, Gordon MB, Roché S, Rodriguez N, et al. Epidemiological modelling of the 2005 French riots: a spreading wave and the role of contagion. Scientific reports. 2018;8(1):1–20.

[16] Ogata Y. On Lewis' simulation method for point processes. IEEE transactions on information theory. 1981;27(1):23–31.

[17] Aubrun C, Benzaquen M, Bouchaud JP. Hawkes processes with infinite mean intensity. Physical Review E. 2022;105(3):L032101.

[18] Bormetti G, Calcagnile LM, Treccani M, Corsi F, Marmi S, Lillo F. Modelling systemic price cojumps with Hawkes factor models. Quantitative Finance. 2015;15(7):1137–1156.

[19] Marcaccioli R, Bouchaud JP, Benzaquen M. Exogenous and endogenous price jumps belong to different dynamical classes. Journal of Statistical Mechanics: Theory and Experiment. 2022;2022(2):023403.

[20] Hawkes AG. Spectra of some self-exciting and mutually exciting point processes. Biometrika. 1971;58(1):83–90.

[21] Chicheportiche R, Bouchaud JP. The fine-structure of volatility feedback I: Multi-scale self-reflexivity. Physica A: Statistical Mechanics and its Applications. 2014;410:174–195.

[22] Bacry E, Jaisson T, Muzy JF. Estimation of slowly decreasing hawkes kernels: application to high-frequency order book dynamics. Quantitative Finance. 2016;16(8):1179–1201.

[23] Hawkes AG. Point spectra of some mutually exciting point processes. Journal of the Royal Statistical Society: Series B (Methodological). 1971;33(3):438–443.

[24] Brémaud P, Massoulié L. Hawkes branching point processes without ancestors. Journal of applied probability. 2001:122–135.

[25] Xie QT. New quasi-exactly solvable periodic potentials. Journal of Physics A: Mathematical and Theoretical. 2011;44(28):285302.

[26] Blanc P, Chicheportiche R, Bouchaud JP. The fine structure of volatility feedback II: Overnight and intraday effects. Physica A: Statistical Mechanics and its Applications. 2014;402:58–75.

1. Notations

The notations and formulae described on this page will be used throughout the paper.

$\mathbb{A}$	Matrix
$\mathbb{A}^{\top}$	Matrix transpose
$\mathbb{A}(\tau,\tau)$	Matrix time diagonal
$\mathbb{A}_{ii}$	Matrix diagonal
$\mathbf{X}$	Vector
$(N_t^i)_{t\geq 0}$	Quadratic Hawkes process describing the time of the price changes of asset i
$(N_t^c)_{t\geq 0}$	Quadratic Hawkes process describing the times of co-jumps for 2 price processes in the correlated case
$(P_t^i)_{t\geq 0}$	Price process of the asset i
ψ	Tick size
$\lambda_{i,t}$	Intensity of the QHawkes process N^i at time t
$\lambda_{i,\infty}$	Baseline intensity of the QHawkes process N^i
$\mu_{i,t}$	Intensity of the QHawkes process N^i at time t in the correlated case
$\mu_{i,\infty}$	Baseline intensity of the QHawkes process N^i in the correlated case
$n/n_H/n_Q/n_Z$	Endogeneity ratio of the process/Hawkes contribution/regular contribution/Zumbach contribution
ϵ_t	Sign of the price change at time t, $\epsilon_t \in \{-1, 1\}$
L_j^i	Leverage kernel of intensity i reflecting the feedback of the price changes of asset j
K_{jk}^i	Quadratic kernel of intensity i reflecting the feedback of the price changes of assets j and k
Q_{jk}^i	Same as K_{jk}^i for the correlated case
ρ	Correlation of the signs of the price changes dP^1 and dP^2 in the correlated case
$\mathbb{K}_d$	Kernel matrix such that $\mathbb{K}_{d\,ij} = K_{jj}^i$
$\mathbf{K}_\times$	Kernel vector representing cross feedback, composed by K_{jk}^i with $j < k$
$\bar{\boldsymbol{\lambda}}$	Mean Intensity
$\delta_{\cdot,\cdot}$	Kronecker delta: $\delta_{x,y} = 1$, when $x = y$ and $\delta_{x,y} = 0$ otherwise
$\delta(\cdot)$	Dirac
$\mathbb{C}_{ij}$	Covariance $\mathbb{C}_{ij}(\tau) := \mathbb{E}\left(\dfrac{dN_t^i}{dt}\dfrac{dN_{t-\tau}^j}{dt}\right) - \bar{\lambda}_i\bar{\lambda}_j$
$\mathcal{D}_{ijk}$	3 points correlation structure $\mathcal{D}_{ijk}(\tau_1,\tau_2) := \mathbb{E}\left(\dfrac{dN_t^i}{dt}\dfrac{dP_{t-\tau_1}^j}{dt}\dfrac{dP_{t-\tau_2}^k}{dt}\right)$
$\mathbb{D}_\times$	Square (2x2) correlation matrix such that $\mathbb{D}_{\times\,ij} = \mathcal{D}_{ijj}$
$\mathbb{D}_d$	Square (2x2) correlation matrix such that $\mathbb{D}_{\times\,ij} = \mathcal{D}_{ijk}$ with $j \neq k$
$(\phi_d)_{jj}^i$	Time diagonal contribution of $(\mathbb{K}_d)_{jj}^i$
k_{jk}^i	Rank one contribution of K_{jk}^i
H_j^i	Hawkes component of intensity i from feedback j; $H_j^i = n_{H,j}^i h_j(t)$
$n_{H,j}^i$	Endogeneity ratio of Hawkes component H_j^i
Z_j^i	Quadratic component of intensity i from feedback j
$a_{Z,j}^i$	Amplitude of the quadratic component Z_j^i; $a_{Z,i}^j = \sqrt{2n_Z/\omega_i}$
Y_j^i	Cross component of intensity i; $Y_j^i = a_{Z,\times}^i z_1 z_2$
$a_{Z,\times}^i$	Amplitude of the cross component of intensity i
γ	Arbitrary coefficient $\gamma \in (-2, 2)$, such that $a_{Z,\times}^i = 2\gamma n_Z/\sqrt{\omega_1\omega_2}$
$p_\infty(r,\theta) \underset{r\to\infty}{\approx} F(\theta)r^{-\alpha}$	Joint pdf of (z_1, z_2) in polar coordinates in the case ZHawkes without Hawkes with exponential kernels.

2. Conditions for Positive QHawkes Intensities

Inspired by the method in [26], we detail here sufficient conditions for the intensity of the process i to be positive, when considering N assets. We first consider the simple case $\mathbb{L} \equiv 0$. First, considering that the kernels are negligible up to a value q, and using a discrete approximation of the integration, we rewrite the intensity associated with asset i:

$$\lambda_{i,t} = \lambda_{i,\infty} + \sum_{j \leq k}^{N} \iint_{-\infty}^{t} K_{jk}^{i}(t-s, t-u)\, \mathrm{d}P_{s}^{j}\, \mathrm{d}P_{u}^{k} \simeq \lambda_{i,\infty} + \sum_{j \leq k}^{N} \sum_{u=1}^{q} \sum_{s=1}^{q} \mathrm{d}P_{t-u} K_{jk}^{i}(s,u) \mathrm{d}P_{t-s}$$

So we can write:

$$\lambda_{i,t} = \lambda_{i,\infty} + \boldsymbol{r_t}^{\top} \mathbb{K}_i \boldsymbol{r_t}$$

where

$$\mathbb{K}_i = \begin{pmatrix} K_{11}^{i}(1,1) & \dots & K_{11}^{i}(1,q) & \dots & \frac{1}{2}K_{1N}^{i}(1,1) & \dots & \frac{1}{2}K_{1N}^{i}(1,q) \\ \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ K_{11}^{i}(q,1) & \dots & K_{11}^{i}(q,q) & \dots & \frac{1}{2}K_{1N}^{i}(q,1) & \dots & \frac{1}{2}K_{1N}^{i}(q,q) \\ & \vdots & & \ddots & & \vdots & \\ K_{1N}^{i}(1,1) & \dots & K_{1N}^{i}(1,q) & \dots & \frac{1}{2}K_{NN}^{i}(1,1) & \dots & \frac{1}{2}K_{NN}^{i}(1,q) \\ \vdots & \ddots & \vdots & & \vdots & \ddots & \vdots \\ K_{1N}^{i}(q,1) & \dots & K_{1N}^{i}(q,q) & \dots & \frac{1}{2}K_{NN}^{i}(q,1) & \dots & \frac{1}{2}K_{NN}^{i}(q,q) \end{pmatrix}, \qquad \boldsymbol{r_t} = \begin{pmatrix} \mathrm{d}P_{t-1}^{1} \\ \vdots \\ \mathrm{d}P_{t-q}^{1} \\ \vdots \\ \mathrm{d}P_{t-1}^{N} \\ \vdots \\ \mathrm{d}P_{t-q}^{N} \end{pmatrix}$$

So $\mathbb{K}_i$ is a symmetric blocs matrix where the bloc j of the diagonal is $\mathrm{bloc}_{jj} = \begin{pmatrix} K_{11}^{i}(1,1) & \dots & K_{jj}^{i}(1,q) \\ \vdots & \ddots & \vdots \\ K_{jj}^{i}(q,1) & \dots & K_{jj}^{i}(q,q) \end{pmatrix}$ and the bloc

(k,j) when $k < j$ is $\mathrm{bloc}_{kj} = \frac{1}{2}\begin{pmatrix} K_{kj}^{i}(1,1) & \dots & K_{kj}^{i}(1,q) \\ \vdots & \ddots & \vdots \\ K_{kj}^{i}(q,1) & \dots & K_{kj}^{i}(q,q) \end{pmatrix}$ (when $j < k$ the cross kernels is K_{jk}^{i}).

Thus, in this case, where $\mathbb{L} \equiv 0$, sufficient condition to keep the intensity positive is to have $\mathbb{K}_i$ positive semi-definite.

If we now consider the case where $\mathbb{L} \neq 0$, following the same method, we can write:

$$\lambda_{i,t} = \lambda_{i,\infty} + \mathbf{L}_i \boldsymbol{r_t} + \boldsymbol{r_t}^{\top} \mathbb{K}_i \boldsymbol{r_t}$$

where $\boldsymbol{r_t}$ and $\mathbb{K}_i$ are the same as before and $\mathbf{L}_i$ is

$$\mathbf{L}_i = \begin{pmatrix} L_1^i(1) & \dots & L_1^i(q) & \dots & L_N^i(1) & \dots & L_N^i(q) \end{pmatrix}^{\top}$$

Assuming $\mathbb{K}_i$ is invertible, one can complete the square by writing:

$$\begin{aligned} \lambda_{i,t} &= \lambda_{i,\infty} + \mathbf{L}_i \boldsymbol{r_t} + \boldsymbol{r_t}^{\top} \mathbb{K}_i \boldsymbol{r_t} \\ &= \lambda_{i,\infty} + \left(\boldsymbol{r_t} + \frac{1}{2}\mathbb{K}_i^{-1}\mathbf{L}_i\right)^{\top} \mathbb{K}_i \left(\boldsymbol{r_t} + \frac{1}{2}\mathbb{K}_i^{-1}\mathbf{L}_i\right) - \frac{1}{4}\mathbf{L}_i^{\top} \mathbb{K}_i^{-1} \mathbf{L}_i \end{aligned}$$

In conclusion, a sufficient condition for intensities to stay positive when $\mathbb{L} \neq 0$ is to have all kernels $\mathbb{K}_i$ positive definite and

$$\lambda_{i,\infty} \geq \frac{1}{4}\mathbf{L}_i^{\top} \mathbb{K}_i^{-1} \mathbf{L}_i, \qquad \forall i$$

3. Stationary Condition in Bivariate Case

In the bivariate case, we no longer have $\mathbb{E}(\mathrm{d}P_u^i \mathrm{d}P_u^j) = 0$ when $i \neq j$. Thus, when computing $\mathbb{E}(\boldsymbol{\lambda_t})$ to find the mean intensity and thus the stationary condition, we need to compute the mean of cross feedback.

We first introduce $m(dt, \mathrm{d}\xi^1, \mathrm{d}\xi^2)$ the joint Punctual Poisson Measure associate to the jump processes of the prices (P^1, P^2). So it is a pure jump process with i.i.d jump sizes $(\xi^1 \xi^2)$ of common law p on $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$. We assume $\int_{\mathbb{R}} \xi^1 \xi^2 p(\mathrm{d}\xi^1, \mathrm{d}\xi^2) = \psi_1 \psi_2$.

$$
\begin{aligned}
\frac{1}{\psi_1 \psi_2} \mathbb{E}\Big(\int_{-\infty}^t \int_{-\infty}^t K_{12}(t-s, t-u) \mathrm{d}P_s^1 \mathrm{d}P_u^2 \Big) &= \frac{1}{\psi_1 \psi_2} \mathbb{E}\Big(\int_{-\infty}^t K_{12}(t-u, t-u) \mathrm{d}P_u^1 \mathrm{d}P_u^2 \Big) \\
&= \frac{1}{\psi_1 \psi_2} \mathbb{E}\Big(\int_{\mathbb{R}} \int_{-\infty}^t K_{12}(t-u, t-u) s_u^1 s_u^2 \xi_u^1 \xi_u^2 m(\mathrm{d}u, \mathrm{d}\xi^1, \mathrm{d}\xi^2) \Big) \\
&= \frac{1}{\psi_1 \psi_2} \mathbb{E}\Big(\int_{\mathbb{R}} \int_{-\infty}^t K_{12}(t-u, t-u) s_u^1 s_u^2 \xi_u^1 \xi_u^2 \mu_u^c \mathrm{d}u p(\mathrm{d}\xi^1, \mathrm{d}\xi^2) \Big) \\
&= \frac{1}{\psi_1 \psi_2} \mathbb{E}\Big(\int_{-\infty}^t K_{12}(t-u, t-u) s_u^1 s_u^2 \mu_u^c \mathrm{d}u \Big) \mathbb{E}\Big(\int_{\mathbb{R}} \xi_u^1 \xi_u^2 p(\mathrm{d}\xi^1, \mathrm{d}\xi^2) \Big) \\
&= \int_{-\infty}^t K_{12}(t-u, t-u) \mathbb{E}(s_u^1 s_u^2) \mathbb{E}(\mu_u^c) \mathrm{d}u \\
&= \rho \overline{\mu^c} \int_{-\infty}^t K_{12}(t-u, t-u) \mathrm{d}u
\end{aligned}
$$

4. Yule-Walker Equations: Complement

We write the Yule-Walker equations for the matrix and one for $\mathbb{D}_\times$, the calculations can be made as in Appendix 1 of [5]:

$$
\begin{aligned}
\mathbb{D}_\times(\tau_1, \tau_2) =& \int_{\tau_1^+}^{+\infty} \mathbb{K}_{\mathrm{d}}(u) \mathbb{D}_\times(\tau_1 - u, \tau_2 - u) \mathrm{d}u \\
&+ 2 \int_{\tau_1^+}^{+\infty} \mathbb{K}_{\mathrm{d}}(u, \tau_1) \begin{pmatrix} \mathcal{D}_{112}(u - \tau_1, \tau_2 - \tau_1) & 0 \\ 0 & \mathcal{D}_{221}(u - \tau_1, \tau_2 - \tau_1) \end{pmatrix} \mathrm{d}u \\
&+ \int_{\tau_1^+}^{+\infty} \mathbf{K}_\times(u, \tau_1) \begin{pmatrix} 0 \\ \mathcal{D}_{211}(u - \tau_1, \tau_2 - \tau_1) \end{pmatrix}^\top \mathrm{d}u + \int_{\tau_1^+}^{+\infty} \mathbf{K}_\times(\tau_1, u) \begin{pmatrix} \mathcal{D}_{122}(u - \tau_1, \tau_2 - \tau_1) \\ 0 \end{pmatrix}^\top \mathrm{d}u
\end{aligned}
\tag{A.1}
$$

In the bivariate case, we need to consider $\mathrm{d}N^c$. Thus, $\mathbb{C}$ becomes a 3×3 matrix, and the three-point correlation now also considers the components $\mathbb{D}_{cij}$. With this in mind, we can calculate the Yule-Walker equation for $\mathbb{C}$ in the bivariate case:

$$
\begin{aligned}
\mathbb{C}(\tau) =& \bar{\mathbb{\lambda}} \mathbb{Q}_{\mathrm{d}}(\tau) + \int_0^{+\infty} \mathbb{Q}_{\mathrm{d}}(u, u) \mathbb{C}(\tau - u) \mathrm{d}u + 2 \int_0^{+\infty} \int_{u^+}^{+\infty} \mathbb{Q}_{\mathrm{d}}(\tau + u, \tau + v) \mathbb{D}_{\mathrm{d}}(u, v) \mathrm{d}v \mathrm{d}u \\
&+ \int_0^{+\infty} \int_{u^+}^{+\infty} \mathbf{Q}_\times(\tau + u, \tau + v) \begin{pmatrix} \mathcal{D}_{112} \\ \mathcal{D}_{212} \\ \mathcal{D}_{c12} \end{pmatrix}^\top (u, v) \mathrm{d}v \mathrm{d}u + \int_0^{+\infty} \int_{u^+}^{+\infty} \mathbf{Q}_\times(\tau + u, \tau + v) \begin{pmatrix} \mathcal{D}_{112} \\ \mathcal{D}_{212} \\ \mathcal{D}_{c12} \end{pmatrix}^\top (u, v) \mathrm{d}u \mathrm{d}v \\
&+ \rho \int_0^{+\infty} \begin{pmatrix} Q_{12}^1 \\ Q_{12}^2 \\ Q_{12}^c \end{pmatrix}(u) \begin{pmatrix} \mathcal{C}_{c1} \\ \mathcal{C}_{c2} \\ \mathcal{C}_{c3} \end{pmatrix}^\top (\tau - u) \mathrm{d}u + \int_0^{+\infty} \begin{pmatrix} Q_{11}^1 \\ Q_{22}^2 \\ Q_{cc}^c \end{pmatrix}(u) \begin{pmatrix} \mathcal{C}_{c1} \\ \mathcal{C}_{c2} \\ \mathcal{C}_{c3} \end{pmatrix}^T (\tau - u) \mathrm{d}u
\end{aligned}
$$

where $\bar{\mathbb{\lambda}}$ is a 2×2 matrix defined as

$$
\bar{\mathbb{\lambda}} := \begin{pmatrix} \bar{\lambda}_1 & \bar{\mu}_{\mathrm{c}} \\ \bar{\mu}_{\mathrm{c}} & \bar{\lambda}_2 \end{pmatrix}.
$$

5. Asymptotic Behavior of Decaying Power Law Kernels

a. Asymptotic Forms

We consider decaying power law kernels, such that:

$$\mathbb{K}_{\mathrm{d}}(\tau) \underset{\tau \to +\infty}{\sim} \begin{pmatrix} k_{d1}\tau^{-1-\epsilon_{\mathrm{d}}} & k_{12}\tau^{-1-\epsilon_{\mathrm{o}}} \\ k_{21}\tau^{-1-\epsilon_{\mathrm{o}}} & k_{d2}\tau^{-1-\epsilon_{\mathrm{d}}} \end{pmatrix} \quad \mathbb{K}_{\mathrm{d}}(\tau v_1, \tau v_2) \underset{\tau \to +\infty}{\sim} \begin{pmatrix} \tilde{K}_{11}(v_1,v_2)\tau^{-2\rho_{\mathrm{d}}} & \tilde{K}_{12}(v_1,v_2)\tau^{-2\rho_{\mathrm{o}}} \\ \tilde{K}_{21}(v_1,v_2)\tau^{-2\rho_{\mathrm{o}}} & \tilde{K}_{22}(v_1,v_2)\tau^{-2\rho_{\mathrm{d}}} \end{pmatrix}$$

$$\mathbb{K}_{\times}(\tau) \underset{\tau \to +\infty}{\sim} \begin{pmatrix} k_{x1}\tau^{-1-\epsilon_{\times}} \\ k_{x2}\tau^{-1-\epsilon_{\times}} \end{pmatrix} \quad \mathbb{K}_{\times}(\tau v_1, \tau v_2) \underset{\tau \to +\infty}{\sim} \begin{pmatrix} \tilde{K}_{12}^1(v_1,v_2)\tau^{-2\rho_{\times}} \\ \tilde{K}_{12}^2(v_1,v_2)\tau^{-2\rho_{\times}} \end{pmatrix}$$

Where $\tilde{K}_{11}$, $\tilde{K}_{12}$, $\tilde{K}_{21}$, $\tilde{K}_{22}$, $\tilde{K}_{12}^1$ and $\tilde{K}_{12}^2$ are bounded. Given the Yule-Walker equations of Section II B, we expect the correlation structure to have a similar form. So we look for them as decaying power law functions with parameters defined such as:

$$\mathbb{C}(\tau) \underset{\tau \to +\infty}{\sim} \begin{pmatrix} c_{d1}\tau^{-\beta_{\mathrm{d}}} & c_{12}\tau^{-\beta_{\mathrm{o}}} \\ c_{21}\tau^{-\beta_{\mathrm{o}}} & c_{d2}\tau^{-\beta_{\mathrm{d}}} \end{pmatrix} \quad \mathbb{D}_{\mathrm{d}}(\tau v_1, \tau v_2) \underset{\tau \to +\infty}{\sim} \begin{pmatrix} \tilde{\mathcal{D}}_{111}(v_1,v_2)\tau^{-2\delta_{\mathrm{d}}} & \tilde{\mathcal{D}}_{122}(v_1,v_2)\tau^{-2\delta_{\mathrm{o}}} \\ \tilde{\mathcal{D}}_{211}(v_1,v_2)\tau^{-2\delta_{\mathrm{o}}} & \tilde{\mathcal{D}}_{222}(v_1,v_2)\tau^{-2\delta_{\mathrm{d}}} \end{pmatrix}$$

$$\mathbb{D}_{\times}(\tau) \underset{\tau \to +\infty}{\sim} \begin{pmatrix} d_{112}\tau^{-\beta_{\mathrm{d}}^{\times}} & d_{121}\tau^{-\beta_{\mathrm{o}}^{\times}} \\ d_{212}\tau^{-\beta_{\mathrm{o}}^{\times}} & d_{221}\tau^{-\beta_{\mathrm{d}}^{\times}} \end{pmatrix} \quad \mathbb{D}_{\times}(\tau v_1, \tau v_2) \underset{\tau \to +\infty}{\sim} \begin{pmatrix} \tilde{\mathcal{D}}_{112}(v_1,v_2)\tau^{-2\delta_{\times}} & \tilde{\mathcal{D}}_{121}(v_1,v_2)\tau^{-2\delta_{\times}} \\ \tilde{\mathcal{D}}_{212}(v_1,v_2)\tau^{-2\delta_{\times}} & \tilde{\mathcal{D}}_{221}(v_1,v_2)\tau^{-2\delta_{\times}} \end{pmatrix}$$

As in [5], we make the following hypothesis on the exponents: $\rho_{\mathrm{d}} > \frac{1}{2}$, $\rho_{\mathrm{o}} > \frac{1}{2}$, $\rho_{\times} > \frac{1}{2}$, so the first and second moments are finite. Moreover, we focus on the cases:

- Non critical case ($n_H < 1$), where we assume $0 < \epsilon_{\mathrm{d}} < 1$, $0 < \epsilon_{\mathrm{o}} < 1$,
- Critical case ($n_H = 1$), where we assume $0 < \epsilon_{\mathrm{d}} < \frac{1}{2}$, $0 < \epsilon_o < \frac{1}{2}$,

In the non-critical case ($n < 1$), the method consists in replacing $\mathbb{K}_{\mathrm{d}}$, $\mathbb{K}_{\times}$, $\mathbb{C}$, $\mathbb{D}_{\mathrm{d}}$ and $\mathbb{D}_{\times}$ by their asymptotic expressions presented above in the Yule-Walker equations (Eq. (20), Eq. (23) and Eq. (A.1)) and study the limit $\tau \to +\infty$.

In the following, we describe the method to find the relationship between the exponents of the correlation structures and the exponents of the kernels in the critical case. In a next section, we give the resulting relationships for both the critical and non critical case.

b. Method for the Asymptotic Study in the Critical Case of QHawkes process

In critical case, when $n_H = 1$, the relationship between auto-correlation structures exponents and kernel exponents can not be determined by looking at the limit $\tau \to +\infty$. Thus, to overcome this difficulty, we use a second method to investigate asymptotic behavior using Fourier-domain.

This method for the linear Hawkes can be found in [23]. We present it now for the quadratic Hawkes process. For the sake of simplicity we limit ourselves to the 1D case here, the multivariate case can be worked out similarly.

The definitions of the autocorrelation structures in the 1D case can be found in [5] and are substantially similar to those used here.
We define the Fourier transform of a function f such as:

$$\hat{f}(\omega) = \int_{\mathbb{R}} f(t)e^{-i\omega t}\mathrm{d}t$$

Step 1: Find the regularity of the Yule-Walker terms:
As for the linear case, we start from the Yule-Walker equation on $\mathcal{C}$ for $\tau \neq 0$ (see equation (9) in [5]):

$$\mathcal{C}(\tau) = K(\tau)\bar{\lambda} + \int_0^{+\infty} K(\tau - u)\mathcal{C}(u)\mathrm{d}u + 2\int_{0+}^{+\infty}\int_{u+}^{+\infty} K(\tau + u, \tau + r)\mathcal{D}(u,r)\mathrm{d}r\mathrm{d}u, \tag{A.2}$$

and its extension in 0 in Fourier domain gives:

$$\hat{\mathcal{C}}^*(\omega) = \bar{\lambda} + \hat{\mathcal{C}}(\omega). \tag{A.3}$$

To use the Fourier transform, we need to extend K and $\mathcal{D}$ for $\tau, \tau_1, \tau_2 < 0$. Thus, we consider the function K defined on $\mathbb{R}$ with $K(\tau) = 0$ for $\tau < 0$ similarly for $\mathcal{D}$. Thus, we expect to have $\hat{K}(\omega)$ and $\hat{\mathcal{D}}(\omega)$ regular in the half plan $\text{Im}(\omega) < 0$.

The 2 first terms of Eq. (A.2) are the same as in the Yule-Walker equation for the linear case. Hence, the regularity arguments are the same as in [23].

However, we need to study the last term of Eq. (A.2), $C_3(\tau) = 2 \int_{0+}^{+\infty} \int_{u+}^{+\infty} K(\tau + u, \tau + r)\mathcal{D}(u, r)\mathrm{d}r\mathrm{d}u$.

In order to study the regularity we decompose ω into $\omega = \omega_R + i\omega_I$. Then, switching integration order and using Chasles relation we obtain:

$$\hat{C}_3(\omega) = 2\int_0^{+\infty}\int_{u+}^{+\infty}\int_{-\infty}^{-u} e^{-i\omega_R\tau}e^{\omega_I\tau}K(\tau+u,\tau+r)\mathcal{D}(u,r)d\tau drdu \qquad \Big\}\hat{C}_{3_1}(\omega)$$

$$+ 2\int_0^{+\infty}\int_{u+}^{+\infty}\int_{-u}^{0} e^{-i\omega_R\tau}e^{\omega_I\tau}K(\tau+u,\tau+r)\mathcal{D}(u,r)d\tau drdu \qquad \Big\}\hat{C}_{3_2}(\omega)$$

$$+ 2\int_0^{+\infty}\int_{u+}^{+\infty}\int_{0}^{+\infty} e^{-i\omega_R\tau}e^{\omega_I\tau}K(\tau+u,\tau+r)\mathcal{D}(u,r)d\tau drdu \qquad \Big\}\hat{C}_{3_3}(\omega)$$

We study each term $\hat{C}_{3_1}$, $\hat{C}_{3_2}$ and $\hat{C}_{3_3}$ separately. We first have:

$$\hat{C}_{3_1}(\omega) = 2\int_0^{+\infty}\int_{u+}^{+\infty}\int_{-\infty}^{-u} e^{-i\omega_R\tau}e^{\omega_I\tau}K(\tau+u,\tau+r)\mathcal{D}(u,r)d\tau drdu$$

Here $\tau < -u$ so $\tau + u < 0$, so $K(\tau + u, \tau + r) = 0$, so $\hat{C}_{3_1}(\omega) = 0$.

For the second term, we switch integration order, and we obtain:

$$\hat{C}_{3_2}(\omega) = 2\int_{-\infty}^{0}\int_{-\tau}^{+\infty}\int_{u+}^{+\infty} e^{-i\omega_R\tau}e^{\omega_I\tau}K(\tau+u,\tau+r)\mathcal{D}(u,r)drdud\tau$$

In this case, we have $\tau \geq -u$, so $\tau + u \geq 0$, and $r \geq u^+ \geq -\tau$, so $K(\tau + u, \tau + r) > 0$. We are only interested in $\tau \to -\infty$, so $\hat{C}_{3_2}(\omega)$ is regular in the half plan $\omega_I > 0$.

For $\hat{C_{3_3}}(\omega)$, we have $\tau + u > 0$ and $\tau + r > 0$, so $K(\tau + u, \tau + r) \geq 0$. Since we are only worried about $\tau \to +\infty$. $\hat{C_{3_3}}(\omega)$ is regular in half plan $\omega_I < 0$.

With those results, we now introduce the Yule-Walker equation A.2 in Fourier domain and, as in [23], the function $\hat{B}$:

$$\begin{cases} \hat{\mathcal{C}}(\omega) &= \hat{K}(\omega)D + \hat{K}(\omega)\hat{\mathcal{C}}(\omega) + \hat{C_{3_2}}(\omega) + \hat{C_{3_3}}(\omega) \\ \hat{B}(\omega) &= \hat{K}(\omega)D + \hat{K}(\omega)\hat{\mathcal{C}}(\omega) + \hat{C_{3_2}}(\omega) + \hat{C_{3_3}}(\omega) - \hat{\mathcal{C}}^*(\omega) \end{cases} \qquad (A.4)$$

Thus,

$$\hat{\mathcal{C}}(\omega) = (1 - \hat{K}(\omega))^{-1}(\hat{K}(\omega)D - \hat{B}(\omega) + \hat{C_{3_2}}(\omega) + \hat{C_{3_3}}(\omega)) \qquad (A.5)$$

Since $\mathcal{C}$ is even, we have $(\hat{\mathcal{C}}^{\top}(\omega) = \hat{\mathcal{C}}(-\omega))$ and Eq. (A.5) becomes:

$$(\hat{K}(\omega)D + \hat{C_{3_2}}(\omega) + \hat{C_{3_3}}(\omega) - \hat{B}(\omega))^{\top}(1 - \hat{K}^{\top}(\omega))^{-1} = (1 - \hat{K}(-\omega))^{-1}(\hat{K}(-\omega)D + \hat{C_{3_2}}(-\omega) + \hat{C_{3_3}}(-\omega) - \hat{B}(-\omega))$$

Multiplying by $1 - \hat{K}^{\top}(\omega)$ on the right side and by $1 - \hat{K}(-\omega)$ on the left side, we then develop the expression and mark the regularity propriety below each term.

$$\underbrace{D\hat{K}^{\top}(\omega)}_{\text{Im}(\omega)<0} + \underbrace{\hat{C_{3_2}}^{\top}(\omega)}_{\text{Im}(\omega)>0} + \underbrace{\hat{C_{3_3}}^{\top}(\omega)}_{\text{Im}(\omega)<0} - \underbrace{\hat{B}^{\top}(\omega)}_{\text{Im}(\omega)>0} - \underbrace{\hat{K}(-\omega)\hat{C_{3_2}}^{\top}(\omega)}_{\text{Im}(\omega)>0} - \hat{K}(-\omega)\hat{C_{3_3}}^{\top}(\omega) + \underbrace{\hat{K}(-\omega)\hat{B}^{\top}(\omega)}_{\text{Im}(\omega)>0} =$$

$$\underbrace{\hat{K}(-\omega)D}_{\text{Im}(\omega)>0} + \underbrace{\hat{C_{3_2}}(-\omega)}_{\text{Im}(\omega)<0} + \underbrace{\hat{C_{3_3}}(-\omega)}_{\text{Im}(\omega)>0} - \underbrace{\hat{B}(-\omega)}_{\text{Im}(\omega)<0} - \underbrace{\hat{C_{3_2}}(-\omega)\hat{K}^{\top}(\omega)}_{\text{Im}(\omega)<0} - \hat{C_{3_3}}(-\omega)\hat{K}^{\top}(\omega) + \underbrace{\hat{B}(-\omega)\hat{K}^{\top}(\omega)}_{\text{Im}(\omega)<0}$$

We notice we have a problem with $\hat{C_{3_3}}(-\omega)\hat{K}^\top(\omega)$ and $\hat{K}(-\omega)\hat{C_{3_3}}^\top(\omega)$. In fact, for each of them one factor is regular for $\mathrm{Im}(\omega) > 0$ and the other one for $\mathrm{Im}(\omega) < 0$... If we still reorder the terms to have the left side of the equality regular in the half plan $\mathrm{Im}(\omega) < 0$ and the right side of the equality regular in the half plan $\mathrm{Im}(\omega) > 0$, without defining the regularity plan for the two problematic terms, we obtain the following:

$$\underbrace{D\hat{K}^\top(\omega) + \hat{C_{3_3}}^\top(\omega) - \hat{C_{3_2}}(-\omega) + \hat{B}(-\omega) + \hat{C_{3_2}}(-\omega)\hat{K}^\top(\omega) - \hat{B}(-\omega)\hat{K}^\top(\omega)}_{\mathrm{Im}(\omega)<0} - \hat{K}(-\omega)\hat{C_{3_3}}^\top(\omega) =$$

$$\underbrace{\hat{K}(-\omega)D + \hat{C_{3_3}}(-\omega) - \hat{C_{3_2}}^\top(\omega) + \hat{B}^\top(\omega) + \hat{K}(-\omega)\hat{C_{3_2}}^\top(\omega) - \hat{K}(-\omega)\hat{B}^\top(\omega)}_{\mathrm{Im}(\omega)>0} - \hat{C_{3_3}}(-\omega)\hat{K}^\top(\omega) \tag{A.6}$$

We now need to study in details the regularity of the problematic terms.

Step 2: Develop $\hat{C_{3_3}}(-\omega)\hat{K}^\top(\omega)$ and $\hat{K}(-\omega)\hat{C_{3_3}}^\top(\omega)$

We have

$$\hat{K}(-\omega)\hat{C_{3_3}}^\top(\omega) = 2\int_0^{+\infty}\int_{u+}^{+\infty}\int_0^{+\infty} \left(\hat{K}(-\omega)e^{-i\omega\tau}\right)\mathcal{D}^\top(u,r)K^\top(\tau+u,\tau+r)\mathrm{d}\tau\mathrm{d}r\mathrm{d}u$$

and

$$\hat{C_{3_3}}(-\omega)\hat{K}^\top(\omega) = 2\int_0^{+\infty}\int_{u+}^{+\infty}\int_0^{+\infty} K(\tau+u,\tau+r)\mathcal{D}(u,r)\left(e^{i\omega\tau}\hat{K}^\top(\omega)\right)\mathrm{d}\tau\mathrm{d}r\mathrm{d}u$$

We look at $\hat{K}(-\omega)e^{-i\omega\tau}$ with the change of variables $\phi(t) = t + \tau$ and the Chasles relation:

$$\hat{K}(-\omega)e^{-i\omega\tau} = \int_0^{+\infty}\int_{-\infty}^{+\infty} e^{i\omega t}K(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau = \int_0^{+\infty}\int_{-\infty}^{-\tau} e^{i\omega t}K(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau \tag{A.7}$$

$$+ \int_0^{+\infty}\int_{-\tau}^{0} e^{i\omega t}K(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau \tag{A.8}$$

$$+ \int_0^{+\infty}\int_0^{+\infty} e^{i\omega t}K(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau \tag{A.9}$$

The first term, (A.7), $\int_0^{+\infty}\int_{-\infty}^{-\tau} e^{i\omega t}K(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau$ is null, because $K(t+\tau)$ is null for $t \leq -\tau$. If we switch the integration order in the second term, (A.8), we have $\int_0^{+\infty}\int_{-\tau}^{0} e^{i\omega t}K(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau = \int_{-\infty}^{0}\int_{-t}^{+\infty} e^{i\omega t}K(t+\tau,t+\tau)\mathrm{d}\tau\mathrm{d}t$ and then, we need to have a convergent exponential $e^{i\omega t}$ when $t \to -\infty$, so the second term is regular in half plan $\mathrm{Im}(\omega) < 0$. And the last term, (A.9), $\int_0^{+\infty}\int_0^{+\infty} e^{i\omega t}K(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau$ is regular in half plan $\mathrm{Im}(\omega) > 0$.

We now need to consider the term $e^{i\omega\tau}\hat{K}^\top(\omega)$, similarly, we have:

$$e^{i\omega\tau}\hat{K}^\top(\omega) = \int_0^{+\infty}\int_{-\infty}^{+\infty} e^{-i\omega t}K^\top(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau = \int_0^{+\infty}\int_{-\infty}^{-\tau} e^{-i\omega t}K^\top(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau \tag{A.10}$$

$$+ \int_0^{+\infty}\int_{-\tau}^{0} e^{-i\omega t}K^\top(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau \tag{A.11}$$

$$+ \int_0^{+\infty}\int_0^{+\infty} e^{-i\omega t}K^\top(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau \tag{A.12}$$

For same reasons as before, first term, (A.10), is null, second term, (A.11), is regular in half plan $\mathrm{Im}(\omega) > 0$, and last term, (A.12), is regular in half plan $\mathrm{Im}(\omega) < 0$.

Wrapping up those results, we have

$$\hat{K}(-\omega)\hat{C}_{33}^\top(\omega) = 2\int_0^{+\infty}\int_{u+}^{+\infty}\int_0^{+\infty}\int_{-\tau}^{0} e^{i\omega t}K(t+\tau,t+\tau)\mathrm{d}t\mathcal{D}^\top(u,r)K^\top(\tau+u,\tau+r)\mathrm{d}\tau\mathrm{d}r\mathrm{d}u \qquad \Big\}F_1(\omega)$$

$$+ 2\int_0^{+\infty}\int_{u+}^{+\infty}\int_0^{+\infty}\int_0^{+\infty} e^{i\omega t}K(t+\tau,t+\tau)\mathrm{d}t\mathcal{D}^\top(u,r)K^\top(\tau+u,\tau+r)\mathrm{d}\tau\mathrm{d}r\mathrm{d}u \qquad \Big\}F_2(\omega)$$

$$= F_1(\omega) + F_2(\omega)$$

with $F_1(\omega)$ being regular for $\text{Im}(\omega) < 0$ and $F_2(\omega)$ for $\text{Im}(\omega) > 0$.

Similarly,

$$
\begin{aligned}
\hat{C_{3_3}}(-\omega)\hat{K}^\top(\omega) =& 2\int_0^{+\infty}\int_{u+}^{+\infty}\int_0^{+\infty} K(\tau+u,\tau+r)\mathcal{D}(u,r)\int_{-\tau}^0 e^{-i\omega t}K^\top(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau\mathrm{d}r\mathrm{d}u \qquad \Big\} E_1(-\omega) \\
&+ 2\int_0^{+\infty}\int_{u+}^{+\infty}\int_0^{+\infty} K(\tau+u,\tau+r)\mathcal{D}(u,r)\int_0^{+\infty} e^{-i\omega t}K^\top(t+\tau,t+\tau)\mathrm{d}t\mathrm{d}\tau\mathrm{d}r\mathrm{d}u \qquad \Big\} E_2(-\omega) \\
=& E_1(-\omega) + E_2(-\omega)
\end{aligned}
$$

with $E_1(-\omega)$ being regular for $\text{Im}(\omega) > 0$ and $E_2(-\omega)$ for $\text{Im}(\omega) < 0$.

Step 3: Wrap up

We can now go back to Eq. (A.6) which becomes:

$$
\begin{aligned}
D\hat{K}^\top(\omega) + \hat{C_{3_3}}^\top(\omega) - \hat{C_{3_2}}(-\omega) + \hat{B}(-\omega) + \hat{C_{3_2}}(-\omega)\hat{K}^\top(\omega) - \hat{B}(-\omega)\hat{K}^\top(\omega) - F_1(\omega) + E_2(-\omega) = \\
\hat{K}(-\omega)D + \hat{C_{3_3}}(-\omega) - \hat{C_{3_2}}^\top(\omega) + \hat{B}^\top(\omega) + \hat{K}(-\omega)\hat{C_{3_2}}^\top(\omega) - \hat{K}(-\omega)\hat{B}^\top(\omega) - E_1(-\omega) + F_2(\omega)
\end{aligned}
$$

We then have the left side regular in half plan $\text{Im}(\omega) < 0$ and the right one in half plan $\text{Im}(\omega) > 0$. So as in [23], we can say that the lower side is null, and obtain an expression for $\hat{B}$:

$$
0 = \hat{K}(-\omega)D + \hat{C_{3_3}}(-\omega) - \hat{C_{3_2}}^\top(\omega) + \hat{B}^\top(\omega) + \hat{K}(-\omega)\hat{C_{3_2}}^\top(\omega) - \hat{K}(-\omega)\hat{B}^\top(\omega) - E_1(-\omega) + F_2(\omega)
$$

$$
\implies \begin{cases} \hat{B}^\top(\omega) &= -\Big(\hat{K}(-\omega)D + \hat{C_{3_3}}(-\omega) - \hat{C_{3_2}}^\top(\omega) + \hat{K}(-\omega)\hat{C_{3_2}}^\top(\omega) - E_1(-\omega) + F_2(\omega)\Big)^\top (1 - \hat{K}^\top(-\omega))^{-1} \\ \hat{B}(\omega) &= -\Big(D\hat{K}^\top(-\omega) + \hat{C}_{3_3}^\top(-\omega) - \hat{C}_{3_2}(\omega) + \hat{C}_{3_2}(\omega)\hat{K}^\top(-\omega) - E_1^\top(-\omega) + F_2^\top(\omega)\Big)(1 - \hat{K}^\top(-\omega))^{-1} \end{cases}
$$

Then we inject the expression of $\hat{B}$ in $\hat{\mathcal{C}}_*$, and after some manipulations where we use $\hat{C_{3_3}}(\omega)\hat{K}^\top(-\omega) = E_1(\omega) + E_2(\omega)$:

$$
\begin{aligned}
\hat{\mathcal{C}}(\omega) =& (1 - \hat{K}(\omega))^{-1}\Big(\hat{K}(\omega)D + \hat{C_{3_3}}(\omega) - \hat{K}(\omega)D\hat{K}^\top(-\omega) - E_1(\omega) - E_2(\omega) \\
& + D\hat{K}^\top(-\omega) + \hat{C}_{3_3}^\top(-\omega) - E_1^\top(-\omega) + F_2^\top(\omega)\Big)(1 - \hat{K}^\top(-\omega))^{-1}.
\end{aligned}
$$

Note that in 1D, the transpose and commutative operations do not really matter. However, in order to use this method in the multivariate case we presented the method paying attention to it.

In 1D, we have $F_1(\omega) = E_1(\omega)$ and $F_2(\omega) = E_2(\omega)$, the transpose are equal the element itself, and the products are commutative, thus:

$$
\hat{\mathcal{C}}(\omega) = (1 - \hat{K}(\omega))^{-1}\Big(\hat{K}(\omega)D + \hat{C_{3_3}}(\omega) - \hat{K}(\omega)D\hat{K}(-\omega) - E_1(\omega) + \hat{K}(-\omega)D + \hat{C_{3_3}}(-\omega) - E_1(-\omega)\Big)(1 - \hat{K}(-\omega))^{-1}.
$$

Finally, using Eq. (A.3), we obtain for QHawkes in 1D:

$$
\hat{\mathcal{C}}^\star(\omega) = (1 - \hat{K}(\omega))^{-1}\Big(D + \hat{C_{3_3}}(\omega) + \hat{C_{3_3}}(-\omega) - E_1(\omega) - E_1(-\omega)\Big)(1 - \hat{K}(-\omega))^{-1}.
$$

For the multivariate case the matrices we integrate in C_{3_3}, E_1, E_2, F_1 and F_2, will be a mix of $\mathbb{D}_\mathrm{d}$, $\mathbb{D}_\times$, $\mathbb{K}_\mathrm{d}$ and $\mathbf{K}_\times$.

6. Asymptotic Study- Results Tables

Based on the notations in Appendix 5 a, we give here the results of the asymptotic behavior of covariance structures when considering power law decaying kernels.

In both non critical and critical case, we have the same result for the exponent of non time diagonal parts $\delta_\times = \rho_\times + \delta_o - \frac{1}{2}$ and $\delta_o = \rho_o$. The exponent $\delta_\mathrm{d} = \rho_\mathrm{d}$ can take 2 values:

1. $\delta_\mathrm{d} = \rho_\mathrm{d}$
2. $\delta_\mathrm{d} = 2\rho_\times + \delta_o - \frac{1}{2}$

From empirical observations, we see that it makes sense if the diagonal terms persist longer in time than the non diagonal terms. Hence, we expect $\delta_\mathrm{d} < \delta_o$. Considering this hypothesis, we would only keep the first case, $\delta_\mathrm{d} = \rho_\mathrm{d} < \rho_o$.

For the time diagonal exponent $(\beta_\mathrm{d}, \beta_o, \beta_\mathrm{d}^\times, \beta_o^\times)$, we need to make a difference between the non critical case $(n_H < 1)$, and the critical case $(n_H = 1)$. We call $\epsilon = \min(\epsilon_\mathrm{d}, \epsilon_o)$ and $\rho = \frac{1}{2}\min(\delta_\mathrm{d} + \rho_\mathrm{d}, \delta_o + \rho_o, \delta_\mathrm{d} + \rho_o, \delta_o + \rho_\mathrm{d}, \delta_\times + \rho_\times) = \frac{1}{2}\min(\rho_\mathrm{d} + \rho_o, 2\rho_o, 2\rho_2, \delta_\times + \rho_\times)$.

Non Critical		
	if $\frac{3+\epsilon_{\mathrm d}}{4} < \min(\rho_{\mathrm d}, \rho_{\times} + \frac{\rho_{\mathrm o}}{2} + \frac{1}{4})$	if $\frac{3+\epsilon_{\mathrm d}}{4} > \min(\rho_{\mathrm d}, \rho_{\times} + \frac{\rho_{\mathrm o}}{2} + \frac{1}{4})$
$\beta_{\mathrm d}$	$\beta_{\mathrm d} = 1 + \epsilon_{\mathrm d}$	$\beta_{\mathrm d} = \min(4\rho_{\times} + 2\rho_{\mathrm o} - 1, 4\rho_{\mathrm d} - 2)$
	if $\frac{3+\epsilon_{\mathrm o}}{4} < \min(\frac{\rho_{\mathrm d}+\rho_{\mathrm o}}{2}, \rho_{\times} + \frac{\rho_{\mathrm o}}{2} + \frac{1}{4})$	if $\frac{3+\epsilon_{\mathrm o}}{4} > \min(\frac{\rho_{\mathrm d}+\rho_{\mathrm o}}{2}, \rho_{\times} + \frac{\rho_{\mathrm o}}{2} + \frac{1}{4})$
$\beta_{\mathrm o}$	$\beta_{\mathrm o} = 1 + \epsilon_{\mathrm o}$	$\beta_{\mathrm o} = \min(4\rho_{\times} + 2\rho_{\mathrm o} - 1, 2(\rho_{\mathrm d} + \rho_{\mathrm o}) - 2)$
Critical		
	if $\rho < \frac{3}{2}$	if $\rho > \frac{3}{2}$
$\beta_{\mathrm d} = \beta_{\mathrm o} = \beta$	$\beta = 4\rho - 2 - 2\epsilon$	$\beta = 1 - 2\epsilon$

Table I: $\mathbb{C}$ exponents for both non critical and critical case, when using the following notations $\epsilon = \min(\epsilon_{\mathrm d}, \epsilon_{\mathrm o})$ and $\rho = \frac{1}{2}\min(\delta_{\mathrm d} + \rho_{\mathrm d}, \delta_{\mathrm o} + \rho_{\mathrm o}, \delta_{\mathrm d} + \rho_{\mathrm o}, \delta_{\mathrm o} + \rho_{\mathrm d}, \delta_{\times} + \rho_{\times})$

Non Critical		
	if $\rho_{\times} < 2\rho_{\mathrm d} - \frac{1}{2}$	if $\rho_{\times} > 2\rho_{\mathrm d} - \frac{1}{2}$
$\beta_{\mathrm d}^{\times}$	$\beta_{\mathrm d}^{\times} = 2\rho_{\times} + \rho_{\mathrm o} - 1$	$\beta_{\mathrm d}^{\times} = 2\rho_{\mathrm d} + \delta_{\times} - 1$
	if $\rho_{\times} < 2\rho_{\mathrm o} - \frac{1}{2}$	if $\rho_{\times} > 2\rho_{\mathrm d} - \frac{1}{2}$
$\beta_{\mathrm o}^{\times}$	$\beta_{\mathrm o}^{\times} = 2\rho_{\times} + \rho_{\mathrm o} - 1$	$\beta_{\mathrm o}^{\times} = 2\rho_{\mathrm o} + \delta_{\times} - 1$
Critical		
	if $\frac{2\rho_{\times}+\delta_o}{3} > \frac{2}{3}$ and $\frac{2\rho_d+\delta_{\times}}{3} > \frac{2}{3}$	else
$\beta_{\mathrm d}^{\times}$	$\beta_{\mathrm d}^{\times} = 1 - \epsilon$	$\beta_{\mathrm d}^{\times} = \min(2\rho_{\mathrm d} + \delta_{\times} - \epsilon, 2\rho_{\times} + \delta_{\mathrm o} - \epsilon)$
	if $\frac{2\rho_{\times}+\delta_o}{3} > \frac{2}{3}$ and $\frac{2\rho_o+\delta_{\times}}{3} > \frac{2}{3}$	else
$\beta_{\mathrm o}^{\times}$	$\beta_{\mathrm o}^{\times} = 1 - \epsilon$	$\beta_{\mathrm o}^{\times} = \min(2\rho_{\mathrm o} + \delta_{\times} - \epsilon, 2\rho_{\times} + \delta_{\mathrm o} - \epsilon)$

Table II: $\mathbb{D}_{\times}$ "time diagonal" exponents for both non critical and critical case, when using the following notations $\epsilon = \min(\epsilon_{\mathrm d}, \epsilon_{\mathrm o})$ and $\rho = \frac{1}{2}\min(\delta_{\mathrm d} + \rho_{\mathrm d}, \delta_{\mathrm o} + \rho_{\mathrm o}, \delta_{\mathrm d} + \rho_{\mathrm o}, \delta_{\mathrm o} + \rho_{\mathrm d}, \delta_{\times} + \rho_{\times})$

7. Computation of the Infinitesimal Generator

We consider the processes, introducing a time scale T > 0 that will eventually diverge, we define the processes $\bar{h}_1^T(t) = h_1(tT)$, $\bar{h}_2^T(t) = h_2(tT)$, $\bar{z}_1^T(t) = z_1(tT)$ and $\bar{z}_2^T(t) = z_2(tT)$, $\bar{N}_t^{1T} = N_{tT}^1$, $\bar{N}_t^{2T} = N_{tT}^2$ with the parameters (β_{1T}, β_{2T}) and $(\omega_{1T}, \omega_{2T})$ that may depend on T. In the decaying exponential case, we have

$$
\begin{cases}
d(h_1)_t &= \beta_1(-(h_1)_t dt + dN_t^1) \\
d(h_2)_t &= \beta_2(-(h_2)_t dt + dN_t^2) \\
d(z_1)_t &= -\omega_1(z_1)_t dt + \omega_1 dP_t^1 \\
d(z_2)_t &= -\omega_2(z_2)_t dt + \omega_2 dP_t^2
\end{cases}
\implies
\begin{cases}
d(\bar{h}_1^T)_t &= -\beta_{1T}((\bar{h}_1^T)_t T dt + n_H dN_t^{1T}) \\
d(\bar{h}_2^T)_t &= -\beta_{2T}((\bar{h}_2^T)_t T dt + n_H dN_t^{2T}) \\
d(\bar{z}_1^T)_t &= -\omega_{1T}(\bar{z}_1^T)_t T dt + \omega_{1T} dP_t^{1T} \\
d(\bar{z}_2^T)_t &= -\omega_{2T}(\bar{z}_2^T)_t T dt + \omega_{2T} dP_t^{2T}
\end{cases}
$$

The intensity of, respectively, the first and second process are given by $T[\lambda_{\infty}^1 + n_{H,1}^1 \bar{h}_1^T + n_{H,2}^1 \bar{h}_2^T + (a_{Z,1}^1 \bar{z}_1^T)^2 + (a_{Z,2}^1 \bar{z}_2^T)^2 + a_{Z,\times}^1 \bar{z}_1^T \bar{z}_2^T]$ and $T[\lambda_{\infty}^2 + n_{H,1}^2 \bar{h}_1^T + n_{H,2}^2 \bar{h}_2^T + (a_{Z,1}^2 \bar{z}_1^T)^2 + (a_{Z,2}^2 \bar{z}_2^T)^2 + a_{Z,\times}^2 \bar{z}_1^T \bar{z}_2^T]$. Moreover, the price processes can either go up or down with same probability $\frac{1}{2}$.

Thus, the infinitesimal generator can be written using *Theorem 1.22 of Oksendal and Sulem*:

$$Af^T(h_1, h_2, z_1, z_2) = - T\beta_{1T}h_1\partial_{h_1}f - T\beta_{2T}h_2\partial_{h_2}f - \omega_{1T}Tz_1\partial_{z_1}f - \omega_{2T}Tz_2\partial_{z_2}f$$
$$+ T[\lambda_\infty^1 + n_{H,1}^1 h_1 + n_{H,2}^1 h_2 + (a_{Z,1}^1 z_1)^2 + (a_{Z,2}^1 z_2)^2 + a_{Z,\times}^1 z_2 z_1]\Big($$
$$\frac{1}{2}f(h_1 + \beta_{T1}, h_2, z_1 + \omega_{1T}, z_2) + \frac{1}{2}f(h_1 + \beta_{T1}, h_2, z_1 - \omega_{1T}, z_2) - f(h_1, h_2, z_1, z_2)\Big)$$
$$+ T[\lambda_\infty^2 + n_{H,1}^2 h_1 + n_{H,2}^2 h_2 + (a_{Z,1}^2 z_1)^2 + (a_{Z,2}^2 z_2)^2 + a_{Z,\times}^2 z_2 z_1]\Big($$
$$\frac{1}{2}f(h_1, h_2 + \beta_{T2}, z_1, z_2 - \omega_{2T}) + \frac{1}{2}f(h_1, h_2 + \beta_{T2}, z_1, z_2 + \omega_{2T}) - f(h_1, h_2, z_1, z_2)\Big)$$

As in [5], we use the scaling $\beta_{1T} = \frac{\bar\beta_1}{T}$, $\beta_{2T} = \frac{\bar\beta_2}{T}$, $\omega_{1T} = \frac{\bar\omega_1}{T}$, $\omega_{2T} = \frac{\bar\omega_2}{T}$ and the Taylor development when $T \to +\infty$ which result in:

$$Af^T(h_1, h_2, z_1, z_2) = - \bar\beta_1 h_1\partial_{h_1}f - \bar\beta_2 h_2\partial_{h_2}f - \bar\omega_1 z_1\partial_{z_1}f - \bar\omega_2 z_2\partial_{z_2}f$$
$$+ T[\lambda_\infty^1 + n_{H,1}^1 h_1 + n_{H,2}^1 h_2 + (a_{Z,1}^1 z_1)^2 + (a_{Z,2}^1 z_2)^2 + a_{Z,\times}^1 z_2 z_1]\Big(\frac{\bar\beta_1}{T}\partial_{h_1}f + \frac{(\bar\omega_1)^2}{2T}\partial_{z_1 z_1}^2 f + o(\frac{1}{T})\Big)$$
$$+ T[\lambda_\infty^2 + n_{H,1}^2 h_1 + n_{H,2}^2 h_2 + (a_{Z,1}^2 z_1)^2 + (a_{Z,2}^2 z_2)^2 + a_{Z,\times}^2 z_2 z_1]\Big(\frac{\bar\beta_2}{T}\partial_{h_2}f + \frac{(\bar\omega_2)^2}{2T}\partial_{z_2 z_2}^2 f + o(\frac{1}{T})\Big)$$

Thus, the infinitesimal generator is given by:

$$Af^\infty(h_1, h_2, z_1, z_2) = - \bar\beta_1 h_1\partial_{h_1}f - \bar\beta_2 h_2\partial_{h_2}f - \bar\omega_1 z_1\partial_{z_1}f - \bar\omega_2 z_2\partial_{z_2}f$$
$$+ [\lambda_\infty^1 + n_{H,1}^1 h_1 + n_{H,2}^1 h_2 + (a_{Z,1}^1 z_1)^2 + (a_{Z,2}^1 z_2)^2 + a_{Z,\times}^1 z_2 z_1]\Big(n_H\bar\beta_1\partial_{h_1}f + \frac{(\bar\omega_1)^2}{2}\partial_{z_1 z_1}^2 f\Big)$$
$$+ [\lambda_\infty^2 + n_{H,1}^2 h_1 + n_{H,2}^2 h_2 + (a_{Z,1}^2 z_1)^2 + (a_{Z,2}^2 z_2)^2 + a_{Z,\times}^2 z_2 z_1]\Big(n_H\bar\beta_2\partial_{h_2}f + \frac{(\bar\omega_2)^2}{2}\partial_{z_2 z_2}^2 f\Big)$$

One can also include co-jumps, which would change the coefficients above and add a term of the form $\rho\partial_{z_1 z_2}^2 f$, where ρ is the correlation between the Poisson processes driving 1 and 2.

When considering only ZHawkes without Hawkes ($n_H = 0$), we have $h_1 = h_2 = 0$, thus the infinitesimal generator results in:

$$Af^\infty(z_1, z_2) = - \bar\omega_1 z_1\partial_{z_1}f - \bar\omega_2 z_2\partial_{z_2}f$$
$$+ [\lambda_\infty^1 + (a_{Z,1}^1 z_1)^2 + (a_{Z,2}^1 z_2)^2 + a_{Z,\times}^1 z_2 z_1]\frac{(\bar\omega_1)^2}{2}\partial_{z_1 z_1}^2 f + [\lambda_\infty^2 + (a_{Z,1}^2 z_1)^2 + (a_{Z,2}^2 z_2)^2 + a_{Z,\times}^2 z_2 z_1]\frac{(\bar\omega_2)^2}{2}\partial_{z_2 z_2}^2 f$$

8.   General ODE for $F(\theta)$

We write here the general ODE on $F(\theta)$ for any parameters a_Z's and ω's:

$$\Big(\omega_1 + \omega_2 + (a_{Z,1}^1 \omega_1)^2 + (a_{Z,2}^2 \omega_2)^2 - \alpha(\omega_1 + 2(a_{Z,1}^1 \omega_1)^2)\cos^2(\theta) - \alpha(\omega_2 + 2(a_{Z,2}^2 \omega_2)^2)\sin^2(\theta)$$
$$- (a_{Z,\times}^2 (\omega_2)^2 + a_{Z,\times}^1 (\omega_1)^2)\alpha\cos\theta\sin\theta$$
$$+ \frac{1}{8}\alpha(\alpha + (2 + \alpha)\cos(2\theta))(\omega_1)^2 \big((a_{Z,1}^1)^2 + (a_{Z,2}^1)^2 + ((a_{Z,1}^1)^2 - (a_{Z,2}^1)^2)\cos(2\theta) + a_{Z,\times}^1 \sin(2\theta)\big)$$
$$+ \frac{1}{8}\alpha(\alpha - (2 + \alpha)\cos(2\theta))(\omega_2)^2 \big((a_{Z,1}^2)^2 + (a_{Z,2}^2)^2 + ((a_{Z,1}^2)^2 - (a_{Z,2}^2)^2)\cos(2\theta) + a_{Z,\times}^2 \sin(2\theta)\big)\Big)F(\theta)$$
$$+ \Big(a_{Z,\times}^2 (\omega_2)^2 \cos^2(\theta) - a_{Z,\times}^1 (\omega_1)^2 \sin^2(\theta) - \cos\theta\sin\theta\big(\omega_1 - \omega_2 + 2(a_{Z,1}^1 \omega_1)^2 - 2(a_{Z,2}^2 \omega_2)^2\big) \tag{A.13}$$
$$+ \frac{1}{4}(1 + \alpha)(\omega_1)^2 \sin(2\theta)((a_{Z,1}^1)^2 + (a_{Z,2}^1)^2 + ((a_{Z,1}^1)^2 - (a_{Z,2}^1)^2)\cos(2\theta) + a_{Z,\times}^1 \sin(2\theta))$$
$$- \frac{1}{4}(1 + \alpha)(\omega_2)^2 \sin(2\theta)((a_{Z,1}^2)^2 + (a_{Z,2}^2)^2 + ((a_{Z,1}^2)^2 - (a_{Z,2}^2)^2)\cos(2\theta) + a_{Z,\times}^2 \sin(2\theta))\Big)F'(\theta)$$
$$+ \frac{1}{4}\Big((\omega_1)^2 \sin^2(\theta)((a_{Z,1}^1)^2 + (a_{Z,2}^1)^2 + ((a_{Z,1}^1)^2 - (a_{Z,2}^1)^2)\cos(2\theta) + a_{Z,\times}^1 \sin(2\theta))$$
$$+ (\omega_2)^2 \cos^2(\theta)((a_{Z,1}^2)^2 + (a_{Z,2}^2)^2 + ((a_{Z,1}^2)^2 - (a_{Z,2}^2)^2)\cos(2\theta) + a_{Z,\times}^2 \sin(2\theta))\Big)F''(\theta) = 0$$