

Multivariate Quadratic Hawkes Processes – Part II: Non-Parametric Empirical Calibration

Cecilia Aubrun,^{1, 2, 3} 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: September 26, 2025)

This is the second part of our work on Multivariate Quadratic Hawkes (MQHawkes) Processes, devoted to the calibration of the model defined and studied analytically in Aubrun, C., Benzaquen, M., & Bouchaud, J. P., Quantitative Finance, 23(5), 741–758 (2023). We propose a non-parametric calibration method based on the general method of moments applied to a coarse-grained version of the MQHawkes model. This allows us to bypass challenges inherent to tick by tick data. Our main methodological innovation is a multi-step calibration procedure, first focusing on “self” feedback kernels, and then progressively including cross-effects. Indeed, while cross-effects are significant and interpretable, they are usually one order of magnitude smaller than self-effects, and must therefore be disentangled from noise with care. For numerical stability, we also restrict to pair interactions and only calibrate bi-variate QHawkes, neglecting higher-order interactions. Our main findings are: (a) While cross-Hawkes feedback effects have been empirically studied previously, cross-Zumbach effects are clearly identified here for the first time. The effect of recent trends of the E-MINI futures contract onto the volatility of other futures contracts is especially strong; (b) We have identified a new type of feedback that couples past realized covariance between two assets and future volatility of these two assets, with the pair E-MINI/TBOND as a case in point; (c) A cross-leverage effect, whereby the sign of the return of one asset impacts the volatility of another asset, is also clearly identified. The cross-leverage effect between the E-MINI and the residual volatility of single stocks is notable, and surprisingly universal across the universe of stocks that we considered.

Keywords: Multivariate QHawkes, QGARCH, non-parametric calibration

I. INTRODUCTION

The increasing amount of high-frequency data in financial markets has made it possible to study and calibrate microstructure models. An increasingly renowned family of such models has been invented by the late Alan G. Hawkes [1, 2]. Initially introduced to model seismic activity, the adaptability and interpretability of Hawkes processes have made them highly appealing for financial data [3, 4]. They allow in particular to identify the feedback loop that underlies some of the most prominent stylized facts that characterize financial prices.

In a previous paper [5], we introduced a Multivariate version of the Quadratic Hawkes (QHawkes) process proposed by Blanc et al. [6]. However, the calibration of MQHawkes processes on empirical data is far from trivial and poses specific difficulties that the present work aims at resolving.

Many methods exist to calibrate Hawkes processes on empirical data, using either events binned in a regular time grid, or successive time events. A good review is presented in Appendix C of [7]. Methods fall into two

categories: parametric and non-parametric, referring to the way the feedback “kernel” is specified.

Among the parametric approaches frequently employed, the maximum likelihood [8, 9] and Expectation Maximisation (EM) method [10] stand out as prominent examples. Non-parametric methods offer the advantage of agnostic kernel forms, as they do not impose specific shapes. The primary non-parametric approaches include the Expectation Maximisation-based method [11, 12], the minimisation of contrast function [13, 14], and the method of moments or Wiener-Hopf-based method promoted by Bacry et al. [4, 15–18]. Additionally, the *method of cumulants*, a non-parametric method [19, 20], allows one to assess the average endogeneity of the process without having to fully estimate the feedback kernels.

In the present study we want to remain agnostic regarding the kernel shape and implement the method of moments on binned data for several pairs of stocks and futures (henceforth called “futures”). This can later be used as an initial condition for maximum likelihood estimation. Our main methodological innovation is to perform the calibration of the MQHawkes in successive steps, first focusing on “self” (univariate) feedback kernels, and then progressively including cross-effects. Such a procedure is needed for the calibration to con-

* michael.benzaquen@polytechnique.edu

verge towards stable and meaningful results. Indeed, while cross-effects are significant and interpretable, they are usually one order of magnitude smaller than self-effects, and must therefore be disentangled from noise with care. We also separate quadratic effects (invariant when the sign of returns are flipped) from linear effects (that change sign when sign of returns are flipped), again to improve identifiability.

For the sake of numerical robustness, we restrict to bi-variate calibrations and think of a full multivariate calibration of the MQHawkes model as resulting from the combination of all pairwise calibrations. Along the way, we use the pair E-MINI vs. TBOND as a canonical test case for our method. We also simplify our model to deal with single stocks, by only considering interactions between the market factor (i.e. the E-MINI in our case) and the individual residual component of each stock separately, within the context of a one factor model.

The outline of the paper is as follows. Section II recalls the definition of the MQHawkes model we aim to calibrate. Section III describes the data and the pre-processing methods we apply to it. Section IV presents the key ingredients for the calibration: a discrete-time (bin) approximation, the set of covariance functions and the calibration method. Finally, the last sections are dedicated to our calibration results on empirical data (futures and single stocks) and their interpretation. Calibration on synthetic data, as a proof of concept, is given in Appendix A. Many technical details can be found in other Appendices.

The main takeaways of our study are the following:

- The two well-studied “self” feedback effects, namely the Hawkes feedback (i.e. activity \rightarrow activity [3, 7]) and the Zumbach feedback (i.e. squared trend \rightarrow activity [6]) have “cross” counterparts. While the cross-Hawkes terms were empirically studied previously (see e.g. [7, 21]) cross-Zumbach effects are clearly identified here for the first time (see also [5]). The effect of recent trends of the E-MINI futures contract onto the volatility of other futures contracts is especially strong.
- We have identified a new type of feedback that couples past realized *covariance* between two assets and future *volatility* of these two assets, with the pair E-MINI-TBOND serving as a striking illustrative example.
- A cross-leverage effect, whereby the sign of the return of one asset impacts the volatility of another asset, is also clearly identified. Once again, the effect is particularly strong in the case of the E-MINI on other futures, although the sign of the returns of the TBOND also plays a significant role. The cross-leverage effect between the E-MINI and

the residual volatility of single stocks is also noteworthy, and appears to be remarkably universal across the set of stocks we examined (see Fig. 9 below).

II. MQHAWKES & MQARCH MODELS

In this section we recall essential ingredients of the original (point process) MQHawkes model presented in [5]. We then formulate a corresponding multivariate quadratic ARCH model (MQARCH) that can be seen as a coarse-grained version of the MQHawkes model, see [22] and below. Such a formulation is better adapted to the 1-minute binned data at our disposal (see below) and could be used as a starting point for a more fine-grained calibration using tick by tick multivariate data.

A. MQHawkes

Let $(N_t)_{t \geq 0}$ be a Hawkes process, that is an inhomogenous Poisson process with intensity λ_t defined from the past realisations of the process itself, given by:

$$\lambda_t = \lambda_\infty + \int_{-\infty}^t \phi(t-u) dN_u \quad (1)$$

with

$$n = \int_{\mathbb{R}} \phi(u) du < 1. \quad (2)$$

In this context, λ_∞ denotes the baseline intensity, ϕ is the feedback kernel and n represents the “endogeneity ratio”, which must be strictly less than one to ensure that the process reaches a stationary state.

QHawkes processes, akin to Hawkes processes, are inhomogeneous Poisson processes reliant on past occurrences for their intensity calculations. In the context of financial markets, the event process $(N_t)_{t > 0}$ represents the sequence of price changing events, where the price change (in tick units) is defined by $dP_t = \epsilon_t dN_t$ [6] where $\epsilon_t = \pm 1$ is a random sign. The intensity self-exciting feedback loop is then defined on the return stream dP_t rather than on the events dN_t . To wit, the sign of the returns influences the process intensity as [6]:

$$\begin{aligned} \lambda_t = & \lambda_\infty + \int_{-\infty}^t L(t-s) dP_s \\ & + \iint_{-\infty}^t K(t-s, t-u) dP_s dP_u, \end{aligned} \quad (3)$$

where λ_∞ is again the baseline intensity, $L(\cdot)$ is the so-called leverage kernel (capturing the increase of activity following a drop in prices) and $K(\cdot, \cdot)$ is the quadratic Hawkes kernel, with a time-diagonal part that boils down to the standard Hawkes feedback and a non-diagonal part that captures the path-dependency of the activity on past price changes [22–24].

In a multivariate framework, that is encompassing N financial assets with prices $(P_t^i)_{t \geq 0, i=1,\dots,N}$, we associate such price processes with jump processes $(N_t^i)_{t \geq 0}$, with:

$$dP_t^i = \epsilon_i dN_t^i, \quad (4)$$

where each jump process $(dN_t^i)_{t \geq 0}$ is a conditionally independent¹ Quadratic Hawkes process with intensity $(\lambda_{i,t})_{t \geq 0}$ such that:

$$\begin{aligned} \lambda_{i,t} = & \lambda_{i,\infty} + \sum_{j=1}^N \int_{-\infty}^t L_j^i(t-s) dP_s^j \\ & + \sum_{j \leq k}^N \iint_{-\infty}^t K_{jk}^i(t-s, t-u) dP_s^j dP_u^k. \end{aligned} \quad (5)$$

In this notation, the superscript in the kernels indicates which asset is affected by the feedback, while the subscripts identify the assets responsible for it.

We will generically refer to the kernels K_{jk}^i with $j = k$ as the “diagonal” feedback terms. These components describe the quadratic feedback from two price changes of the same asset j onto the activity of asset i . Similarly, K_{jk}^i with $j < k$ outlines cross-effects from two different assets j, k onto asset i ’s activity. As we argued in [5], one can neglect the case where the three indices i, j, k are all different, since this would correspond to scenario where the instantaneous return correlation between two assets impacts the activity on a third one. Although not impossible a priori, we will not consider this possibility, which opens the path to a complete calibration of an N -dimensional MQHawkes from $N(N-1)/2$ pairwise calibrations.

Finally, we distinguish the “time-diagonal” of a kernel $K_{jk}(\tau_1, \tau_2)$, by which we mean $K_{jk}(\tau, \tau)$, from its “diagonal”, which refers to the diagonal components in asset space K_{ii} .

B. MQARCH

Here, we give the equations defining the multivariate, quadratic ARCH model that represents a coarse-grained version of the MQHawkes continuous point process. Specifically, relying on the work of Blanc *et al.* [6]

on the univariate QHawkes, the QHawkes intensity λ_t can be approximated by the squared volatility as follows:

$$\frac{\left(\sigma_t^{(\Delta t)}\right)^2}{\Delta t} \xrightarrow[\Delta t \rightarrow 0^+]{} \lambda_t, \quad (6)$$

where $\sigma_t^{(\Delta t)}$ is a volatility estimate over the time bin of size Δt .

Thus, the Hawkes continuous point process dN_t can be approximated by the discrete volatility process $\sigma^{(\Delta t)}$ over a time bin of size Δt . This approximation critically depends on the choice of the time bin size Δt , which in turn significantly influences the calibration of the underlying QHawkes model.

Indeed if $\Delta t \gg 1/\bar{\lambda}$, the QHawkes relation $\Delta P_t = \pm \Delta N_t$, is no longer valid. Consequently, some of the terms in the Yule-Walker equations (relating kernels to covariances [5]) would need to be modified. In particular, the relation between the diagonal contribution $K(s, s)\Delta P_{t-s}^2$ and the Hawkes contribution $\phi(s)\Delta N_{t-s}$ is lost.

Relying on the QARCH framework introduced in [22, 25, 26], we define the MQARCH model in the bi-variate case, for the time dependent volatility $\sigma_{i,t}^2$ of two assets $i, j \in (1, 2)$, as follows:

$$\begin{aligned} \sigma_{i,t}^2 = & \sigma_{i,\infty}^2 + \sum_{j=1}^2 \sum_{\tau=1}^{+\infty} L_j^i(\tau) r_{j,t-\tau} + \sum_{j=1}^2 \sum_{\tau=1}^{+\infty} \phi_j^i(\tau) r_{j,t-\tau}^2 \\ & + 2 \sum_{j=1}^2 \sum_{\tau_1=1}^{+\infty} \sum_{\tau_2=\tau_1+1}^{+\infty} K_j^i(\tau_1, \tau_2) r_{j,t-\tau_1} r_{j,t-\tau_2} \\ & + \sum_{\tau=1}^{+\infty} \phi_{\bar{i}}^i(\tau) (r_{i,t-\tau} r_{\bar{i},t-\tau} - C) \\ & + \sum_{\tau_1=1}^{+\infty} \sum_{\tau_2 \neq \tau_1}^{\infty} K_{\bar{i}}^i(\tau_1, \tau_2) r_{i,t-\tau_1} r_{\bar{i},t-\tau_2}, \end{aligned} \quad (7)$$

where we define $\bar{i} = 2$ if $i = 1$ and vice-versa, and $C = \mathbb{E}[r_{1,t} r_{2,t}]$ is the equal time covariance between the returns of the two assets 1, 2, and where all returns are de-drifted ($\mathbb{E}[r_{i,t}] = 0$).

In this framework, the returns $r_{i,t}$ are no longer defined at microscale but are related – as in all ARCH models – to the volatility through

$$r_{i,t} = \sigma_{i,t} \xi_{i,t}, \quad (8)$$

where $\xi_{i,t}$ are independent random variables of zero mean and variance unity, possibly non-Gaussian and non-identical for different assets.² The MQARCH

¹ This means that for given intensities $\lambda_{i,t}$, the inhomogeneous Poisson processes dN_t^i are independent.

² We assume here that all non-equal time covariances are zero, but see Section III B 2 below.

framework is more consistent with the data used in this paper (see Section III), which is based on binned 1-min returns.

Additionally, using the fact that $\mathbb{E}[\xi_{i,t}\xi_{j,t'}] = \delta_{ij}\delta_{tt'}$, the mean squared volatility is determined through the following relation:

$$\begin{pmatrix} \overline{\sigma_1^2} \\ \overline{\sigma_2^2} \end{pmatrix} = \begin{pmatrix} 1 - n_1^1 & n_1^2 \\ n_1^2 & 1 - n_2^2 \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{1,\infty}^2 \\ \sigma_{2,\infty}^2 \end{pmatrix}, \quad (9)$$

where

$$n_j^i := \sum_{\tau=1}^{+\infty} \phi_j^i(\tau). \quad (10)$$

Eq. (9) is of course only valid if the spectral radius of the feedback matrix (n_j^i) is less than unity. Note that this relation does not involve the quadratic kernels K nor the cross kernel $\phi_{\times}, K_{\times}$.

III. DATA

We consider a set of 6 futures contracts (E-MINI, E-MINI-3 (3 months futures), NASDAQ, DOW JONES, CRUDE OIL and TBOND) for every trading day from 2013 to 2023, and 317 single stocks belonging to the S&P500 during the same period.

A. Data description

For each studied asset i , we collect the opening (o), closing (c), highest (h) and lowest (l) prices for 1-minute intervals for every trading day. We only consider the bins between 10 am and 3 pm to exclude specially high volatility at market openings (as participants react to overnight news) and at market closings (as some participants close positions or complete the execution of their metaorders before the market session ends).

From these price time series, we compute the two quantities needed for the calibration for each 1-minute bin, the log-returns of asset i $r_{t,d}^i$ and the corresponding volatility $\sigma_{t,d}^i$, as follows:

$$r_{t,d}^i = \log \left(\frac{c_{t,d}^i}{o_{t,d}^i} \right) \quad (11)$$

$$\sigma_{t,d}^i = \frac{1}{3} \frac{h_{t,d}^i - l_{t,d}^i}{o_{t,d}^i} + \frac{2}{3} \frac{|c_{t,d}^i - o_{t,d}^i|}{o_{t,d}^i} \quad (12)$$

where d accounts for the day and t for the 1-minute bin within day d . For single stocks, a special treatment will be implemented, as we will decompose stock returns into factor and idiosyncratic components before extracting the volatility of each bin, see Section VI B.

B. Data pre-processing

After collecting the log-returns $r_{t,d}^i$ and volatilities $\sigma_{t,d}^i$ time series, we proceed with two steps of data pre-processing: normalisation and “martingalisation”. The first step removes any remaining intraday pattern of the volatility and standardises the volatility across the entire studied period. The “martingalisation” process allows us to remove any linear correlation between returns, imposing that $\mathbb{E}[r_t r_{t-\tau}] = 0$. We now detail the two steps in turn.

1. Normalisation

The objectives of these normalization procedures are twofold: first, to make the time series stationary over time, and second, to standardise the activity throughout the day.

To achieve stationarity and remove intraday seasonalities, each 1-minute bin return is normalized by an estimate of the realized volatility over the past 100 days as described in Eqs. (13) and (14):

$$r_{t,d}^i \leftarrow \frac{r_{t,d}^i}{\sqrt{\frac{1}{100} \sum_{k=d-100}^d (r_{t,k}^i)^2}}, \quad (13)$$

$$\sigma_{t,d}^i \leftarrow \frac{\sigma_{t,d}^i}{\sqrt{\frac{1}{100} \sum_{k=d-100}^d (\sigma_{t,k}^i)^2}}. \quad (14)$$

Without this normalization, the daily activity profile exhibits marked a U-shape or even a J-shape, as it is the case for the E-MINI and the TBOND as demonstrated in Figure 1, with low-frequency variations of the volatility due to the length of the time series.

Even after such a normalisation, there is still a residual intraday profile that we suppress by the following operation:

$$\sigma_{t,d}^i \leftarrow \frac{\sigma_{t,d}^i}{\sqrt{\langle (\sigma_{t,d}^i)^2 \rangle_d}}, \quad (15)$$

$$r_{t,d}^i \leftarrow \frac{r_{t,d}^i}{\sqrt{\langle (\sigma_{t,d}^i)^2 \rangle_d}}, \quad (16)$$

where $\langle \dots \rangle_d$ means an average over all days in the time series.

2. “Martingalisation”

Upon examining empirical data, it appears that at the 1-minute time scale, returns show negligible auto-

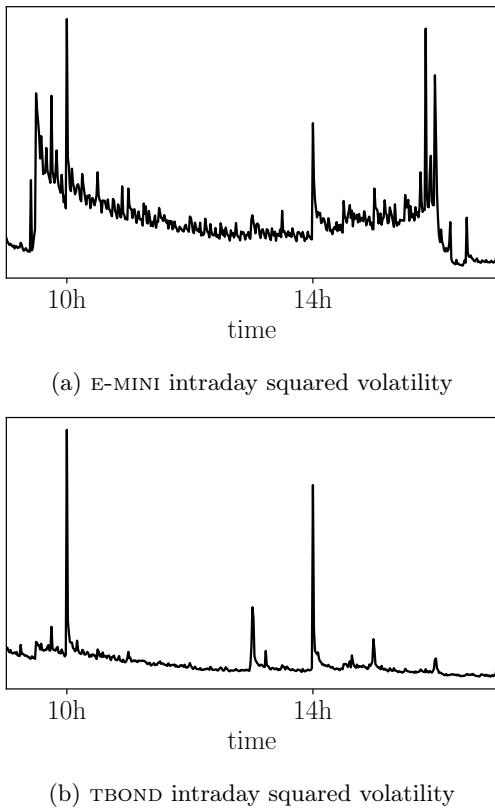


Figure 1: Average intraday profiles of the squared volatility Bachelier estimates, as defined by Equation (12), for a period spanning from 2013 to 2023.

correlations, with a notable non-zero correlation between r_t and r_{t-1} . Although this correlation is weak and only significant over short time scales, it is essential to address this weakly non-martingale behavior because in the QHawkes framework presented in [5], price returns are assumed to be martingales, and many properties rely on this assumption. Therefore, we perform by hand a “martingalisation” of the returns time series, meaning that the component of the returns at time t that is predictable from time $t-1$ is removed to isolate the return’s “surprise” component, as follows (see also [17] and Appendix B for more details):

$$r_{t,d}^i \leftarrow r_{t,d}^i - \mathbb{E}[r_{t,d}^i | \mathcal{F}_{t-1,d}]. \quad (17)$$

Such a procedure also removes any (small) average value of the returns. The data pre-processing is completed by normalising the returns one last time by their standard deviation.

The subsequent section describes the calibration method to be applied on the returns and volatility time series.

IV. CALIBRATION METHOD

This section outlines the calibration method we implemented. For the sake of clarity, we present here the main elements and Appendices C and D provide more comprehensive details.

To stay agnostic on the shape of the kernels, we chose to implement a method of moments. This method relies on linear relationships between the kernels we want to recover and quantities that we can compute from data, that are covariance structures. Both were introduced in [5] for the MQHawkes model (see Equations (16), (17), (20) and (23) in Ref. [5]).

The discretization of the data imposes these equations to be adapted; however, the same covariance structures are still a key ingredient to the calibration. We thus start by discussing the data binning and the modifications that follow.

A. Approximation for discrete data – toward the MQARCH model

As we just mentioned, the advantage of the method of moments is that it is non-parametric, unlike the Maximum Likelihood method for instance. The method of moments relies on the computation of covariance structures, such as those introduced in [5], in terms of the different feedback kernels, without having to specify their shape.

The estimation of these covariance structures can be done directly using the times of events, i.e. the times of price changes in the QHawkes model, which necessitates working with high frequency data. However, high frequency data is sometimes difficult to obtain and/or to process, due to the sheer quantity of data. Such data can also be heavily impacted by microstructure noise [27]. In particular, high frequency, particularly tick-by-tick, price changes reflect strong bid-ask bounce effects (strong mean reversion), which is not of interest in this study. Therefore, we choose to work with aggregated data, to which we apply the martingalisation process of Section III B 2 to effectively eliminate any residual bid-ask bounce effect.

The subsequent section introduces the “adapted” covariance structures needed to characterise the kernels of Equation (7), that are $(L_j^i)_{i,j \in \{1,2\}}$, $(\phi_j^i)_{i,j \in \{1,2\}}$, $(K_j^i)_{i,j \in \{1,2\}}$, $(\phi_x^i)_{i \in \{1,2\}}$ and $(K_x^i)_{i \in \{1,2\}}$ and baseline values $(\sigma_{i,\infty})_{i \in \{1,2\}}$ of Equation (7) which cannot be directly observed in data.

B. Covariance structures

1. Definitions

The method of moments relies on the fact that kernels can be deduced from covariance functions, which can easily be estimated from empirical data. Given the MQGARCH specification of the model delineated in Equation (7), we must consider adapting the covariance structures and the corresponding Yule-Walker system, derived for the event-based, point process MQHawkes model.

Let us detail the number of unknowns to be determined. Considering a time grid up to lag q ,

- $(L_j^i(\tau))_{i,j \in \{1,2\}, 1 \leq \tau \leq q}$ and $(\phi_j^i(\tau))_{i,j \in \{1,2\}, 1 \leq \tau \leq q}$ bring each $q \times 2 \times 2$ unknowns
- $(K_j^i(\tau_1, \tau_2))_{i,j \in \{1,2\}, 1 \leq \tau_1 < \tau_2 \leq q}$ brings $\frac{q(q-1)}{2} \times 2 \times 2$ unknowns. Indeed, since $(K_j^i)_{i,j \in \{1,2\}}$ are time-symmetric, it is sufficient to only determine the upper triangle entries.
- $(\phi_\times^i(\tau))_{i \in \{1,2\}, 1 \leq \tau \leq q}$ bring $q \times 2$ unknowns and $(K_\times^i(\tau_1, \tau_2))_{i \in \{1,2\}, 1 \leq \tau_1, \tau_2 \leq q, \tau_1 \neq \tau_2}$ bring $q \times (q-1) \times 2$ unknowns.

All this adds up to $4q^2 + 6q$ unknowns.

Now, given the form of the model, we need covariance structures containing information on both the time-diagonal of the kernels ($\tau_1 = \tau_2$) and their off-time-diagonal $\tau_1 \neq \tau_2$.

The first quantity of interest is the covariance between volatilities and realized squared returns, for all $\tau > 0$, for $i, j \in \{1, 2\}$:

$$\mathcal{C}_{ij}(\tau) := \mathbb{E}(\sigma_{i,t}^2 r_{j,t-\tau}^2) - \mathbb{E}(\sigma_i^2) \mathbb{E}(\sigma_j^2) \quad (18)$$

where we have noted that $\mathbb{E}(r_j^2) \equiv \mathbb{E}(\sigma_j^2)$.

As we shall see, computing $(\mathcal{C}_{ij}(\tau))_{i,j \in \{1,2\}, 1 \leq \tau \leq q}$ allows to set $q \times (2 \times 2)$ equations and predominantly shape the kernels $(\phi_i^j)_{i,j \in \{1,2\}}$.

We also define two relevant three-point correlation structures, for $0 < \tau_1 < \tau_2$,³

$$\mathcal{D}_{ij}(\tau_1, \tau_2) := \mathbb{E}((\sigma_{i,t}^2 - \langle \sigma_{i,t}^2 \rangle_t) r_{j,t-\tau_1} r_{j,t-\tau_2}) \quad (19)$$

and, for $0 < \tau_1, \tau_2$,

$$\mathcal{D}_{\times j}(\tau_1, \tau_2) := \mathbb{E}((\sigma_{j,t}^2 - \langle \sigma_{j,t}^2 \rangle_t) r_{j,t-\tau_1} r_{\bar{j},t-\tau_2}). \quad (20)$$

\mathcal{D}_{ij} is time-symmetric and allows to fix $q(q-1)/2 \times (2 \times 2)$ equations. It mainly shapes the kernels $(K_j^i)_{i,j \in \{1,2\}}$.

$\mathcal{D}_{\times j}$ fixes $q \times q \times (2 \times 1)$ equations, and mainly determines $(\phi_\times^i)_{i \in \{1,2\}}$ and $(K_\times^i)_{i \in \{1,2\}}$.

The above covariance structures fix $6q^2 + 2q$ values and enable one to characterize the kernels $(\phi_j^i)_{i,j \in \{1,2\}}$, $(K_j^i)_{i,j \in \{1,2\}}$, $(\phi_\times^i)_{i \in \{1,2\}}$ and $(K_\times^i)_{i \in \{1,2\}}$ up to a lag q , after which the covariance structures and kernels are considered negligible. Moreover, Eq. (9) enables to retrieve the baseline values from $(\phi_j^i)_{i,j \in \{1,2\}}$.

Finally, to recover the leverage feedback that measures the influence of past signed returns on future volatility and determine the missing $4q$ quantities, we define an additional two-point correlation:

$$V_{ij}(\tau) := \mathbb{E}(\sigma_{i,t}^2 r_{j,t-\tau}). \quad (21)$$

$(V_{ij}(\tau))_{i,j \in \{1,2\}, 1 \leq \tau \leq q}$ allows to fix $q \times (2 \times 2)$ equations and predominantly shapes the kernels $(L_i^j)_{i,j \in \{1,2\}}$.

Let us note that some of the aforementioned covariance structures involve fourth moments of financial returns which, due to their fat-tailed distribution, might be undefined and could lead to numerical instability. An interesting alternative would be to use mixed moments as used in Chicheportiche *et al.* (2014) [26]. To avoid obscuring our message, the corresponding Yule-Walker equations linking the feedback kernels to these modified covariance structures are established in Appendix C.

2. Fits and noise reduction

In order to reduce the numerical noise, the four covariance functions $\mathcal{C}, \mathcal{D}, \mathcal{D}_\times, \mathcal{V}$ computed on empirical data are approximated (“smoothed”) in the following way:

- $(\mathcal{C}_{ij})_{i,j \in \{1,2\}}$ and the time-diagonal of $(\mathcal{D}_{\times j})_{j \in \{1,2\}}$ are smoothed using their corresponding best fits

$$\tau \rightarrow \frac{n \exp(-\beta\tau)}{(1 + \gamma\tau)^\alpha}.$$

Figure 2 and Figure 3 illustrate the approximation of $(\mathcal{C}_{ij})_{i,j \in \{1,2\}}(\tau)$ and $(\mathcal{D}_{\times j}(\tau, \tau))_{j \in \{1,2\}}$, respectively, for the pair E-MINI vs. TBOND. These figures suggest that $(\mathcal{C}_{ij})_{i,j \in \{1,2\}}$ and the time-diagonal of $(\mathcal{D}_{\times j})_{j \in \{1,2\}}$ both exhibit long-range correlations (note the log-log scale in Figure 2). This observation justifies the use of a power law function to smooth the empirical values; it also means that volatility movements that happened early in the day still have influence on the end-of-day volatility.

³ Note that $\mathcal{D}_{ij}(\tau, \tau) \equiv \mathcal{C}_{ij}(\tau)$.

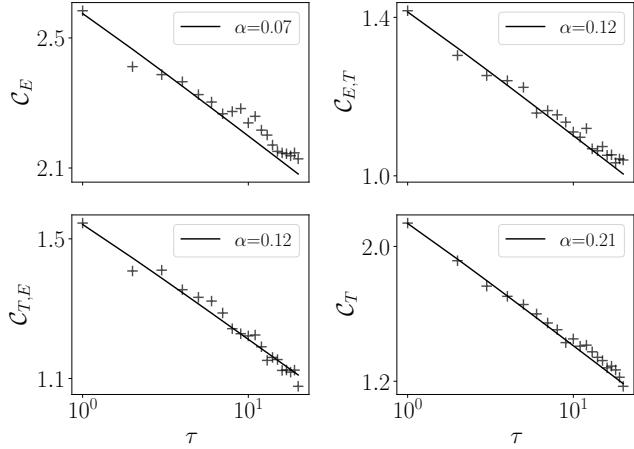


Figure 2: Two-points correlation $C_{ij}(\tau)$, as defined in Equation (18), calculated for the pair E-MINI vs TBOND. The scatter crosses represent the empirical data, while the solid black line represents the fit used for the calibration, given by $\tau \rightarrow n \exp(-\beta\tau)(1 + \gamma\tau)^{-\alpha}$. The parameter α is specified in each subplot's labels and the other parameters, (β, γ, n) are given in Table IV in Appendix F. Note that the subplots are presented on a log-log scale.

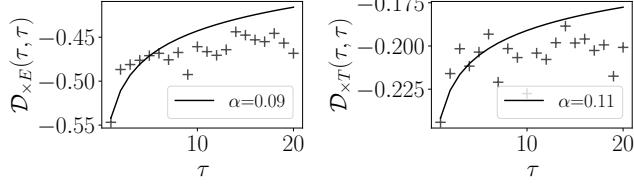


Figure 3: Three-points correlation ($\mathcal{D}_{\times j}$) for the pair E-MINI vs TBOND, as defined by Equation (20). The scatter crosses represent the empirical data, whereas the solid black line represents the fit used for the calibration, described by $\tau \rightarrow n \exp(-\beta\tau)(1 + \gamma\tau)^{-\alpha}$. The parameter α is specified in each subplot's labels and the other parameters, (β, γ, n) are given in Table V in Appendix F.

- We observe that the largest eigenvalue the off-diagonal part of $(\mathcal{D}_{ij}(\tau_1, \tau_2))_{1 \leq \tau_1, \tau_2 \leq q}$ for all $i, j \in \{1, 2\}$ is a lot more significant than the subsequent ones. This suggests using a rank-one approximation for these four matrices, where we furthermore smooth the first eigenvector, which is replaced by its best fit of

$$\tau \rightarrow a \exp(-b\tau).$$

For instance, Figure 4 presents the first eigenvector and its corresponding fit of the 3-points cor-

relation \mathcal{D}_{ij} for the pair E-MINI vs TBOND.

Let us note that, the ‘‘martingalisation’’ process described in Section III B 2 facilitates this exponential fit. Indeed, without applying the ‘‘martingalisation’’ process, the first eigenvector of the 3-point correlation matrix \mathcal{D}_{ij} would reflect the strong mean-reversion observed at this scale, characterized by the first two points having opposite signs.

We also observe from the eigenvalues of the off-diagonal part of $(\mathcal{D}_{\times j})_{j \in \{1, 2\}}$ that such coefficients are mostly noise hovering around zero, giving us an hint on the off-diagonal of the kernels (K_x^j) .

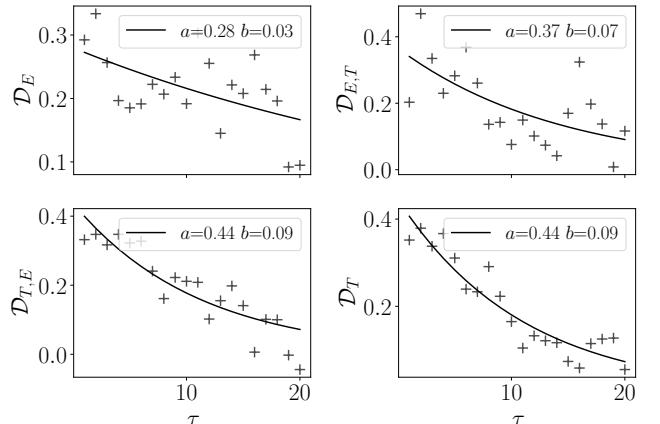


Figure 4: First eigenvector of the 3-points correlation \mathcal{D}_{ij} , as defined in Equation (19), for the pair E-MINI vs TBOND. The scatter plots are the empirical values and the solid black line is the exponential fit $\tau \rightarrow a \exp(-b\tau)$ with the corresponding values of a, b given in the caption.

- Finally, $(\mathcal{V}_{ij})_{i,j \in \{1, 2\}}$ is approximated by its best fit using $\tau \rightarrow a \exp(-b\tau)$. Figure 5 presents the approximation of all $(\mathcal{V}_{ij})_{i,j \in \{1, 2\}}$ for the pair E-MINI vs. TBOND, with their corresponding fit.

3. Empirical results: discussion

Let us comment on the empirical results obtained above.

1. First of all, the standard ‘‘ARCH’’ correlations $(C_{ii})_{i \in \{1, 2\}}$ between past square returns and future volatilities are strong and decay as power-laws, as well documented in the literature. We find that the corresponding cross-correlations between E-MINI and TBOND are of the same order of magnitude and also decay as power-laws.

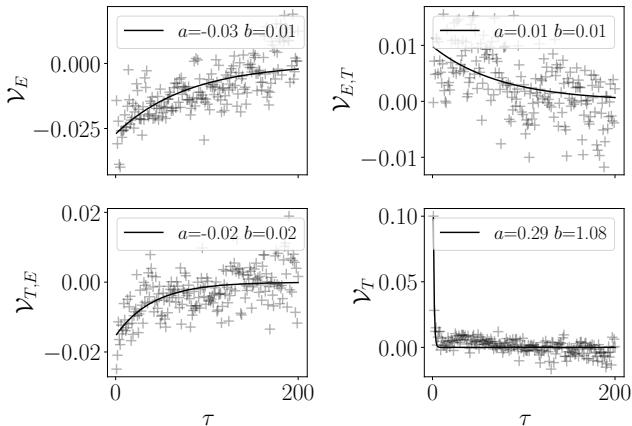


Figure 5: Two-points correlation $(\mathcal{V}_{ij})_{i,j \in \{1,2\}}$, as defined in Equation (21), calculated for the pair E-MINI vs TBOND. The scatter crosses represent the empirical data, whereas the solid black line represents the fit used for the calibration with the corresponding values of a, b given in the caption.

2. Turning to the time-off-diagonal part of $(\mathcal{D}_{ij}(\tau_1, \tau_2))$, we have seen that it is well reproduced by a rank-one matrix, with a corresponding vector that is well fitted by an exponential decay in τ . Following Ref. [6], this is interpreted as self- and cross-Zumbach effects: the exponential moving average of the past returns of both E-MINI and TBOND (i.e., the “trend”) is correlated with higher future volatilities.
3. Moreover, Figure 3 demonstrates that the diagonals of $(\mathcal{D}_{\times,j})$ for $j \in \{\text{E,T}\}$ are negative and decay slowly, where here and henceforth we abbreviate subscripts E-MINI and TBOND by E, T respectively. Since the returns of the E-MINI and the TBOND are negatively correlated during the studied period (correlation coefficient is $C_{\text{ET}} \approx -0.15$), $(\mathcal{D}_{\times,j}) < 0$ means that when past realized covariance is more negative than its average value -0.15 , then future volatility tends to be higher than usual, and conversely, when the returns of the E-MINI and the TBOND are less negatively correlated, then future volatility tends to be higher than usual. However, the time-diagonal of $(\mathcal{D}_{\times,j})$ also captures pure volatility effects which could explain such negative values as we know that volatility begets volatility. Only the determination of the influence kernels K_{\times}^j can ascertain the impact of realized correlations in the determination of volatility.
4. Finally, Figure 5 recovers the well known self-leverage effect on the E-MINI: negative past returns are correlated to higher future volatility.

But one also clearly sees that negative returns of the E-MINI increases the volatility of the TBOND as well (left panels). The upper right panel suggests that the cross-leverage correlation of the TBOND on the E-MINI is positive, although more noisy. This is however consistent with the negative correlation between the E-MINI and TBOND returns, and could be a spillover effect. Again, the determination of the influence kernels L_i^j will resolve this issue. It is also notable that the TBOND self-leverage effect (lower right panel) is one order of magnitude larger but decays very fast with lag τ .

C. Calibration using Yule-Walker equations

This section aims at delineating the calibration method, and illustrating in on the example of the pair E-MINI–TBOND. Appendices C and D provide further details.

1. A 4-step procedure

As explained above, most cross-asset effects can be captured by pair-wise interactions only, so our calibration strategy focuses on the two-dimensional (2D) version of the MQHawkes/MQGARCH model which accounts for leverage, linear and quadratic effects. In order to maintain numerical stability, the calibration process is broken down in successive steps, which take advantage of the fact that the different terms in the MQGARCH model have different orders of magnitude and different symmetries: self-feedback effects (K_i^i) are much larger than cross-feedback effects ($K_{\bar{i}}^i$), which are themselves larger than realized covariance effects (K_{\times}^i).

All steps rely on linear relationships between the covariance and the kernels, the so-called Yule-Walker system of equations. (Appendix C details those equations system for the model specified in Equation (7)). Solving such linear system of equations is achieved by considering that after a lag q the kernels and the covariance structures are negligible (see Appendix D for details on the implementation of the Yule-Walker systems for the 2D-QGARCH). We now outline the main steps of the calibration.

To accurately distinguish leverage feedback (which break the $r \rightarrow -r$ symmetry) from quadratic and cross-feedback effects, we initially estimate the linear, quadratic, and cross kernels using symmetrized $r \rightarrow -r$ data. Subsequently, we estimate the leverage kernels on

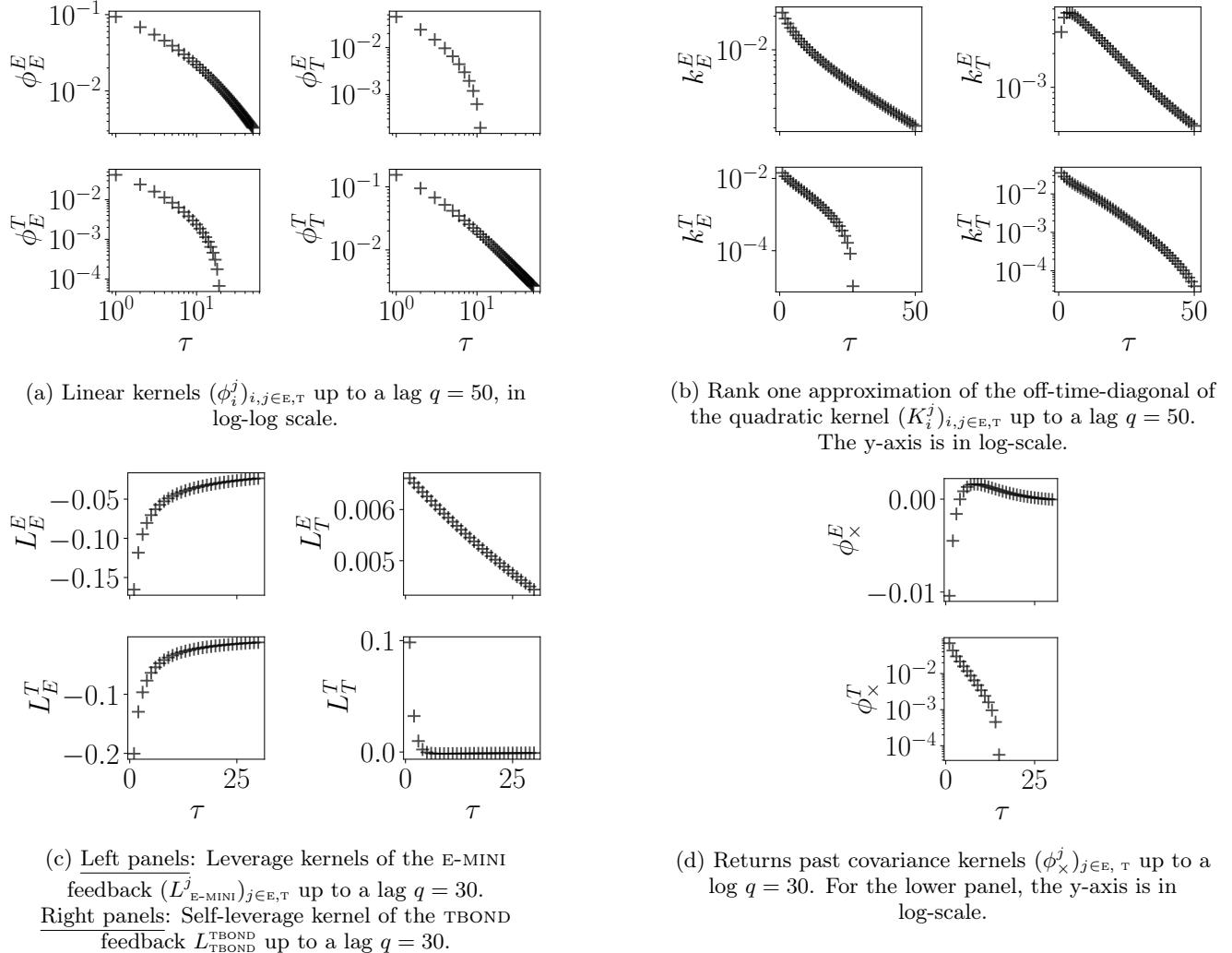


Figure 6: Results of the 2D QGARCH calibration on the pair E-MINI vs. TBOND.

the original dataset, excluding from \mathcal{V} the contributions accounted for by the linear feedback.

Specifically, to completely eliminate the leverage effect, we duplicate our datasets as follows. For each time series r_t, σ_t , we generate a mirror time series $-r_t, \sigma_t$ that we append to compute all correlation functions, with averages noted \mathbb{E}' . Hence leverage correlations are identically zero for this expanded data set, while all other correlations structures are strictly identical. Indeed,

$$\mathbb{E}'(\sigma_t r_{t-\tau}) = \frac{1}{2}[\mathbb{E}(\sigma_t r_{t-\tau}) - \mathbb{E}(\sigma_t r_{t-\tau})] = 0,$$

while, for example,

$$\mathbb{E}'(\sigma_t r_{t-\tau}^2) = \frac{1}{2}[\mathbb{E}(\sigma_t r_{t-\tau}^2) + \mathbb{E}(\sigma_t r_{t-\tau}^2)] \equiv \mathbb{E}(\sigma_t r_{t-\tau}^2).$$

The linear, quadratic and cross kernels are then estimated from in the following three-step procedure,

with a final step to retrieve the leverage kernels.

1. Since it is numerically stable, as shown in Appendix A, the univariate QHawkes feedback kernels of each asset are estimated up to a certain lag q . From this first step, we obtain $(\phi_i^i(\tau))_{i \in \{1,2\}, 1 \leq \tau \leq q}$ and $(K_i^i(\tau_1, \tau_2))_{i \in \{1,2\}, 1 \leq \tau_1 < \tau_2 \leq q}$.
2. Considering the cross contributions null in a first pass ($\phi_X = k_X = 0$), we estimate the cross linear and quadratic contributions characterised by the values $(\phi_j^i(\tau))_{i,j \in \{1,2\}, i \neq j, 1 \leq \tau \leq q}$ and $(K_j^i(\tau_1, \tau_2))_{i,j \in \{1,2\}, i \neq j, 1 \leq \tau_1 < \tau_2 \leq q}$. After this second step, the estimation of the baseline activity $\sigma_{1,\infty}^2$ and $\sigma_{2,\infty}^2$ can be achieved using Equation (9).

3. Accounting for the contributions of the kernels obtained in steps 1 and 2, we estimate the cross feedback components up to a certain lag q , i.e., $(\phi_{\times}^i(\tau))_{i \in \{1,2\}, 1 \leq \tau \leq q}$ and $(K_{\times}^i(\tau_1, \tau_2))_{i \in \{1,2\}, 1 \leq \tau_1, \tau_2 \leq q, \tau_1 \neq \tau_2}$.
4. Once these three steps are completed, the leverage kernels, $(L_i^j(\tau))_{i,j \in \{1,2\}, 1 \leq \tau \leq q}$, can then be estimated up to a lag q on the original dataset (r, σ) while removing the contributions of the quadratic kernels (we refer to Section D 7 of Appendix D for further details).

Note that the calibration using Maximum Likelihood estimation, whose implementation is detailed in Appendix G, has also been implemented and provides similar results as the method of moments. Notably, the Maximum Likelihood estimation is particularly efficient when the GMM outcome is used as an initial guess.

2. Some general empirical observations

First, as already mentioned above, we observe that the empirical cross-covariance structure $\mathcal{D}_{\times j}(\tau_1, \tau_2)$ is close to zero for $\tau_1 \neq \tau_2$. This means that the product $r_{i,t-\tau_1} r_{i,t-\tau_2}$ is uncorrelated with $\sigma_{i,t}^2$ unless $\tau_1 = \tau_2$: the realized correlation between the returns of two assets only affects future volatility when these returns are simultaneous. This observation leads to $K_{\times}^j(\tau_1, \tau_2) = 0$ when $\tau_1 \neq \tau_2$. Consequently, only the time-diagonal $\phi_{\times}^i(\tau) := K_{\times}^i(\tau, \tau)$ will be considered hereafter.

Second, consistent with the empirical findings of [6], the calibration of the off-time-diagonal of K_i^j reveals that these kernels can then be well approximated by a rank-one matrix, such that $K_i^j(\tau_1, \tau_2) \approx k_i^j(\tau_1)k_i^j(\tau_2)$. Indeed, the first eigenvalue of the matrices $K_i^j(\tau_1, \tau_2)$ with a diagonal of zeros, is substantially larger than the subsequent ones. Besides, the time-diagonal of these kernels is directly represented by $\phi_i^j(\tau)$.

V. CALIBRATION OF THE 2D-QGARCH ON FUTURES PAIRS

A. Results for the pair E-MINI–TBOND

Implementing the above procedure on the time series of the E-MINI and the TBOND resulted in the kernels presented in Figure 6. Figure 6b displays the rank-one approximation $k_i^j(\tau)$ of the off-time-diagonal of the quadratic kernels $(K_i^j)_{i,j \in \{E,T\}}$. Let us comment on the

results.

As already mentioned above, it appears that the covariance structures forms, \mathcal{C} , \mathcal{D} , \mathcal{D}_{\times} and \mathcal{V} (see Figures 2, 4, 3 and 5, respectively), shape the resulting kernels.

Notably, the long-range nature of $(\mathcal{C}_{ii})_{i \in \{E,T\}}$ is reflected in the kernels $(\phi_i^i)_{i \in \{E,T\}}$ shown in Figure 6a. The cross-influence Hawkes kernels ϕ_i^i have a smaller amplitude and decay faster but still appear to be very significant.

Figure 6b exhibits the shape of the “Zumbach” kernels k_i^j , i.e. the filters that determine the trends on the E-MINI and the TBOND which feedback on future volatility of both contracts. Of note: (i) the amplitude and range of the self-kernels k_i^i are similar for E-MINI and the TBOND and (ii) the cross-kernels k_i^j are similar magnitude as the self-kernels, although – perhaps surprisingly – the decay of k_e^t is faster than that of k_t^e : trends on the TBOND has a longer lasting effect on the volatility of the E-MINI than the other way round.

Turning now to Figure 6c, we see that the negative exponential shape of the influence of the E-MINI past returns on both σ_E^2 and σ_T^2 is well-captured by the kernels $(L_E^i)_{i \in \{E,T\}}$. Additionally, the leverage kernel capturing the feedback of TBOND returns on its own volatility, depicted in the lower right panel of Figure 6c, exhibits a positive shape that rapidly decays to zero. This observation is consistent with the shape of \mathcal{V}_{TBOND} in Figure 5, and with the negative correlation between the E-MINI and the TBOND returns. Consistently with the observation of Figure 5, the leverage kernel representing the feedback of the TBOND returns on the E-MINI volatility is weakly positive, again consistent with the negative correlation between the E-MINI and the TBOND returns. Note that stronger leverage signals are expected for larger time scales, as the leverage effect predominantly manifests at the daily scale, as demonstrated by the first figure of [5].

Finally, Figure 6d represents the perhaps less trivial of all feedback kernels, $(\phi_{\times}^j)_{j \in \{E,T\}}$, that describe the influence of the realized covariance between the returns of the E-MINI and the TBOND on future volatility of both assets. These graphs indicate that when the realized covariance is more negative than usual (a.k.a. a flight to quality mode), it tends to decrease the TBOND future volatility while it increases the E-MINI volatility for short time horizon before perhaps reverting sign for longer time horizons. Note that this is not exactly the effect that was previously discussed for $(\mathcal{D}_{\times j})_{j \in \{E,T\}}$ in Figure 3. In fact, the kernels $(\phi_{\times}^j)_{j \in \{E,T\}}$ reveal solely the effect of past returns covariance, while $(\mathcal{D}_{\times j})_{j \in \{E,T\}}$ conflate feedback effects arising from the Hawkes kernels.

B. Extension to other pairs of futures contracts

We now extend the 2D-QGARCH calibration over all pairs composed by the following six futures: four on very correlated indices: E-MINI, E-MINI-3 (3 months futures), NASDAQ, and DOW JONES, one on commodities (CRUDE OIL) and finally TBOND.

The temporal structure of all kernels are found to be very similar to those obtained for the pair E-MINI–TBOND, so we chose to only display here the L1-norm of the corresponding kernels in Figure 7; see Appendix E for their full profiles.

1. Volatility feedback (Hawkes) kernels ϕ_j^i

The primary observation from Figure 7 concerns the kernels characterising the influence of past volatility on future volatility, ϕ_j^i , shown in Figure 7a. The strong diagonal of Figure 7a indicates that most of the system endogeneity originates from the Hawkes feedback of the asset’s own volatility on itself. Additionally, Figure 7a demonstrates that the system operates near criticality, as the diagonal values are all above 0.6 and the largest eigenvalues of the pairwise matrices $\|\phi_j^i\|_1$ are between 0.73 and 0.90. Of note is the rather strong column corresponding to the E-MINI as a trigger of volatility for all other markets, whereas the column corresponding to NASDAQ is surprisingly weak.

2. Zumbach trend feedback kernels k_j^i

The norms of the rank-one approximations k_j^i of the off-time-diagonal of the quadratic kernels K_j^i , in Figure 7b, are consistent with their profiles in Appendix E. The strong diagonal of Figure 7b shows that the self-Zumbach feedback is quite strong compared to the cross-Zumbach effects. Furthermore, some futures present particularly trend quadratic feedback, like CRUDE OIL.

As for the Hawkes feedback, the first column of Figure 7b demonstrates that the cross-Zumbach effect arising from the E-MINI on futures on other markets is quite strong compared to the other cross-Zumbach kernels, meaning that the S&P500 index past trends predominantly increase the volatility of the other futures on indices. Similarly, the last column of Figure 7b shows that the Zumbach feedback of the TBOND on stock indices is significant and consistent across contracts.

3. Leverage feedback kernels L_j^i

It appears that leverage kernels are not always relevant at the intraday level, as evidenced by the many zeros in Figure 7c. In fact, certain leverage kernels derived from our calibration process exhibit several constraints inherent to our framework, prompting the exclusion of the corresponding norms and profiles depicted in Figures 7c and 21. The primary limitation is the finite dimensionality of the calibration, which is further emphasized by our strategy to remove the influence of overnight effects by averaging correlations solely within individual days. This approach is particularly significant in the context of leverage kernels, as the leverage effect predominantly appears at the daily scale. Additionally, leverage kernels are computed only during the fourth and last step of our method, following the removal of the linear kernels’, ϕ , contribution from the correlation \mathcal{V} . Consequently, the final correlation used to derive L is notably influenced by the outcomes of the initial correlation step.

When present, except in the case of the TBOND as discussed above, the leverage feedback is negative, consistent with the idea that large negative returns trigger increased volatility. Notably, the DOW JONES futures⁴ seems to be particularly affected by the leverage effect from both itself and other futures on indices. We also point out again that the model is calibrated intraday, whereas the leverage effect predominantly appears on a daily scale.

4. Covariance feedback kernels ϕ_{\times}^i

Figure 7d reports on the influence of past realized covariance of the returns of the two assets involved in each pair on the volatility of both assets in the pair.

The pair E-MINI–E-MINI-3 exhibits a particularly strong returns covariance feedback compared to the other pairs (see the upper left 2×2 matrix of Figure 7d), which means that an increased past covariance between the returns of the E-MINI and those of the E-MINI-3 tends to increase the volatility of both futures.

Additionally, the volatility of the (TBOND) and of the (CRUDE OIL) appear to be significantly impacted by the past covariance of their returns with the returns of US stock-indices, as demonstrated by the last two rows of Figure 7d. Specifically, for all the pairs combining the CRUDE OIL with stock-indices, the returns covariance feedback kernels ϕ_{\times} exhibit a negative norm, while the

⁴ US index.

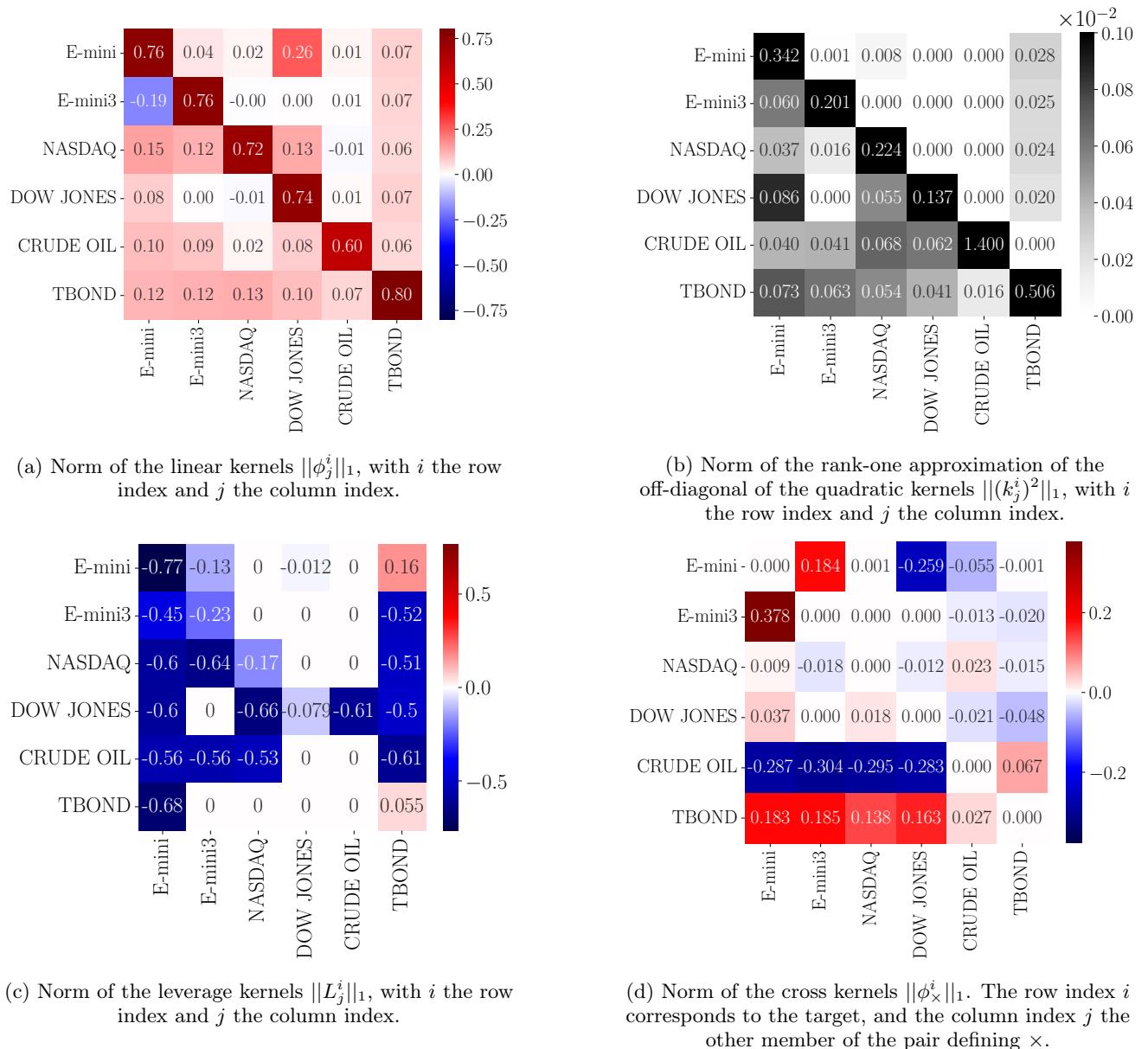


Figure 7: Norms of the kernels from the calibration of 2D-QGARCHs on pairs of futures on indices. For the four sub-figures, the x-labels determine the index i corresponding to the “source” and the y-labels determine the index j corresponding to the “target” of the feedback.

pairs combining a stock-index with the TBOND, display a positive returns covariance feedback.

During the studied period, CRUDE OIL and stocks were positively correlated while stocks and bonds were negatively correlated. Therefore, the sign of the norm of the returns covariance feedback kernels $\|\phi_x^i\|$ for pairs combining stock-indices with the TBOND and for pairs combining stock-indices with the CRUDE OIL conveys the same meaning in both cases: when the absolute value of the past returns covariance increases, it tends

to decrease future volatility of both TBOND and CRUDE OIL. The opposite effect on the E-MINI is much weaker and has the opposite sign with the TBOND, as already noted above. As for the pair CRUDE OIL–TBOND, an increased, in absolute value, of the past returns covariance tends to decrease TBOND volatility while it increases CRUDE OIL volatility.

VI. A MQARCH FACTOR-MODEL

To calibrate on a large set of stocks, while accommodating for the specific feedback from the index to which they belong, or from more general factors we define and calibrate a 2D-QGARCH factor model. The framework is outlined below.

A. Model framework

The starting point is a 1-factor model. Specifically, the returns of stock i are decomposed into a sum of their market exposure, represented by the factor component f_0 , with an exposure coefficient β_i , and a residual component $e_{i,t}$, which is idiosyncratic to the stock dynamics:

$$r_t^i = \beta_i f_{0,t} + e_{i,t}. \quad (22)$$

Empirically, f_0 represents the market mode and will be directly sourced from the E-MINI returns studied in the previous section, while β_i is estimated from the correlations between the returns of stock i and f_0 .

The particularity of such framework, and our main interest, lies in the non-linear dynamics of f_0 and e_i . Specifically, we posit here that the volatility of the market factor f_0 is unaffected by the dynamics of residuals e_i and given by a 1D-QGARCH, while the e_i s themselves follow a 2D-QGARCH model with (i) self-feedback and (ii) feedback from the market factor. This will account for the empirical observation that the volatility of the residuals e_i depends on that of the market mode f_0 [28–31].

Correspondingly, the volatility of f_0 is written as

$$\begin{aligned} \sigma_{0,t}^2 = & \sigma_{0,\infty}^2 + \sum_{k=1}^{+\infty} L_0^0(k) f_{0,t-k} + \sum_{k=1}^{+\infty} \phi_0^0(k) f_{0,t-k}^2 \\ & + 2 \sum_{k_1=1}^{+\infty} \sum_{k_2=k_1+1}^{+\infty} K_0^0(k_1, k_2) f_{0,t-k_1} f_{0,t-k_2}, \end{aligned}$$

while the volatility of e_i is specified as

$$\begin{aligned} \sigma_{i,t}^2 = & \sigma_{i,\infty}^2 + \sum_{k=1}^{+\infty} L_i^i(k) e_{i,t-k} + \sum_{k=1}^{+\infty} L_0^i(k) f_{0,t-k} \\ & + \sum_{k_1=1}^{+\infty} \sum_{k_2=1}^{+\infty} K_i^i(k_1, k_2) e_{i,t-k_1} e_{i,t-k_2} \\ & + \sum_{k_1=1}^{+\infty} \sum_{k_2=1}^{+\infty} K_0^i(k_1, k_2) f_{0,t-k_1} f_{0,t-k_2} \\ & + \sum_{k_1,k_2=1}^{+\infty} K_\times^i(k_1, k_2) e_{i,t-k_1} f_{0,t-k_2} \end{aligned} \quad (23)$$

Scrutinizing this last expression, we see that the idiosyncratic volatility, σ_i^2 , is influenced by the past returns of both the factor and the residual, characterised by (i) the leverage kernels L_\times and L_i respectively, (ii) by the past trends of both the factor and the residual, described by the quadratic kernels K_0^i and K_i^i respectively and (iii) a covariance feedback effect described by K_\times^i . Note that since the correlation between factor f_0 and residuals e_i is zero by construction, we do not need to subtract the average C_{0i} from $e_{i,t-k} f_{0,t-k}$ in the last term.

In fact, empirical data suggests that the covariance feedback effect between e_i and f_0 is negligible on intraday timescales but may become relevant at lower frequencies (at variance with the case of the E-MINI – TBOND pair). We leave this question open for future investigations and will drop the K_\times^i terms henceforth. As before, we can decompose the quadratic kernels to highlight the diagonal feedback (Hawkes feedback of the activity onto itself), which primarily characterizes the endogeneity of the system, and the off-diagonal feedback (Zumbach feedback of the trends) as follows:

$$\begin{aligned} \sigma_{i,t}^2 = & \sigma_{i,\infty}^2 + \sum_{k=1}^{+\infty} L_i^i(k) e_{i,t-k} + \sum_{k=1}^{+\infty} L_0^i(k) f_{0,t-k} \\ & + \sum_{k=1}^{+\infty} \phi_i^i(k) e_{i,t-k}^2 + \sum_{k=1}^{+\infty} \phi_0^i(k) f_{0,t-k}^2 \\ & + 2 \sum_{k_1=1}^{+\infty} \sum_{k_2=k_1+1}^{+\infty} K_i^i(k_1, k_2) e_{i,t-k_1} e_{i,t-k_2} \\ & + 2 \sum_{k_1=1}^{+\infty} \sum_{k_2=k_1+1}^{+\infty} K_0^i(k_1, k_2) f_{0,t-k_1} f_{0,t-k_2} \end{aligned} \quad (24)$$

The goal of this section is to investigate the values of $\sigma_{i,\infty}^2$ and the features of the various feedback kernels determining the residual future volatility for US stocks, in particular the role of the index factor f_0 on this volatility since we know from the work of Refs. [28–31] that the equal time correlation between residual volatilities and the market factor is non-trivial. Of particular interest are the Zumbach self-trend and cross-trend effects.

The subsequent section is dedicated to presenting the data used in this section, followed by a discussion of the results obtained.

B. Data processing

Our initial dataset contains 1-minute returns $(r^i)_{i=1,\dots,317}$ of 317 US stocks spanning from 2013 to 2023 and the 1-minute E-MINI returns, f_0 , for the same period. We consider only stocks which belonged to the S&P500 index for the whole studied period.

This initial dataset is then used to deduce the time series $(e_i)_i$ according to Equation (22) and taking, for all i , β_i as the covariance coefficients between the returns of stock i and the returns of the E-MINI. The squared volatility time series, $\sigma_{i,t}^2$ and $\sigma_{f_0,t}^2$, are then approximated by the squared 1-minute returns, $e_{i,t}^2$ and $f_{0,t}^2$ respectively.

At this stage, we thus end up with two time series for the E-MINI and for each stock of our 317 US stocks sample: one characterising the returns and the second one representing the squared volatility. These times series are then processed as before following the steps described in Section III B.

The method of moments is then used to calibrate the 1D-QGARCH on the E-MINI, resulting in the values of $\sigma_{0,\infty}^2$, $(L_0(k))_{1 \leq k \leq q}$ and $(K_0(k_1, k_2))_{1 \leq k_1, k_2 \leq q}$, where q is the lag beyond which feedback effects are considered negligible.

Finally, the calibration of the 2D-QGARCH on the 317 stocks residuals, adapting the method described in Section IV and Appendix D, is performed, with results presented in the next section.

C. MQARCH 1-factor model: Calibration Results

The kernels L_0^0 , ϕ_0^0 and k_0^0 , where ϕ_0^0 and k_0 are respectively the time-diagonal of K_0^0 and the rank-one approximation of the off-time-diagonal of K_0 , result from the calibration of a 1D-QGARCH on the E-MINI, and are therefore identical to those already computed in the previous section and shown in Figure 6.

Figure 8 presents the *cross-sectionally averaged* profiles of the quadratic kernels shaping the volatility of the stock idiosyncratic dynamics, i.e. the volatility feedback kernels $\langle \phi_i^i \rangle_i$ and $\langle \phi_0^i \rangle_i$, and the rank-one approximation of the off-time-diagonal of the feedback kernels, $\langle k_i^i \rangle_i$ and $\langle k_0^i \rangle_i$.

Interestingly, on average the norm of the residual volatility feedback kernel is close to one ($\langle \|\phi_i^i\| \rangle_i = 0.94$), supporting the aforementioned results suggesting that endogeneity arises chiefly from the self-volatility feedback. Additionally, the kernels characterising the residual volatility feedback on itself exhibit a long-range, power-law behaviour, as demonstrated by the log-log scale of the left panel of Figure 8. This means that the idiosyncratic activity that occurred early in the day influences the realized idiosyncratic volatility throughout the entire day (recall that we have removed any intraday seasonal pattern in the data). The thin grey area on the left panel of Figure 8 indicates the cross-sectional dispersion of individual kernels around the sample mean. We thus conclude that the Hawkes kernels $\phi_i^i(\tau)$ are quite uniform across stocks, which is

by no means a trivial result.

Moreover, the average endogeneity arising from the market contribution is not negligible, reaching up to $\|\phi_0^i\|_{\max} = 0.12$, with values hovering around $\langle \|\phi_0^i\| \rangle_i = 0.02$. Past volatility of the market factor appears to predominantly influence the residual volatility of stocks in the Financials, Real Estate, and Industrials sectors. Additionally, the log-log scale of the second panel from the left of Figure 8 can be fitted by a power-law function truncated by an exponential.

As previously found, the off-time-diagonal of the trend feedback kernels, as characterised by their rank-one approximation, have lower values than the diagonal feedback kernels with a norm averaging at $\langle \|\kappa_i^2\| \rangle_i = 0.012$. This is consistent with the results of the calibration of the QHawkes on stocks in [6]. The average profile of k_i^i s does not reflect strong differences between stocks for small lags, but such differences tend to increase with lag as demonstrated by the large grey area in the third panel from the left of Figure 8. This might however be due to numerical noise, which becomes more important for large lags. Notably, the higher the residual activity, the longer the range of the residual Zumbach feedback. Additionally, the Real Estate and Utilities sectors, which show the highest average levels of idiosyncratic activity, consistently present trend feedback kernels with the highest norms and longer temporal spans on average, specifically, $\langle \|(k_i^i)^2\| \rangle_{i \in \text{real estate}} = 0.017$ and $\langle \|(k_i^i)^2\| \rangle_{i \in \text{utilities}} = 0.018$.

The kernels characterising the market past trends contribution to future residual volatility exhibit much smaller but still clearly discernible values, with a norm averaging $\langle \|(k_0^i)^2\| \rangle_i = 4.7 \cdot 10^{-4}$, as shown by the profile in the right panel of Figure 8. Interestingly, the self-Zumbach effect of the index increases the activity of the index, which, in turn, increases the activity of the stock through ϕ_0^i , with a net effect of the same order of magnitude as the direct coupling k_0^i . Once again, the Real Estate and Utilities sectors stand out as having the most significant direct trend index feedback.

Surprisingly, we found highly disparate results across our sample of stocks for the self-leverage kernels L_i^i . In fact, some stocks exhibit positive residual leverage feedback. Note however that the leverage effect predominantly manifests itself at the daily scale.

Nevertheless, the leverage feedback L_0^i of the index on the volatility of the residuals is quite stable across stocks and its average profile is represented in Figure 9, together with a thin grey zone indicating a very small amount of cross-sectional dispersion. The negative exponential shape of L_0^i indicates that the E-MINI negative past returns directly increase the residual volatility of single stocks. This sheds a complementary light to the results presented in [32].

Finally, the endogeneity ratios of these 2D-

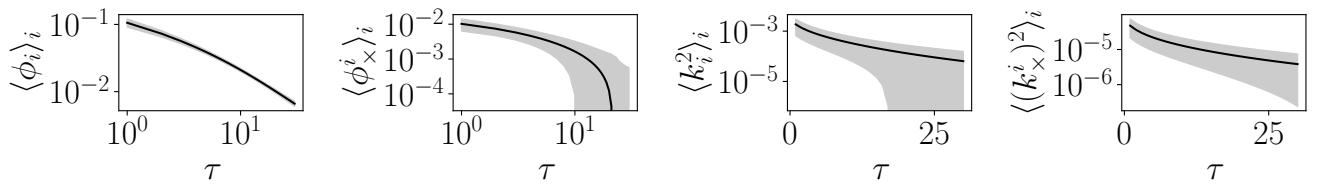


Figure 8: Average profiles of kernels for the dynamics of the residuals e_i . The solid black line represents the average profile and the grey area indicates the range of one standard deviation across the stocks.

QGARCH-1 factor systems⁵ are very close to 1, with all values exceeding 0.955 and an average of 0.96, indicating that the systems operate near criticality. Correspondingly the baseline fraction of the stock idiosyncratic activity averages at $\langle \sigma_{i,\infty}^2 \rangle_i = 0.04$.

To recap, the calibration of our model on single stocks yields two main results. First, most of the system endogeneity stems from the long-range feedback of the idiosyncratic volatility, consistent with previous findings on the calibration of futures on indices. Second, past volatility and negative returns of the index both significantly impact the residual volatility of stocks, supporting empirical findings that the idiosyncratic volatility of stocks is strongly coupled to that of the market as a whole [28–31].

A natural extension of this framework would be to investigate a factor model with more than one factor, thereby characterising the influence of the index not only on the stock’s idiosyncratic component but also on the various industrial sectors, whose dynamics would be accounted for by the additional factors.

VII. CONCLUSION

Let us summarise our work. Building upon the work of Blanc *et al.* [6] and our theoretical work presented in

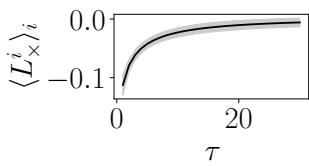


Figure 9: Average profile of the leverage kernel from the E-MINI on the residual L_x^i .

[5], we presented a non-parametric calibration method to characterise the feedback mechanisms that influence the dynamics of future volatility. To overcome the challenges associated with estimating an underlying MQHawkes process from aggregated data, the micro-scale MQHawkes framework is approximated by a discrete MGQARCH framework. The key to this transformation is assimilating the Hawkes intensity as the squared volatility, up to a multiplicative factor.

We introduced the key ingredients and steps required to calibrate the 2D-MQARCH framework using the method of moments, employing the pair E-MINI vs. TBOND as a pedagogical example. The primary idea of the calibration is to build a linear system of equations using the covariance structures, which are directly observable in real data, to deduce the kernel functions. The difficulty we had to overcome to get stable and meaningful results is that these covariance structures have quite different orders of magnitude and relate to different symmetries of the feedback terms. This has led us to propose a robust 4-step calibration scheme, where dominant effects are first accounted for, before including weaker effects.

Using this calibration method on pairs of futures revealed that the strongest feedback effect on future volatility originates from the past volatility itself and the sign of past returns (“leverage”), followed by the self-Zumbach effect. Cross effects are smaller, but clearly significant and interpretable. In particular, our analysis demonstrated that the past trends and returns of the E-MINI contribute significantly to the activity of other futures, compared to other cross-feedback effects. Interestingly, we have discovered a new stylized fact: realized correlation between the E-MINI and other futures contracts does influence the volatility of these contracts.

To investigate the feedback mechanisms on stock volatility while distinguishing market dynamics from stock idiosyncratic dynamics, we combined the 2D-QGARCH framework with a 1-factor model, which allowed us to disentangle the self-influence from the influence of the market factor on the idiosyncratic activity of single stocks. Calibrating the 2D-QGARCH 1-factor model on 317 US stocks proved to be quite stable across stocks. Notably, while most of the system endogeneity stems from the influence of past residual

⁵ The endogeneity ratio of the 2D-QGARCH-1 factor model for stock i is the largest eigenvalue of $\begin{pmatrix} \|\phi_0^0\| & 0 \\ \|\phi_0^i\| & \|\phi_i^0\| \end{pmatrix}$, which is thus directly given by $\max(\|\phi_0^0\|, \|\phi_i^0\|)$.

activity on itself, the past returns and past volatility of the market factor still significantly impact the future volatility of the stock idiosyncratic component. In our analysis, we observed some disparities among industrial sectors. It would be interesting to further develop a MQARCH multi-factor framework to investigate the feedback loops between the market, the industrial sectors and the stock idiosyncratic component. We leave such a study for future research.

Finally, it would be interesting to undertake a direct calibration of the MQHawkes process using synchronized, tick by tick data. However, some of the feedback terms that we have discussed seem to arise at lower frequencies, so for many purposes extending our MQARCH calibration method to daily data would also be insightful.

ACKNOWLEDGEMENTS

We would like to thank Jérôme Garnier-Brun and Salma Elomari, for their insights on data analysis and calibration. We are also grateful to Natascha Hey and Jutta Kurth for their research internships which gave a lot of insights on the calibration on tick-by-tick data and on multivariate processes. This research was conducted within the Econophysics & Complex Systems Research Chair, under the aegis of the Fondation du Risque, the Fondation de l’Ecole polytechnique, the Ecole polytechnique and Capital Fund Management.

-
- [1] Hawkes AG. Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1971;33(3):438-43.
 - [2] Hawkes AG. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*. 1971;58(1):83-90.
 - [3] Hawkes AG. Hawkes processes and their applications to finance: a review. *Quantitative Finance*. 2018;18(2):193-8.
 - [4] Bacry E, Muzy JF. Hawkes model for price and trades high-frequency dynamics. *Quantitative Finance*. 2014;14(7):1147-66.
 - [5] Aubrun C, Benzaquen M, Bouchaud JP. Multivariate quadratic Hawkes processes—part I: theoretical analysis. *Quantitative Finance*. 2023;23(5):741-58.
 - [6] Blanc P, Donier J, Bouchaud JP. Quadratic Hawkes processes for financial prices. *Quantitative Finance*. 2017;17(2):171-88.
 - [7] Bacry E, Mastromatteo I, Muzy JF. Hawkes processes in finance. *Market Microstructure and Liquidity*. 2015;1(01):1550005.
 - [8] Ogata Y. Estimators for Stationary Point Processes. *Ann Inst Statist Math*. 1978;30(Part A):243-61.
 - [9] Ogata Y, Akaike H. On linear intensity models for mixed doubly stochastic Poisson and self-exciting point processes. *Selected Papers of Hirotugu Akaike*. 1998:269-74.
 - [10] Veen A, Schoenberg FP. Estimation of space-time branching process models in seismology using an em-type algorithm. *Journal of the American Statistical Association*. 2008;103(482):614-24.
 - [11] Lewis E, Mohler G. A nonparametric EM algorithm for multiscale Hawkes processes. *Journal of nonparametric statistics*. 2011;1(1):1-20.
 - [12] Marsan D, Lengline O. Extending earthquakes' reach through cascading. *Science*. 2008;319(5866):1076-9.
 - [13] Hansen NR, Reynaud-Bouret P, Rivoirard V. Lasso and probabilistic inequalities for multivariate point processes. 2015.
 - [14] Reynaud-Bouret P, Schbath S. Adaptive estimation for Hawkes processes; application to genome analysis. 2010.
 - [15] Bacry E, Jaisson T, Muzy JF. Estimation of slowly decreasing hawkes kernels: application to high-frequency order book dynamics. *Quantitative Finance*. 2016;16(8):1179-201.
 - [16] Bacry E, Muzy JF. Second order statistics characterization of Hawkes processes and non-parametric estimation. arXiv preprint arXiv:14010903. 2014.
 - [17] 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.
 - [18] Bacry E, Dayri K, Muzy JF. Non-parametric kernel estimation for symmetric Hawkes processes. Application to high frequency financial data. *The European Physical Journal B*. 2012;85:1-12.
 - [19] Hardiman SJ, Bouchaud JP. Branching-ratio approximation for the self-exciting Hawkes process. *Physical Review E*. 2014;90(6):062807.
 - [20] Achab M, Bacry E, Gaiffas S, Mastromatteo I, Muzy JF; PMLR. Uncovering causality from multivariate Hawkes integrated cumulants. 2017;1-10.
 - [21] 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):1137-56.
 - [22] 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.
 - [23] Guyon J, Lekeufack J. Volatility is (mostly) path-dependent. *Quantitative Finance*. 2023;23(9):1221-58.
 - [24] Guyon J, Parent L. The Discrete-Time 4-Factor Path-Dependent Volatility Model: Calibration under P and Q. Available at SSRN 5382733. 2025.
 - [25] Sentana E. Quadratic ARCH models. *The Review of Economic Studies*. 1995;62(4):639-61.
 - [26] Chicheportiche R, Bouchaud JP. The fine-structure of volatility feedback I: Multi-scale self-reflexivity. Phys-

- ica A: Statistical Mechanics and its Applications. 2014;410:174-95.
- [27] Bacry E, Delattre S, Hoffmann M, Muzy JF. Modelling microstructure noise with mutually exciting point processes. Quantitative finance. 2013;13(1):65-77.
- [28] Cizeau P, Potters M, Bouchaud JP. Correlation structure of extreme stock returns. 2001.
- [29] Lillo F, Mantegna RN. Symmetry alteration of ensemble return distribution in crash and rally days of financial markets. The European Physical Journal B-Condensed Matter and Complex Systems. 2000;15:603-6.
- [30] Allez R, Bouchaud JP. Individual and collective stock dynamics: intra-day seasonalities. New Journal of Physics. 2011;13(2):025010.
- [31] Chicheportiche R, Bouchaud JP. A nested factor model for non-linear dependencies in stock returns. Quantitative Finance. 2015;15(11):1789-804.
- [32] Reigneron PA, Allez R, Bouchaud JP. Principal regression analysis and the index leverage effect. Physica A: Statistical Mechanics and its Applications. 2011;390(17):3026-35.
- [33] Ogata Y. On Lewis' simulation method for point processes. IEEE transactions on information theory. 1981;27(1):23-31.
- [34] Bouchaud JP, Bonart J, Donier J, Gould M. Trades, quotes and prices: financial markets under the microscope. Cambridge University Press; 2018.
- [35] Aubrun C. Unraveling Financial Market Quakes: Exploring Endogenous Volatility Dynamics in Interconnected Markets. Institut Polytechnique de Paris; 2024.

Appendix A: Calibration - Proof of concept

This appendix provides a proof of concept for the proposed calibration method, demonstrating its results using synthetic data. Notably, we discuss the challenges associated with accurately retrieving the (Q)Hawkes process, an event-by-event point process, from aggregated data. These difficulties justify our decision to approximate the QHawkes process with a QGARCH model, as described by Equation (C1).

This appendix is organised as follows. It begins by comparing the calibration of the univariate (Q)Hawkes model with that of the univariate (Q)GARCH model. This comparison shows that, in certain regimes, the calibration of the Hawkes becomes biased due to data binning whereas the QGARCH remains stable across a wider range of parameters.

The second part focuses on calibrating the multivariate (Q)GARCH using the method of moments, providing evidence that the calibration method is reliable.

To compare the calibration of the (Q)Hawkes with the calibration on the (Q)GARCH, we generate datasets in two ways:

- The thinning algorithm [33] provide time and mark of events for the (Q)Hawkes, enabling the construction of price time series, according to the definition of the QHawkes, i.e. $dP = \pm dN$. These time series are then aggregated into 1 minute bins to create the volatility time series σ^2 and the 1-minute returns time series dP . These synthetic time series are used to test the calibration of the (Q)Hawkes model.
- The simulation of the (Q)GARCH, as described in Equation (C1), directly results in two time series at 1 minute timescale: the volatility σ^2 and the returns r . These time series are used to test the calibration of the QGARCH.

In both cases, we use exponential feedback kernels with the general form:

$$\phi(t) = n_H \beta \exp(-\beta t) \quad \& \quad k(t) = \sqrt{2n_Z \omega} \exp(-\omega t)$$

1. Univariate cases: comparison between the (Q)Hawkes and the (Q)GARCH models

This first part is dedicated to show and discuss the calibration results of the univariate (Q)Hawkes and of the univariate (Q)GARCH. It begins with the linear case and then deals with the quadratic extension.

a. Univariate linear case



Figure 10: Calibration of a 1D linear Hawkes with exponential kernel, for two different sets of parameters. In both cases, $n_H = 0.7$ and $\beta = 0.04 \text{ min}^{-1}$ and only the baseline varies. The scatter plot is the kernel obtained by the method of moments and the plain line is the kernel used to generate the synthetic data and that we want to recover.

As a reminder, the linear Hawkes process is a point process N , whose intensity is defined as follows:

$$\lambda_t = \lambda_\infty + \int_{-\infty}^t \phi(t-s) dN_s \quad \text{with returns defined as} \quad dP_t = \pm dN_t.$$

In the case of a linear GARCH process, the squared volatility σ^2 is driven by

$$\sigma_t^2 = \sigma_\infty^2 + \sum_{k=1}^{+\infty} \phi(k) r_{t-k}^2 \quad \text{with returns defined as } r_t = \sigma_t \xi_t, \quad \text{with } \xi_t \sim \mathcal{N}(0, 1).$$

The connection between the two models lies in the relationship between the GARCH squared volatility and the Hawkes intensity: $\frac{\sigma^2}{\Delta} \equiv \lambda_t$, where Δ is the time bin used to compute σ [6]. It is interesting to note that, by definition, the Hawkes model is a continuous process describing each event of price change whereas the GARCH model is inherently discretized in time.

The objective of the calibration is to retrieve the kernel ϕ . Specifically, in the case of an exponential kernel, $\phi = n_H \beta \exp(-\beta t)$, we aim to determine the parameters n_H , β and λ_∞ (or σ_∞^2). When the shape of the kernel is unknown, as it is often the case with empirical data, adopting a non-parametric method such as the general method of moments (GMM) allows us to remain agnostic about the kernel's shape. In this appendix, we implement the GMM on synthetic data to test its validity.



Figure 11: Illustration of the estimation of the covariance structure $C(\tau) = \frac{1}{\lambda} \left(\mathbb{E}\left(\frac{dN_t}{dt} \frac{dN_{t-\tau}}{dt}\right) - \bar{\lambda}^2 \right)$ from the generation of a 1D linear Hawkes with exponential kernel, for 2 different sets of parameters. In both cases, $n_H = 0.7$ and $\beta = 0.04 \text{ min}^{-1}$ and only the baseline varies. The scatter plot is the obtained covariance C , while the plain line is the theoretical covariance C [34].

Two key elements about the method of moments should be kept in mind. First, it relies on autocovariance structure, and in particular in the linear case, on the autocovariance of the process's activity $C(\tau) = \frac{1}{\lambda} \left(\mathbb{E}\left(\frac{dN_t}{dt} \frac{dN_{t-\tau}}{dt}\right) - \bar{\lambda}^2 \right)$ (for the GARCH, an equivalent expression is $C(\tau) = \mathbb{E}(\sigma_t^2 r_{t-\tau}^2) - \bar{\sigma}^2 \bar{r}^2$). These values form the linear system of equations required to recover ϕ (for more comprehensive details see [6] and Appendix D, in particular Section D 3). Secondly, the method of moments yields the values $(\phi(\tau))_{1 \leq \tau \leq q}$, where q is the lag after which the covariance and kernels are considered null.

We now proceed to the presentation and discussion of the calibration results of the Hawkes process.

Figure 10 shows the estimations of the linear kernels ϕ calibrated on synthetic data generated by linear Hawkes processes for two sets of parameters. Specifically, we compare the results obtained from a simulation with $\bar{\lambda} dt = 0.03$ to those from a second simulation with $\bar{\lambda} dt = 1.7$. The observations are manifolds. Firstly, although the calibration is noisy, the method of moments successfully retrieves the exponential shape of the input kernel ϕ . Secondly, while the calibration method remains consistent across both parameter sets, it appears to yield better results when $\bar{\lambda} dt \ll 1$. In this regime, the data binning to obtain σ^ϕ from N and P ensures that there is at most one event per bin, and so $\sigma^\phi \approx dN$. However, when $\bar{\lambda} dt \geq 1$, there are more events within each bin, leading to an overestimation of the activity.

For illustration, Figure 11 compares the autocovariance, C , of the synthetic activity with its theoretical values [34]. As shown in Figure 11a, when $\bar{\lambda} dt \ll 1$, the estimated autocovariance matched its theoretical values. In contrast, Figure 11b demonstrates that C is overestimated when $\bar{\lambda} dt = 1.7$. These estimation errors in C result in a biased estimation of the kernel ϕ as the Yule-Walker system to find ϕ relies solely on C .

For real data, due to the high activity of financial markets, we are situated in a regime where $\bar{\lambda} dt \geq 6$. Hence, if the underlying process is a linear Hawkes, we expect the autocovariance C to be significantly overestimated. For the linear Hawkes process, it is possible to determine a correction to apply to C to retrieve the correct theoretical value and thus to correct the bias in the resulting kernel ϕ . However, extending this correction to the quadratic case is challenging.

In comparison, as shown in Figure 12, for instance, the calibration of the univariate linear GARCH remains stable across a wide range of parameters, provided that an appropriate number of lags q is chosen (see Appendix D). A consistent choice is to take $q > \frac{3}{\beta}$.

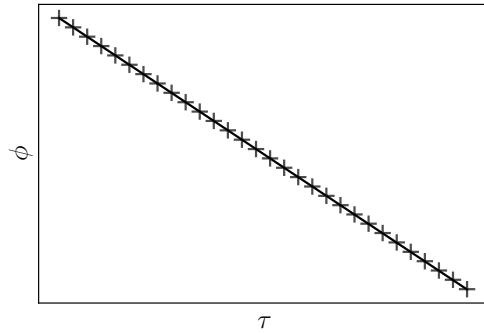


Figure 12: Proof of concept - Calibration of a 1D linear GARCH with exponential kernel. The scatter plot is the kernel obtained by the method of moments and the plain line is the kernel used to generate the synthetic data and that we want to recover. For this simulation $n_H = 0.7$, $\beta = 0.1 \text{ min}^{-1}$ and $\sigma_\infty^2 = 0.1$

Note also that the calibration is much less noisy for the GARCH model compared to the linear Hawkes model. We now turn our attention to the quadratic case.

b. Univariate quadratic case

The quadratic case, for both the QHawkes and the QGARCH, is a lot more delicate. Introduced in [6], the intensity of a QHawkes process is defined by

$$\lambda_t = \lambda_\infty + \int_{-\infty}^t \int_{-\infty}^t K(t-s, t-u) dP_s dP_u \quad \text{with returns defined as } dP_t = \pm dN_t.$$

In the case of a QGARCH process, as defined in [26], the square volatility, σ^2 , is driven by:

$$\sigma_t^2 = \sigma_\infty^2 + \sum_{i=1}^{+\infty} \sum_{j=1}^{+\infty} K(i, j) r_{t-i} r_{t-j} \quad \text{with returns defined as } r_t = \sigma_t \xi, \quad \text{with } \xi \sim \mathcal{N}(0, 1).$$

The connection between the two models remains the same as in the linear case, that is $\frac{\sigma^2}{\Delta} \equiv \lambda_t$.

In the quadratic cases, the kernel K can be understood as a matrix such that $\mathbb{K}_{ij} = K(i, j)$. The objective of the calibration is to estimate the entries of the matrix \mathbb{K} up to a certain lag q ; these entries correspond to the values $(K(i, j))_{1 \leq i, j \leq q}$. Moreover, since K is symmetric, it is only necessary to determine the upper triangle of \mathbb{K} , specifically the values $(K(i, j))_{1 \leq i \leq j \leq q}$.

Following the empirical observation of Blanc *et al.* [6], we adopt the ZHawkes framework which expresses the kernels K as the sum of a diagonal contribution and a rank-one component, meaning that $K(\tau_1, \tau_2) = \phi(\tau_1)\delta_{\tau_1-\tau_2} + k(\tau_1)k(\tau_2)$. For exponential kernels, ϕ and k take the form defined above. Note that the ZHawkes framework also greatly simplifies the simulation and limit the number of parameters to determine. Consistently with the explanation above, the method of moments, as detailed in Appendix D, returns the estimated diagonal of \mathbb{K} up to a certain lag q , that is $K(\tau, \tau)_{\tau \in [\mathbb{I}, q]}$, and its upper triangle values, $K(\tau_1, \tau_2)_{1 \leq \tau_1 < \tau_2 \leq q}$. Therefore, within the ZHawkes framework, the diagonal of \mathbb{K} ($K(\tau, \tau)_{\tau \in [\mathbb{I}, q]}$) provides an estimation of $\phi + k^2$, and the upper triangle values of \mathbb{K} ($K(\tau_1, \tau_2)_{1 \leq \tau_1 < \tau_2 \leq q}$) allow to rebuild the off-diagonal of \mathbb{K} (since \mathbb{K} is symmetric) and approximate it with a rank-one representation, so called hereafter \tilde{k} . Figure 13 shows the result of such calibration on synthetic data generated under the condition $\bar{\lambda}dt \ll 1$. Similar to the linear Hawkes process, this regime ensures that there are very few events per bin, thereby maintaining the relationship $dP = \pm dN$ and approximating $\sigma^2 \approx dN$.

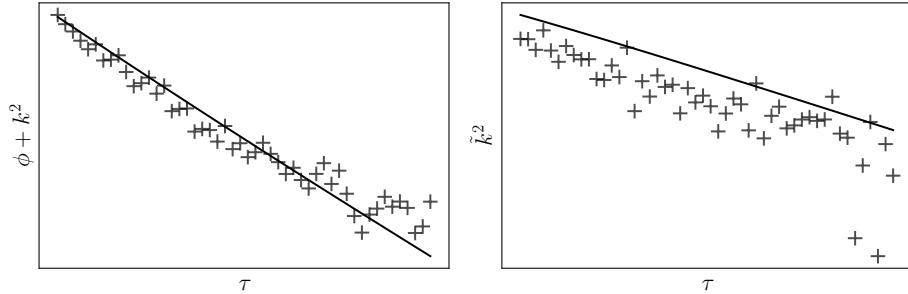


Figure 13: Calibration of a 1D QHawkes with method of moments from a simulation with parameters $n_H = 0.6$, $\beta = 0.04 \text{ min}^{-1}$, $n_Z = 0.2$, $\omega = 0.03 \text{ min}^{-1}$, $\lambda_\infty = 0.003 \text{ min}^{-1}$, hence $\bar{\lambda}dt = 0.015$. The left panel presents the diagonal of the kernel $\mathbb{K}(\tau, \tau) = \phi(\tau) + k^2(\tau)$ up to a lag $q = 50$. The right panel presents the rank-one decomposition of the off-diagonal of \mathbb{K} . In both panels the scatter plot is the component from the method of moments, while the plain line is the theoretical component we want to recover.

From the method of moments estimation, one can reconstruct the entire kernel matrix \mathbb{K} up to a lag q by combining the diagonal and the upper triangular entries, considering that \mathbb{K} is a symmetric $q \times q$ matrix. Subsequently, the parameters n_H , n_Z , β and ω can be retrieved by minimizing $[\hat{K}(\tau_1, \tau_2) - \phi(\tau_1)\delta_{\tau_1-\tau_2} - k(\tau_1)k(\tau_2)]^2$, where \hat{K} is the estimated matrix from the method of moments and ϕ and k are the parametric kernels defined above. The results of such minimisation are illustrated in Figure 14. Note that the obtained values of n_H , n_Z , β and ω can then serve as initial guesses to perform an estimation by maximum likelihood as described in Appendix G.

For the quadratic case, it is essential to remain in the regime where $\bar{\lambda}dt \ll 1$; otherwise, the calibration results deviate significantly from the input values, and the method of moments, even when combined with maximum likelihood estimation, does not yield reliable parameters. Since, for real data at the 1-minute scale, we observe that $\bar{\lambda}dt \gg 1$, properly estimating the QHawkes model as it stands is very challenging. Therefore, following the comparison and the approximation made in [6], we have turned to the QGARCH model introduced in [26]. Indeed, the method of moments to calibrate the QGARCH behaves properly for a large set of parameters as shown by Figure 15.

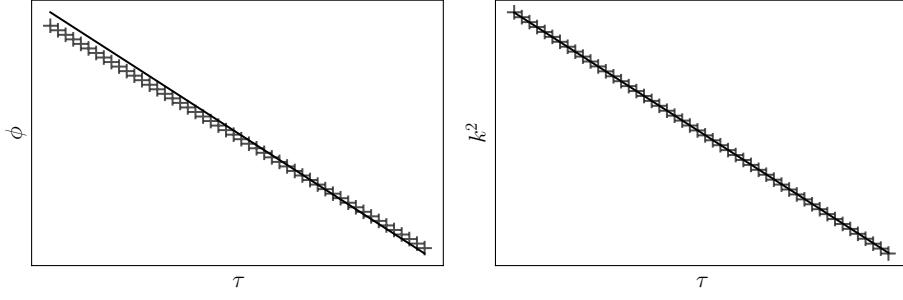


Figure 14: Calibration of a 1D QHawkes with method of moments combined with a parametric minimisation from a simulation with parameters $n_H = 0.6$, $\beta = 0.04 \text{ min}^{-1}$, $n_Z = 0.2$, $\omega = 0.03 \text{ min}^{-1}$, $\lambda_\infty = 0.003 \text{ min}^{-1}$, hence $\bar{\lambda}dt = 0.015$. The left panel presents the kernel ϕ . The right panel present the kernel k . In both panels the scatter plot is the optimised kernel obtained from the method of moments, while the plain line is the theoretical kernel with input parameters, we want to determine.

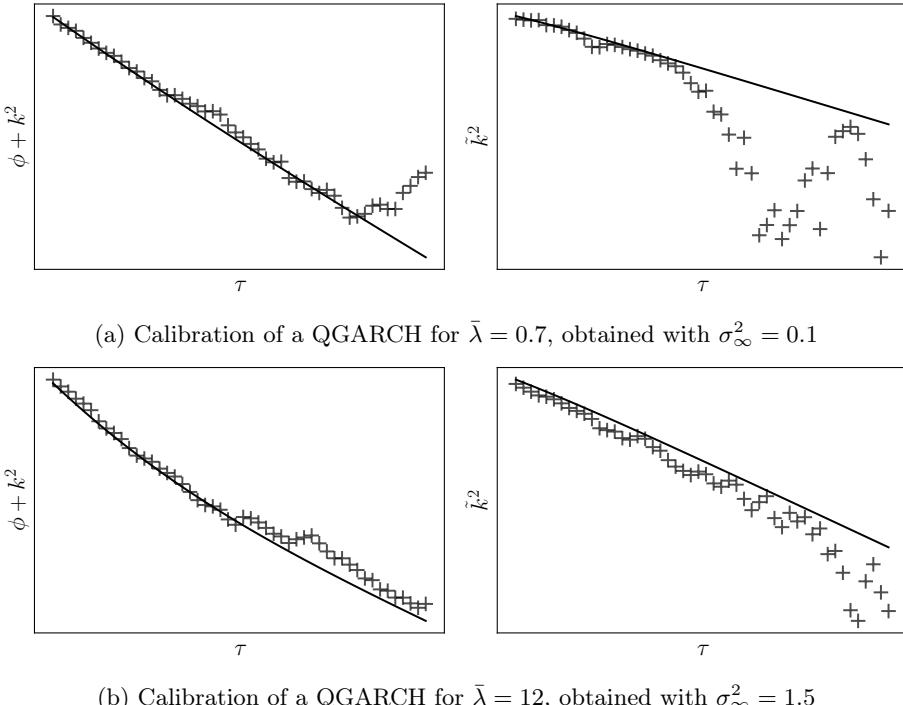


Figure 15: Proof of concept - Calibration of a 1D quadratic GARCH with exponential kernels. The scatter plot are the components obtained by the method of moments and the plain line are the theoretical components used to generate the synthetic data and that we want to recover. The left panel presents the diagonal of the kernel $K = \phi + k^2$ up to a lag $q = 50$. The right panel presents the rank one decomposition of K considering its diagonal has only zeros. For this simulation, we set $n_H = 0.7$, $n_Z = 0.2$, $\beta = 0.06$, $\omega = 0.05$ and σ_∞^2 varies between the two sub-panels.

2. Calibration of the Multivariate QGARCH

Since retrieving the univariate QHawkes from aggregated data is already challenging, for the multivariate case, we present only the calibration results on synthetic data for the MQARCH model as a proof of concept. We first examine the 2D linear GARCH, followed by the 2D quadratic GARCH model.

a. Calibration of 2D linear GARCH

We now consider 2 assets, A and B , with volatility σ_A and σ_B and returns r_A and r_B . The MQARCH volatility of the asset $i \in \{A, B\}$ is driven by

$$\sigma_{i,t}^2 = \sigma_{i,\infty}^2 + \sum_{j \in \{A, B\}} \sum_{k=1}^{+\infty} \phi_j^i(k) r_{j,t-k}^2 \quad \text{with returns defined as } r_{i,t} = \sigma_{i,t} \xi_{i,t}, \quad \text{with } \xi_{i,t} \sim \mathcal{N}(0, 1).$$

In the case of exponential kernels, for all $i, j \in \{A, B\}$, the kernel ϕ_j^i takes the form $\phi_j^i(k) = n_{Hij} \beta_{ij} \exp(-\beta_{ij} k)$. The goal of the calibration is then to determine the eight parameters $(n_{Hij}, \beta_{ij})_{i,j \in \{A, B\}}$. The method of moments works as before. Figure 16 shows our calibration results for the set of parameters presented in Table I. It clearly demonstrates that the method of moments enables to recover the input kernels in the case of the 2D linear GARCH model.

	n_H	β
X_A^A	0.8	0.2
X_B^A	0.2	0.3
X_A^B	0.3	0.3
X_B^B	0.7	0.1

Table I: Input parameters to generate the synthetic 2D GARCH as defined above. Additionally, we set $\sigma_{A,\infty}^2 = 0.05$ and $\sigma_{B,\infty}^2 = 0.1$.

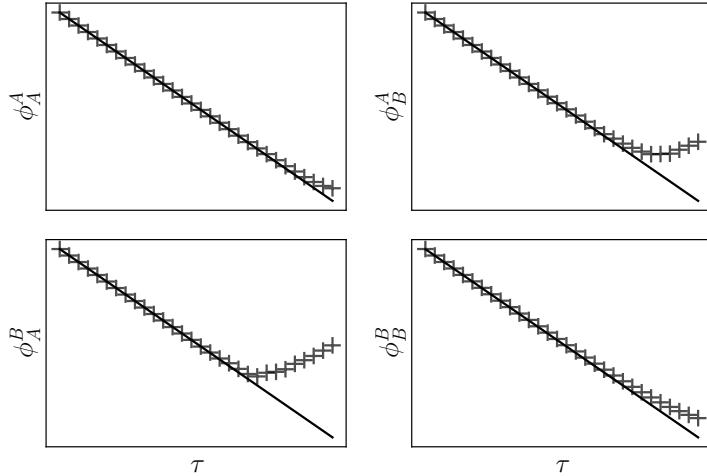


Figure 16: Proof of concept - Calibration of a 2D linear GARCH with exponential kernels. The scatter plots are the kernels obtained by the method of moments and the plain lines are the kernels used to generate the synthetic data and that we want to recover. For this simulation, the parameters are set as defined in Table I.

b. Calibration of 2D quadratic GARCH (without cross trends feedback)

This last section focuses on our model of interest the multivariate quadratic GARCH. The goal is to provide, in a simple case, a proof of concept that the method of moments can be extended to this case and give consistent results.

Here, we consider two assets A and B whose volatility $(\sigma_i)_{i \in \{A, B\}}$ is driven by

$$\sigma_{i,t}^2 = \sigma_{i,\infty}^2 + \sum_{j \in \{A, B\}} \sum_{k_1=1}^{+\infty} \sum_{k_2=1}^{+\infty} K_j^i(k_1, k_2) r_{j,t-k_1} r_{j,t-k_2},$$

with returns defined as $r_{i,t} = \sigma_{i,t} \xi_{i,t}$, with $\xi_{i,t} \sim \mathcal{N}(0, 1)$.

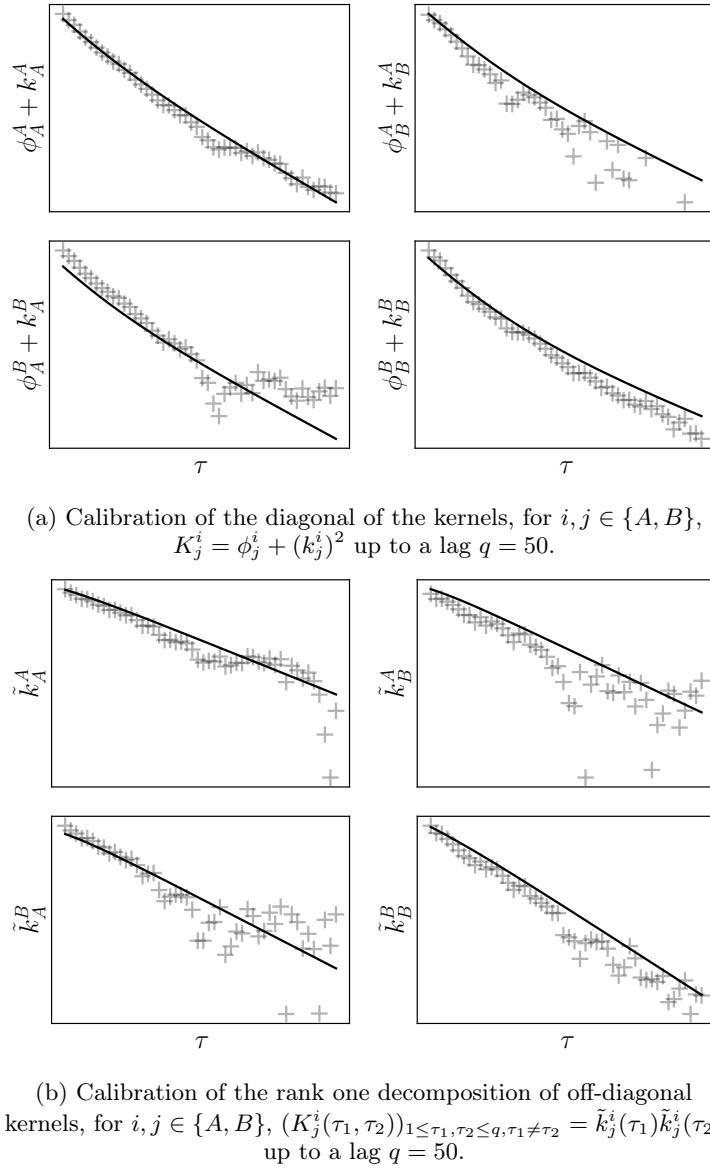


Figure 17: Proof of concept - Calibration of a 2D quadratic GARCH with exponential kernels, with no leverage or cross trend feedback. The scatter plots are the kernels obtained by the method of moments and the plain lines are the kernels used to generate the synthetic data and that we want to recover. For this simulation, the parameters are set as defined in Table II.

Note that, for the sake of simplicity, the leverage and cross-trends feedback (characterised by L and K_x in Equation (7)) are considered null in this framework. As for the univariate quadratic case above (Section A 1 b), we can adopt the ZHawkes framework [6]. Therefore, each kernel K_j^i for $i, j \in \{A, B\}$ can be decomposed as the sum of a diagonal contribution and a rank-one component, meaning that for all $i, j \in \{A, B\}$, K_j^i writes $K_j^i(\tau_1, \tau_2) = \phi_j^i(\tau_1)\delta_{\tau_1-\tau_2} + k_j^i(\tau_1)k_j^i(\tau_2)$.

For better representation, as for the univariate case, each kernel K_j^i is considered as a matrix defined by its lags, up to a certain lag q . Thus, K_j^i is represented by the symmetric matrix $(\mathbb{K}_j^i(\tau_1, \tau_2))_{1 \leq \tau_1, \tau_2 \leq q}$, where, in the ZHawkes framework, the diagonal is $(\phi_j^i(\tau) + (k_j^i)^2(\tau))_{1 \leq \tau \leq q}$ and the upper triangle entries are $(k_j^i(\tau_1)k_j^i(\tau_2))_{1 \leq \tau_1 < \tau_2 \leq q}$. As before, the method of moments (see Appendix D) returns, for all $i, j \in \{A, B\}$, the estimated diagonal and upper triangle values of \mathbb{K}_j^i up to a certain lag q . Subsequently, the off-diagonal of each kernel \mathbb{K}_j^i can be approximated with a rank-one representation, i.e. a vector \tilde{k}_j^i such that $(\mathbb{K}_j^i(\tau_1, \tau_2))_{1 \leq \tau_1, \tau_2 \leq q, \tau_1 \neq \tau_2} = \tilde{k}_j^i(\tau_1)\tilde{k}_j^i(\tau_2)$.

To provide a proof of concept of the calibration of the 2D QGARCH, we generate synthetic time series of the QGARCH according to this framework, meaning that we generated two time series of volatility along with two time series of returns such that

$$\sigma_{i,t}^2 = \sigma_{i,\infty}^2 + \sum_{j \in \{A,B\}} \sum_{k=1}^{+\infty} \phi_j^i(k) r_{j,t-k}^2 + \sum_{k_1=1}^{+\infty} \sum_{k_2=1}^{+\infty} k_j^i(k_1) k_j^i(k_2) r_{j,t-k_1} r_{j,t-k_2},$$

with returns defined as $r_{i,t} = \sigma_{i,t} \xi_{i,t}$ and $\xi_{i,t} \sim \mathcal{N}(0, 1)$, and where kernels take the form $\phi_j^i(\tau) = n_{Hij} \beta_{ij} \exp(-\beta_{ij} \tau)$ and $k_j^i(\tau) = \sqrt{2n_{Zij}\omega_{ij}} \exp(-\omega_{ij}\tau)$. The value of the parameters for the simulation are described in Table II. Subsequently, the method of moments was applied as described in Appendix D considering $L = \phi_x = k_x = 0$. The results of such calibration are shown Figure 17.

	n_H	$\beta (\text{min}^{-1})$	n_Z	$\omega (\text{min}^{-1})$
X_A^A	0.6	0.06	0.2	0.07
X_B^A	0.1	0.1	0.15	0.1
X_A^B	0.2	0.08	0.1	0.09
X_B^B	0.4	0.04	0.21	0.06

Table II: Input parameters to generate the synthetic 2D QGARCH as defined above. Additionally, we set $\sigma_{A,\infty}^2 = 1.2$ and $\sigma_{B,\infty}^2 = 0.8$.

Several observations can be made. First, the calibration appears to successfully replicate the exponential shape and amplitude of the input kernels. However, it is evident that the calibration results do not perfectly align with the input kernels, indicating room for further improvement. Enhancements could involve increasing the number of lags q set in the method of moments, as larger values tend to yield more stable results but also require more computational time. Additionally, another approach to improve the calibration results is to utilize the method of moments' outcome as the initial point for optimization by Maximum Likelihood, as detailed in Appendix G, similar to the procedure used in univariate quadratic Hawkes calibration.

Appendix B: “Surprise price”

We know, for Gaussian variables:

$$\mathbb{E}(X|Y) = \mathbb{E}(X) + \Sigma_{XY}\Sigma_Y^{-1}(Y - \mathbb{E}(Y))$$

So, if we consider returns as Gaussian, and we consider only the lag-1 has influence on the next return:

$$\mathbb{E}(\mathrm{d}P_t|\mathrm{d}P_{t-1}) = \mathbb{E}(\mathrm{d}P_t) + \Sigma_{t,t-1}\Sigma_{t-1}^{-1}(\mathrm{d}P_{t-1} - \mathbb{E}(\mathrm{d}P_{t-1}))$$

Since we centered/normalized the returns $\mathbb{E}(\mathrm{d}P_t) = \mathbb{E}(\mathrm{d}P_{t-1}) = 0$ and $\Sigma_{t-1} = 1$
and $\Sigma_{t,t-1} = \text{cov}(\mathrm{d}P_t, \mathrm{d}P_{t-1}) = \Gamma(1)$. Now in multidimension:

$$\mathbb{E} \left(\begin{pmatrix} \mathrm{d}P_t^1 \\ \mathrm{d}P_t^2 \end{pmatrix} \mid \begin{pmatrix} \mathrm{d}P_{t-1}^1 \\ \mathrm{d}P_{t-1}^2 \end{pmatrix} \right) = \mathbb{E} \left(\begin{pmatrix} \mathrm{d}P_t^1 \\ \mathrm{d}P_t^2 \end{pmatrix} \right) + \Sigma_{t,t-1}\Sigma_{t-1}^{-1} \left[\begin{pmatrix} \mathrm{d}P_{t-1}^1 \\ \mathrm{d}P_{t-1}^2 \end{pmatrix} - \mathbb{E} \left(\begin{pmatrix} \mathrm{d}P_{t-1}^1 \\ \mathrm{d}P_{t-1}^2 \end{pmatrix} \right) \right]$$

Similarly, thanks to the normalization and centralisation we have:

$$\Sigma_{t-1} = \begin{pmatrix} 1 & \text{cov}(\mathrm{d}P^1, \mathrm{d}P^2) \\ \text{cov}(\mathrm{d}P^1, \mathrm{d}P^2) & 1 \end{pmatrix}$$

we note $\text{cov}(\mathrm{d}P^1, \mathrm{d}P^2) = \nu$

$$\Sigma_{t,t-1} = \begin{pmatrix} \text{cov}(\mathrm{d}P_t^1, \mathrm{d}P_{t-1}^1) & \text{cov}(\mathrm{d}P_t^1, \mathrm{d}P_{t-1}^2) \\ \text{cov}(\mathrm{d}P_t^1, \mathrm{d}P_t^2) & \text{cov}(\mathrm{d}P_t^2, \mathrm{d}P_{t-1}^2) \end{pmatrix}$$

So we should have:

$$\begin{cases} \mathrm{d}\tilde{P}_t^1 = \mathrm{d}P_t^1 - \frac{1}{1-\nu^2} \left((\text{cov}(\mathrm{d}P_t^1, \mathrm{d}P_{t-1}^1) - \nu \text{cov}(\mathrm{d}P_t^1, \mathrm{d}P_{t-1}^2)) \mathrm{d}P_{t-1}^1 + (-\nu \text{cov}(\mathrm{d}P_t^1, \mathrm{d}P_{t-1}^1) + \text{cov}(\mathrm{d}P_t^1, \mathrm{d}P_{t-1}^2)) \mathrm{d}P_{t-1}^2 \right) \\ \mathrm{d}\tilde{P}_t^2 = \mathrm{d}P_t^2 - \frac{1}{1-\nu^2} \left((\text{cov}(\mathrm{d}P_t^2, \mathrm{d}P_{t-1}^1) - \nu \text{cov}(\mathrm{d}P_t^2, \mathrm{d}P_{t-1}^2)) \mathrm{d}P_{t-1}^1 + (-\nu \text{cov}(\mathrm{d}P_t^2, \mathrm{d}P_{t-1}^1) + \text{cov}(\mathrm{d}P_t^2, \mathrm{d}P_{t-1}^2)) \mathrm{d}P_{t-1}^2 \right) \end{cases}$$

Appendix C: Yule-Walker equations – MQARCH framework

This appendix details the calibration procedure for a Multivariate Quadratic GARCH (MQARCH) process.

1. MQARCH framework

Consistently with the notations therein, we consider a MQARCH process with $N = 2$ assets, where the squared volatility of asset j , σ_j^2 , is defined by

$$\sigma_{j,t}^2 = \sigma_{j,\infty}^2 + \sum_{i=1}^N \sum_{\tau=1}^{+\infty} L_i^j(\tau) r_{i,t-\tau} + \sum_{i,k=1}^N \sum_{\tau_1,\tau_2=1}^{+\infty} K_{ki}^j(\tau_1, \tau_2) r_{k,t-\tau_1} r_{i,t-\tau_2}. \quad (\text{C1})$$

To ease the development of the Yule-Walker system, we separate the linear, quadratic and cross kernels, then, σ_j^2 writes

$$\begin{aligned} \sigma_{j,t}^2 = & \sigma_{j,\infty}^2 \\ & + \sum_{i=1}^N \sum_{\tau=1}^{+\infty} L_i^j(\tau) r_{i,t-\tau} \\ & + \sum_{i=1}^N \sum_{\tau=1}^{+\infty} \phi_i^j(\tau) r_{i,t-\tau}^2 \\ & + 2 \sum_{i=1}^2 \sum_{\tau_1=1}^{+\infty} \sum_{\tau_2=\tau_1+1}^{+\infty} K_i^j(\tau_1, \tau_2) r_{i,t-\tau_1} r_{i,t-\tau_2} \\ & + \sum_{\tau=1}^{+\infty} \phi_{\times}^j(\tau) (r_{j,t-\tau} r_{\bar{j},t-\tau} - \bar{r}_j \bar{r}_{\bar{j}}) \\ & + \sum_{\tau_1=1}^{+\infty} \sum_{\tau_2=\tau_1+1}^{+\infty} k_{\times}^j(\tau_1, \tau_2) r_{j,t-\tau_1} r_{\bar{j},t-\tau_2} \\ & + \sum_{\tau_1=2}^{+\infty} \sum_{\tau_2=1}^{\tau_1-1} k_{\times}^j(\tau_1, \tau_2) r_{j,t-\tau_1} r_{\bar{j},t-\tau_2}, \end{aligned} \quad (\text{C2})$$

Proof. The transformation of the quadratic term (for $k = i$ in Eq. (C1)) goes as follows

$$\begin{aligned} \sum_{\tau_1,\tau_2=1}^{+\infty} K_i^j(\tau_1, \tau_2) r_{i,t-\tau_1} r_{i,t-\tau_2} &= \sum_{\tau=1}^{+\infty} \phi_i^j(\tau) r_{i,t-\tau}^2 \\ &+ \sum_{\tau_1=2}^{+\infty} \sum_{\tau_2=1}^{\tau_1-1} K_i^j(\tau_1, \tau_2) r_{i,t-\tau_1} r_{i,t-\tau_2} + \sum_{\tau_1=1}^{+\infty} \sum_{\tau_2=\tau_1+1}^{+\infty} K_i^j(\tau_1, \tau_2) r_{i,t-\tau_1} r_{i,t-\tau_2} \\ &= \sum_{\tau=1}^{+\infty} \phi_i^j(\tau) r_{i,t-\tau}^2 + \sum_{\tau_1=2}^{+\infty} \sum_{\tau_2=1}^{\tau_1-1} K_i^j(\tau_1, \tau_2) r_{i,t-\tau_1} r_{i,t-\tau_2} + \sum_{\tau_2=2}^{+\infty} \sum_{\tau_1=1}^{\tau_2-1} K_i^j(\tau_1, \tau_2) r_{i,t-\tau_1} r_{i,t-\tau_2} \\ &\text{thanks to the properties of symmetry,} \\ &= \sum_{\tau=1}^{+\infty} \phi_i^j(\tau) r_{i,t-\tau}^2 + 2 \sum_{\tau_1=2}^{+\infty} \sum_{\tau_2=1}^{\tau_1-1} K_i^j(\tau_1, \tau_2) r_{i,t-\tau_1} r_{i,t-\tau_2} \end{aligned}$$

□

Additionally, if we consider the formulation of Equation (C2), then the mean volatility meets the following condition

$$\overline{\boldsymbol{\sigma}^2} = \boldsymbol{\sigma}_{\infty}^2 + \left(\sum_{\tau=1}^{+\infty} \begin{pmatrix} \phi_1^1 & \phi_2^1 \\ \phi_1^2 & \phi_2^2 \end{pmatrix} (\tau) \right) \overline{\mathbf{r}^2} \iff \overline{\boldsymbol{\sigma}^2} = \boldsymbol{\sigma}_{\infty}^2 + \left(\sum_{\tau=1}^{+\infty} \begin{pmatrix} \phi_1^1 & \phi_2^1 \\ \phi_1^2 & \phi_2^2 \end{pmatrix} (\tau) \right) \overline{\boldsymbol{\sigma}^2}$$

2. Yule-Walker equations for a 4-step calibration

For more stability, we undertake a calibration of the model Eq. (C2) in 4 steps.

During the first three steps, we consider the leverage kernel to be null and aim to determine the other kernels $((\phi_i^j)_{i,j \in \{A,B\}})$, $((K_i^j)_{i,j \in \{A,B\}})$, $((\phi_\times^j)_{j \in \{A,B\}})$ and $((k_\times^j)_{j \in \{1,2\}})$. Therefore, to exclude the leverage effect from real data, we calibrate the linear, quadratic and cross kernels on “symmetrised” data, meaning that we use the datasets $[\sigma, \sigma]$ and $[r, -r]$ where σ and r represent our original datasets. Once step 3 is completed, we use the original datasets, in which the leverage effect is present, to determine the leverage kernels. The calibration is thus decomposed as follows:

1. Determine the QGARCH 1D of the 2 processes $((\phi_i^i)_{i \in \{1,2\}})$ and $((K_i^i)_{i \in \{1,2\}})$
2. Determine the QGARCH 2D cross linear and quadratic kernels $((\phi_j^i)_{i,j \in \{1,2\}, i \neq j})$ and $((K_j^i)_{i,j \in \{1,2\}, i \neq j})$
3. Determine the cross kernels of the QGARCH 2D $((\phi_\times^i)_{i \in \{1,2\}})$ and $((k_\times^i)_{i \in \{1,2\}})$
4. Determine the QGARCH 2D Leverage kernels $((L_i^i)_{i \in \{1,2\}})$.

In order to stay agnostic on the shape of the kernels, for each step, we implement a general method of moments. We rely on a system of Yule-Walker type equations which we describe subsequently. Further details can be found in [35].

a. Covariance structure

The Yule-Walker system of equations and thus, the method of moments, rely on the following covariance structures:

$$\mathcal{C}_{ij}(\tau) = \mathbb{E}(\sigma_{i,t}^2 r_{j,t-\tau}^2) - \bar{\sigma}_i^2 \bar{r}_j^2 \quad (\text{C3})$$

$$\mathcal{C}_{ij}^r(\tau) = \mathbb{E}(r_{i,t}^2 r_{j,s}^2) - \bar{r}_i^2 \bar{r}_j^2 \quad (\text{C4})$$

$$\mathcal{D}_{ijk}(\tau_1, \tau_2) = \mathbb{E}(\sigma_{i,t}^2 r_{j,t-\tau_1} r_{k,t-\tau_2}) \quad (\text{C5})$$

$$\mathcal{D}_{p,(ij)kl}(\tau) = \mathbb{E}((r_{i,t} r_{j,t}) r_{k,t-\tau_1} r_{l,t-\tau_2}) \quad (\text{C6})$$

$$\mathcal{D}_{\times,jAB}(\tau_1, \tau_2) = \mathbb{E}(\sigma_{j,t}^2 r_{A,t-\tau_1} r_{B,t-\tau_2}) \quad (\text{C7})$$

$$\mathcal{V}_{ij}(\tau) = \mathbb{E}(\sigma_{i,t}^2 r_{j,t-\tau}) \quad (\text{C8})$$

$$\mathcal{V}_{ij}^r(\tau) = \mathbb{E}(r_{i,t}^2 r_{j,t-\tau}) \quad (\text{C9})$$

By definition of the model and in simulation, Eq. (C3) and Eq. (C4) should result in the same values. However in practice, there is a slight difference due to noise, which influence the calibration. Hence, in order to give reproducible guidelines we note the difference in what follows.

b. Yule-Walker system for QGARCH 2D without cross terms (steps 1 and 2)

For the first two steps of the calibration, we need equations linking \mathcal{C} and \mathcal{D} to the kernels ϕ and K . The goal is to characterise $((\phi_i^j)_{i,j \in \{1,2\}})$ and $((K_i^j)_{i,j \in \{1,2\}})$ up to a certain lag q . Hence, we want to find the values of, for $i, j \in \{1, 2\}$, $\phi_i^j(\tau)_{\tau \in [\mathbb{1}, q]}$ and $K_i^j(\tau_1, \tau_2)_{1 \leq \tau_1 < \tau_2 \leq q}$ (since K is symmetric, it is enough to find only its upper triangle entries). The Yule Walker equations on $(\mathcal{C}_{ij}(\tau))_{i,j \in \{1,2\}, \tau \in [\mathbb{1}, q]}$ will give $4 \times q$ equations and the Yule Walker equations on $(\mathcal{D}_{ij}(\tau_1, \tau_2))_{i,j \in \{1,2\}, 1 \leq \tau_1 < \tau_2 \leq q}$ will give $4 \times q \frac{q-1}{2}$ equations, allowing to fully characterise $((\phi_i^j)_{i,j \in \{1,2\}})$ and $((K_i^j)_{i,j \in \{1,2\}})$ up to a certain lag q .

Without complicating things, we develop the general equations for N assets.

Lemma 1 (Yule-Walker equation for \mathcal{C}). *for $\tau > 0$, in the framework described by Eq. (C2), with $\phi^\times = k^\times = 0$, the covariance \mathcal{C} writes*

$$\mathcal{C}_{jl}(\tau) = \sum_{i=1}^N \sum_{k=1}^{+\infty} \phi_i^j(k) \mathcal{C}_{il}^r(\tau - k) + 2 \sum_{i=1}^N \sum_{k_2=\tau+1}^{+\infty} \sum_{k_1=k_2+1}^{+\infty} K_i^j(k_1, k_2) \mathcal{D}_{p,lli}(k_1 - \tau, k_2 - \tau)$$

Lemma 2 (Yule-Walker equation for \mathcal{D}). *For $\tau_1 < \tau_2$, in the framework described by Eq. (C2), with $\phi^\times = k^\times = 0$, the covariance \mathcal{D} writes*

$$\mathcal{D}_{jll}(\tau_1, \tau_2) = \sum_{i=1}^N \sum_{k=1}^{\tau_1-1} \phi_i^j(k) \mathcal{D}_{p,ill}(\tau_1 - k, \tau_2 - k) + 2 \sum_{i=1}^N \sum_{k_2=\tau_1+1}^{+\infty} K_i^j(\tau_1, k_2) \mathcal{D}_{p,(il)il}(-\tau_1 + k_2, \tau_2 - \tau_1)$$

c. Yule-Walker system for QGARCH 2D with cross terms (step 3)

The goal of the third step is to characterise $(\phi_{\times}^j)_{j \in \{1,2\}}$ and $(K_{\times}^j)_{j \in \{1,2\}}$ up to a certain lag q . In this case, we need $2 \times q$ equations to characterise $(\phi_{\times}^j)_{j \in \{1,2\}}$ and $2 \times q(q-1)$ equations to characterise $(K_{\times}^j)_{j \in \{1,2\}}$, as K_{\times}^j is not symmetric so we need both the upper and lower triangle entries. The equations on $(\mathcal{D}_{\times}^j(\tau_1, \tau_2))_{j \in \{1,2\}, 1 \leq \tau_1, \tau_2 \leq q}$, developed subsequently, allow to fully characterise the cross kernels $(\phi_{\times}^j)_{j \in \{1,2\}}$ and $(K_{\times}^j)_{j \in \{1,2\}}$.

Lemma 3 (Yule-Walker equation for \mathcal{D}_{\times} for $\tau_1 = \tau_2 = \tau$). *for $\tau > 0$, in the framework described by Eq. (C2), the covariance \mathcal{D}_{\times} writes*

$$\begin{aligned} \mathcal{D}_{\times,j}(\tau, \tau) = & \sigma_{j,\infty}^2 \bar{r}_j \bar{r}_{\bar{j}} \\ & + \sum_{i=1}^N \sum_{k=1}^{+\infty} \phi_i^j(k) \mathcal{D}_{p,(ii)j\bar{j}}(\tau - k, \tau - k) \\ & + 2 \sum_{i=1}^N \sum_{k_1=\tau+1}^{+\infty} \sum_{k_2=k_1+1}^{+\infty} K_i^j(k_1, k_2) \mathcal{D}_{p,(j\bar{j})ii}(k_1 - \tau, k_2 - \tau) \\ & + \sum_{k=1}^{+\infty} \phi_{\times}^j(k) (\mathcal{D}_{p(j\bar{j})j\bar{j}}(\tau - k, \tau - k) - \bar{r}_j \bar{r}_{\bar{j}}^2) \\ & + \sum_{k_2=\tau+1}^{+\infty} \sum_{k_1=k_2+1}^{+\infty} k_{\times}^j(k_1, k_2) \mathcal{D}_{p,(j\bar{j})j\bar{j}}(k_1 - \tau, k_2 - \tau) \\ & + \sum_{k_1=\tau+1}^{+\infty} \sum_{k_2=k_1+1}^{+\infty} k_{\times}^j(k_1, k_2) \mathcal{D}_{p,(j\bar{j})j\bar{j}}(k_1 - \tau, k_2 - \tau) \end{aligned}$$

One of the difficulty in including the cross terms is the non-symmetry of the covariance structure \mathcal{D}_{\times} and of the kernels k_{\times} . Hence, we need to develop equations for $\mathcal{D}_{\times,j}(\tau_1, \tau_2)$ for $\tau_1 < \tau_2$ and $\tau_1 > \tau_2$.

Lemma 4 (Yule-Walker equation for \mathcal{D}_{\times} for $\tau_1 < \tau_2$). *for $\tau_1 < \tau_2$, in the framework described by Eq. (C2), the covariance \mathcal{D}_{\times} writes*

$$\begin{aligned} \mathcal{D}_{\times,j}(\tau_1, \tau_2) = & \sigma_{j,\infty}^2 \bar{r}_j \bar{r}_{\bar{j}} \\ & + \sum_{i=1}^N \sum_{k=1}^{\tau_1-1} \phi_i^j(k) \mathcal{D}_{p,(ii)j\bar{j}}(\tau_1 - k, \tau_2 - k) \\ & + 2 \sum_{i=1}^N \sum_{k_2=\tau_1+1}^{+\infty} K_i^j(\tau_1, k_2) \mathcal{D}_{p,(ij)i\bar{j}}(k_2 - \tau_1, \tau_2 - \tau_1) \\ & + \sum_{k=1}^{\tau_1-1} \phi_{\times}^j(k) (\mathcal{D}_{p(j\bar{j})j\bar{j}}(\tau_1 - k, \tau_2 - k) - \bar{r}_j \bar{r}_{\bar{j}}^2) \\ & + \sum_{k_1=\tau_1+1}^{+\infty} k_{\times}^j(k_1, \tau_1) \mathcal{D}_{p,(j\bar{j})j\bar{j}}(k_1 - \tau_1, \tau_2 - \tau_1) \\ & + \sum_{k_2=\tau_1+1}^{+\infty} k_{\times}^j(\tau_1, k_2) \mathcal{D}_{p,(j\bar{j})j\bar{j}}(k_2 - \tau_1, \tau_2 - \tau_1) \end{aligned}$$

Lemma 5 (Yule-Walker equation for \mathcal{D}_{\times} for $\tau_1 > \tau_2$). *for $\tau_1 > \tau_2$, in the framework described by Eq. (C2), the covariance*

\mathcal{D}_\times writes

$$\begin{aligned}
 \mathcal{D}_{\times,j}(\tau_1, \tau_2) = & \sigma_{j,\infty}^2 \bar{r}_j \bar{r}_{\bar{j}} \\
 & + \sum_{i=1}^N \sum_{k=1}^{\tau_1-1} \phi_i^j(k) \mathcal{D}_{p,(ii)j\bar{j}}(\tau_1 - k, \tau_2 - k) \\
 & + 2 \sum_{i=1}^N \sum_{k_2=\tau_1+1}^{+\infty} K_i^j(\tau_1, k_2) \mathcal{D}_{p,(i\bar{j})ij}(k_2 - \tau_1, \tau_1 - \tau_2) \\
 & + \sum_{k=1}^{\tau_1-1} \phi_\times^j(k) (\mathcal{D}_{(j\bar{j})j\bar{j}}(\tau_1 - k, \tau_2 - k) - \bar{r}_j \bar{r}_{\bar{j}}^2) \\
 & + \sum_{k_1=\tau_2+1}^{+\infty} k_\times^j(k_1, \tau_2) \mathcal{D}_{p,(j\bar{j})jj}(k_1 - \tau_2, \tau_1 - \tau_2) \\
 & + \sum_{k_2=\tau_1+1}^{+\infty} k_\times^j(\tau_1, k_2) \mathcal{D}_{p,(j\bar{j})\bar{j}j}(k_2 - \tau_1, \tau_1 - \tau_2)
 \end{aligned}$$

Appendix D gives more details on who to concretely implement those systems of linear equations between kernels and covariance structures.

d. Yule-Walker system for QGARCH 2D for Leverage kernel (step 4)

Finally, once the linear, quadratic and cross kernels are determined, we can work on the original dataset (not symmetrised) and characterised the leverage kernels. To achieve this, additional Yule-Walker equations are required relying on \mathcal{V} (defined in Equations (C8) and (C9)), we derive them subsequently. The goal is to characterise $(L_i^j)_{i,j \in \{1,2\}}$ up to a certain lag q . Hence, we need $4 \times q$ equations which are obtained using $(\mathcal{V}_{jl}(\tau))_{j,l \in \{1,2\}, 1 \leq \tau \leq q}$.

Lemma 6 (Yule-Walker equation for \mathcal{V}). *for $\tau > 0$, in the framework described by Eq. (C2), the covariance \mathcal{V} writes*

$$\begin{aligned}
 \mathcal{V}_{ji}(\tau) = & L_i^j(\tau) \bar{r}_i \bar{r}_l \\
 & + \sum_{i=1}^N \sum_{k=1}^{\tau} \phi_i^j(k) \mathcal{V}_{il}^r(\tau - k) \\
 & + 2 \sum_{k_2=\tau+1}^{+\infty} K_i^j(\tau, k_2) \mathcal{V}_{(il)j}^r(k_2 - \tau) \\
 & + \sum_{k=1}^{\tau} \phi_\times^j(k) \mathcal{V}_{(j\bar{j})l}^r(\tau - k) \\
 & + \sum_{k_2=\tau+1}^{+\infty} k_\times^j(\tau, k_2) \mathcal{V}_{(jl)\bar{j}}^r(k_2 - \tau) \\
 & + \sum_{k_1=\tau+1}^{+\infty} k_\times^j(k_1, \tau) \mathcal{V}_{(jl)j}^r(k_1 - \tau)
 \end{aligned}$$

Appendix D: Calibration Matrices - QGARCH

The goal of the calibration up to step 3 is to deduce the kernels $(\phi_j^i)_{i,j \in \{1,2\}}$, $(K_j^i)_{i,j \in \{1,2\}}$, $(\phi_\times^i)_{i \in \{1,2\}}$ and $(k_\times^i)_{i \in \{1,2\}}$ that weight the contributions of the past on the squared volatility as defined in Eq. (C2), from the observation of the covariance structures defined in Appendix C 2a. The link between the kernels and the covariance structures is established by the Yule-Walker equations developed in Appendix C. To implement this Yule-Walker system of equations, we propose here a matricel environment. The calibration matrices to deduce the leverage kernels are described in a subsequent section.

1. 2D QGARCH – framework

Consistently, we present a matricel system for the 2D QGARCH, considering the two assets 1 and 2. The goal is to link the kernels

$$\emptyset = \begin{pmatrix} \phi_1^1 & \phi_2^1 \\ \phi_1^2 & \phi_2^2 \end{pmatrix}, \quad \mathbb{K} = \begin{pmatrix} K_1^1 & K_2^1 \\ K_1^2 & K_2^2 \end{pmatrix}, \quad \phi_\times = \begin{pmatrix} \phi_\times^1 \\ \phi_\times^2 \end{pmatrix}, \quad \text{and} \quad \mathbf{k}_\times = \begin{pmatrix} k_\times^1 \\ k_\times^2 \end{pmatrix}$$

to the observable covariance structures

$$\mathbb{C} = \begin{pmatrix} \mathcal{C}_1^1 & \mathcal{C}_2^1 \\ \mathcal{C}_1^2 & \mathcal{C}_2^2 \end{pmatrix}, \quad \mathbb{D}_d = \begin{pmatrix} \mathcal{D}_{111} & \mathcal{D}_{122} \\ \mathcal{D}_{211} & \mathcal{D}_{222} \end{pmatrix} \quad \text{and} \quad \mathcal{D}_\times = \begin{pmatrix} \mathcal{D}_{112} \\ \mathcal{D}_{212} \end{pmatrix}.$$

The goal of this appendix is to describe a system of tensors which describes the relationship between the kernels and the covariance structures up to a certain lag q . We construct the following system:

$$\begin{pmatrix} \mathbb{A}_{1d} & \mathbb{A}_{2d} & 0 & 0 \\ \mathbb{A}_{3d} & \mathbb{A}_{4d} & 0 & 0 \\ \mathbb{A}_{1d\times} & \mathbb{A}_{2d\times} & \mathbb{A}_{1\times} & \mathbb{A}_{2\times} \\ \mathbb{A}_{3d\times} & \mathbb{A}_{4d\times} & \mathbb{A}_{3\times} & \mathbb{A}_{4\times} \end{pmatrix} \begin{pmatrix} \emptyset(\tau) \\ \mathbb{K}(\tau_1, \tau_2)_{\tau_1 < \tau_2} \\ \phi_\times(\tau) \\ \mathbf{k}_\times(\tau_1, \tau_2)_{\tau_1 < \tau_2} \\ \mathbf{k}_\times(\tau_1, \tau_2)_{\tau_1 > \tau_2} \end{pmatrix} = \begin{pmatrix} \mathbb{C}(\tau) \\ \mathbb{D}_d(\tau_1, \tau_1)_{\tau_1 < \tau_2} \\ \mathcal{D}_\times(\tau, \tau) \\ \mathcal{D}_\times(\tau_1, \tau_2)_{\tau_1 < \tau_2} \\ \mathcal{D}_\times(\tau_1, \tau_2)_{\tau_1 > \tau_2} \end{pmatrix}. \quad (\text{D1})$$

Another way to comprehend the system is to see that \emptyset is the time-diagonal of K and thus, one can separate the cross components from the linear and quadratic components, as follows:

$$\begin{pmatrix} \mathbb{A}_d & 0 \\ \mathbb{A}_{d\times} & \mathbb{A}_\times \end{pmatrix} \begin{pmatrix} \mathbb{K}_d \\ \mathbf{K}_\times \end{pmatrix} = \begin{pmatrix} \mathbb{D}_d \\ \mathcal{D}_\times \end{pmatrix}.$$

Let us describe the tensors system above:

- **Covariance structures:** Represented by the right hand in Eq. (D1), the covariance tensors are composed of observable elements.
 - The **red matrices** are correlation tensors of size $q \times 2 \times 2$ (or more generally $q \times N \times N$ if N is the number of assets). They represent the covariance of the squared volatility and the covariance between the volatility and the past trends.
 - The **magenta tensors** are also covariance structures, capturing the influence of the past *cross* trends on future volatility. The time-diagonal $\mathcal{D}_\times(\tau, \tau)$, with $\tau \in \llbracket 1, q \rrbracket$, is a tensor of 2D vectors with shape $q \times 2 \times 1$. Conversely to \mathbb{D}_d , \mathcal{D}_\times is non-time-symmetric, i.e., $\mathcal{D}_\times(\tau_1, \tau_2) \neq \mathcal{D}_\times(\tau_2, \tau_1)$, thus, we need to consider $\mathcal{D}_\times(\tau_1, \tau_2)$ for both $\tau_1 < \tau_2$ and $\tau_2 < \tau_1$, both of which shape $q \frac{q-1}{2} \times 2 \times 1$.
- **Kernels:** The kernels weight the feedback of past realisations on the future square volatility and characterise the process. These kernels are the unknowns we want to recover.
 - The **blue tensors** represent the linear and quadratic we want to calibrate (steps 1 and 2 of Appendix C). The shape of \mathbb{K}_d is, consistently with \mathbb{D}_d , $q + \frac{q(q-1)}{2} \times 2 \times 2$ (\emptyset has shape $q \times 2 \times 2$ and \mathbb{K} has shape $\frac{q(q-1)}{2} \times 2 \times 2$).
 - The **cyan tensors** are the lags of the kernels weighting the feedback of the cross trend. Consistently with \mathcal{D}_\times , their shape is $q + q(q-1) \times 2 \times 1$.
- **Yule-Walker matrices:** These matrices are built upon observables according to the Yule-Walker system of equations developed in Appendix C. We explain briefly here their role, but the subsequent sections are dedicated to describe them precisely.

A_1	A_2	A_3	A_4
\mathbb{A}_{1d}	\mathbb{A}_{2d}	\mathbb{A}_{3d}	\mathbb{A}_{4d}
$\mathbb{A}_{1d\times}$	$\mathbb{A}_{2d\times}$	$\mathbb{A}_{3d\times}$	$\mathbb{A}_{4d\times}$
$\mathbb{A}_{1\times}$	$\mathbb{A}_{2\times}$	$\mathbb{A}_{3\times}$	$\mathbb{A}_{4\times}$

Table III: Calibration matrices with similar structure

- The **violet calibration matrices** relate the linear and quadratic kernels represented by \mathbb{K}_d to the correlations structure \mathbb{D}_d .
- The **teal calibration matrices** relate the kernel of the cross trends \mathbf{K}_\times to the cross correlations \mathcal{D}_\times .
- the **olive components** relate the linear and quadratic kernel \mathbb{K}_d to the cross correlations \mathcal{D}_\times .

The above system already considers a calibration in several steps, otherwise the upper right bloc would not be 0 as \mathbf{K}_\times could contribute to \mathbb{D}_d . Thus, we first solve the upper bloc for \mathbb{K}_d using $\mathbb{A}_d \mathbb{K}_d = \mathbb{D}_d$, and then we solve for \mathbf{K}_\times using $\mathbb{A}_{d\times} \mathbb{K}_d + \mathbb{A}_\times \mathbf{K}_\times = \mathcal{D}_\times$. Let us note that the first calibration step to determine \mathbb{K}_d is also decomposed in two, as we first characterise the self feedback (1D calibration) and then the cross-linear and quadratic feedback. The rest of this appendix is dedicated to the description of the the Yule-Walker matrices which define the system.

2. Building the Yule-Walker Matrices

The Yule-Walker matrices \mathbb{A}_d , $\mathbb{A}_{d\times}$ and \mathbb{A}_\times are built according to the Yule-Walker lemmas of Appendix C. In fact, each matrix (\mathbb{A}_{1d} , \mathbb{A}_{2d} , \mathbb{A}_{3d} , \mathbb{A}_{4d} , $\mathbb{A}_{1d\times}$, $\mathbb{A}_{2d\times}$, $\mathbb{A}_{3d\times}$, $\mathbb{A}_{4d\times}$, $\mathbb{A}_{1\times}$, $\mathbb{A}_{2\times}$, $\mathbb{A}_{3\times}$, $\mathbb{A}_{4\times}$) replicates one term of these equations. Upon scrutiny, it appears that some terms have the same structure, and only the components, i.e. the kernels and the covariance involved in the term, change. In particular, there are 4 different structures to implement: A_1 , A_2 , A_3 and A_4 .

Specifically, restating the Equations of lemmas 1, 2, 3, 4 and 5, we make the link between the matrices and the term they represent.

From lemma 1, we have

$$\mathcal{C}_{jl}(\tau) = \underbrace{\sum_{i=1}^N \sum_{k=1, k \neq \tau}^{+\infty} \phi_i^j(k) \mathcal{C}_{il}^r(\tau - k)}_{\mathbb{A}_{1d} \in \mathbb{A}_d} + \underbrace{2 \sum_{i=1}^N \sum_{k_2=\tau+1}^{+\infty} \sum_{k_1=k_2+1}^{+\infty} K_i^j(k_1, k_2) \mathcal{D}_{p,lii}(k_1 - \tau, k_2 - \tau)}_{\mathbb{A}_{2d} \in \mathbb{A}_d}.$$

From lemma 2, we have, for $\tau_1 < \tau_2$,

$$\mathcal{D}_{jll}(\tau_1, \tau_2) = \underbrace{\sum_{i=1}^N \sum_{k=1}^{\tau_1-1} \phi_i^j(k) \mathcal{D}_{p,ill}(\tau_1 - k, \tau_2 - k)}_{\mathbb{A}_{3d} \in \mathbb{A}_d} + \underbrace{2 \sum_{i=1}^N \sum_{k_2=\tau_1+1}^{+\infty} K_i^j(\tau_1, k_2) \mathcal{D}_{p,(il)il}(-\tau_1 + k_2, \tau_2 - \tau_1)}_{\mathbb{A}_{4d} \in \mathbb{A}_d}.$$

From lemma 3, we have, for, $\tau > 0$

$$\begin{aligned} \mathcal{D}_{\times,j}(\tau, \tau) - \sigma_{j,\infty}^2 \bar{r}_j \bar{r}_{\bar{j}} &= \underbrace{\sum_{i=1}^N \sum_{k=1}^{\tau-1} \phi_i^j(k) \mathcal{D}_{p,(ii)j\bar{j}}(\tau - k, \tau - k)}_{\mathbb{A}_{1d\times} \in \mathbb{A}_{d\times}} \\ &\quad + \underbrace{2 \sum_{i=1}^N \sum_{k_1=\tau+1}^{+\infty} \sum_{k_2=k_1+1}^{+\infty} K_i^j(k_1, k_2) \mathcal{D}_{p,(j\bar{j})ii}(k_1 - \tau, k_2 - \tau)}_{\mathbb{A}_{2d\times} \in \mathbb{A}_{d\times}} \\ &\quad + \underbrace{\sum_{k=1}^{\tau-1} \phi_\times^j(k) (\mathcal{D}_{p(j\bar{j})j\bar{j}}(\tau - k, \tau - k) - \bar{r}_j \bar{r}_{\bar{j}}^{-2})}_{\mathbb{A}_{1\times} \in \mathbb{A}_\times} \\ &\quad + \underbrace{\sum_{k_2=\tau+1}^{+\infty} \sum_{k_1=k_2+1}^{+\infty} k_\times^j(k_1, k_2) \mathcal{D}_{p,(j\bar{j})j\bar{j}}(k_1 - \tau, k_2 - \tau)}_{\mathbb{A}_{2\times} \in \mathbb{A}_\times} \end{aligned}$$

$$+ \underbrace{\sum_{k_1=\tau+1}^{+\infty} \sum_{k_2=k_1+1}^{+\infty} k_{\times}^j(k_1, k_2) \mathcal{D}_{p, (j\bar{j})j\bar{j}}(k_1 - \tau, k_2 - \tau)}_{\mathbb{A}_{2\times} \in \mathbb{A}_{\times}}$$

From lemma 4, we have, for, $\tau_1 < \tau_2$

$$\begin{aligned} \mathcal{D}_{\times, j}(\tau_1, \tau_2) - \sigma_{j, \infty}^2 \overline{r_j r_{\bar{j}}} &= \underbrace{\sum_{i=1}^N \sum_{k=1}^{\tau_1-1} \phi_i^j(k) \mathcal{D}_{p, (ii)j\bar{j}}(\tau_1 - k, \tau_2 - k)}_{\mathbb{A}_{3d\times} \in \mathbb{A}_{d\times}} \\ &+ \underbrace{2 \sum_{i=1}^N \sum_{k_2=\tau_1+1}^{+\infty} K_i^j(\tau_1, k_2) \mathcal{D}_{p, (ij)i\bar{j}}(k_2 - \tau_1, \tau_2 - \tau_1)}_{\mathbb{A}_{4d\times} \in \mathbb{A}_{d\times}} \\ &+ \underbrace{\sum_{k=1}^{\tau_1-1} \phi_{\times}^j(k) (\mathcal{D}_{(j\bar{j})j\bar{j}}(\tau_1 - k, \tau_2 - k) - \overline{r_j r_{\bar{j}}^2})}_{\mathbb{A}_{3\times} \in \mathbb{A}_{\times}} \\ &+ \underbrace{\sum_{k_1=\tau_1+1}^{+\infty} k_{\times}^j(k_1, \tau_1) \mathcal{D}_{p, (\bar{j}j)\bar{j}\bar{j}}(k_1 - \tau_1, \tau_2 - \tau_1)}_{\mathbb{A}_{4\times} \in \mathbb{A}_{\times}} \\ &+ \underbrace{\sum_{k_2=\tau_1+1}^{+\infty} k_{\times}^j(\tau_1, k_2) \mathcal{D}_{p, (jj)\bar{j}\bar{j}}(k_2 - \tau_1, \tau_2 - \tau_1)}_{\mathbb{A}_{4\times} \in \mathbb{A}_{\times}} \end{aligned}$$

From lemma 5, we have, for, $\tau_1 > \tau_2$

$$\begin{aligned} \mathcal{D}_{\times, j}(\tau_1, \tau_2) - \sigma_{j, \infty}^2 \overline{r_j r_{\bar{j}}} &= \underbrace{\sum_{i=1}^N \sum_{k=1}^{\tau_1-1} \phi_i^j(k) \mathcal{D}_{p, (ii)j\bar{j}}(\tau_1 - k, \tau_2 - k)}_{\mathbb{A}_{3d\times} \in \mathbb{A}_{d\times}} \\ &+ \underbrace{2 \sum_{i=1}^N \sum_{k_2=\tau_1+1}^{+\infty} K_i^j(\tau_1, k_2) \mathcal{D}_{p, (i\bar{j})i\bar{j}}(k_2 - \tau_1, \tau_1 - \tau_2)}_{\mathbb{A}_{4d\times} \in \mathbb{A}_{d\times}} \\ &+ \underbrace{\sum_{k=1}^{\tau_1-1} \phi_{\times}^j(k) (\mathcal{D}_{(j\bar{j})j\bar{j}}(\tau_1 - k, \tau_2 - k) - \overline{r_j r_{\bar{j}}^2})}_{\mathbb{A}_{3\times} \in \mathbb{A}_{\times}} \\ &+ \underbrace{\sum_{k_1=\tau_2+1}^{+\infty} k_{\times}^j(k_1, \tau_2) \mathcal{D}_{p, (\bar{j}j)\bar{j}\bar{j}}(k_1 - \tau_2, \tau_1 - \tau_2)}_{\mathbb{A}_{4\times} \in \mathbb{A}_{\times}} \\ &+ \underbrace{\sum_{k_2=\tau_1+1}^{+\infty} k_{\times}^j(\tau_1, k_2) \mathcal{D}_{p, (jj)\bar{j}\bar{j}}(k_2 - \tau_1, \tau_1 - \tau_2)}_{\mathbb{A}_{4\times} \in \mathbb{A}_{\times}} \end{aligned}$$

Table III recaps the different categories. We now proceed to the description of the four structures.

3. A_1 Structure

The matrices we call “ A_1 -like” are the one making the calibration between the diagonal of the observable (\mathbb{C} and \mathcal{D}_\times) and the diagonal of a kernel (ϕ or ϕ_\times). The characteristic equation is

$$E_{A_1}(\tau) = \sum_{u=1}^{+\infty} K(s)\mathcal{D}(\tau - s), \quad \text{where } K \text{ and } \mathcal{D} \text{ are a random kernel and covariance structure.}$$

Example – A_1 relation for $q = 3$

Up to the lag $q = 3$, we then have the following system

$$\begin{cases} E_{A_1}(1) = K(1)\mathcal{D}(0) + K(2)\mathcal{D}(-1) + K(3)\mathcal{D}(-2) \\ E_{A_1}(2) = K(1)\mathcal{D}(1) + K(2)\mathcal{D}(0) + K(3)\mathcal{D}(-1) \\ E_{A_1}(3) = K(1)\mathcal{D}(2) + K(2)\mathcal{D}(1) + K(3)\mathcal{D}(0) \end{cases}$$

so the Yule-Walker matrix is

$$A_1 = \begin{pmatrix} \mathcal{D}(0) & \mathcal{D}(-1) & \mathcal{D}(-2) \\ \mathcal{D}(1) & \mathcal{D}(0) & \mathcal{D}(-1) \\ \mathcal{D}(2) & \mathcal{D}(1) & \mathcal{D}(0) \end{pmatrix} \text{ and when } \mathcal{D} \text{ is even, } A_1 = \begin{pmatrix} \mathcal{D}(0) & \mathcal{D}(1) & \mathcal{D}(2) \\ \mathcal{D}(1) & \mathcal{D}(0) & \mathcal{D}(1) \\ \mathcal{D}(2) & \mathcal{D}(1) & \mathcal{D}(0) \end{pmatrix}$$

Generalizing to any q and for a any covariance \mathcal{D} , we can define

$$f_{A_1}(\mathcal{D}_{up}, \mathcal{D}_{down}, \Sigma) = \begin{pmatrix} \Sigma & \mathcal{D}_{up}(1) & \mathcal{D}_{up}(2) & \mathcal{D}_{up}(3) & \dots & \mathcal{D}_{up}((q-1)) \\ \mathcal{D}_{down}(1) & \Sigma & \mathcal{D}_{up}(1) & \mathcal{D}_{up}(2) & \dots & \mathcal{D}_{up}((q-2)) \\ \mathcal{D}_{down}(2) & \mathcal{D}_{down}(1) & \Sigma & \mathcal{D}_{up}(1) & \dots & \mathcal{D}_{up}((q-3)) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathcal{D}_{down}(q-1) & \mathcal{D}_{down}(q-2) & \dots & \mathcal{D}_{down}(1) & \dots & \Sigma \end{pmatrix}$$

We now just need to determine $(\mathcal{D}_{up}, \mathcal{D}_{down}, \Sigma)$ for each of the “ A_1 -like” matrices.

a. \mathbb{A}_{1d}

\mathbb{A}_{1d} is the same as for the linear GARCH model and we have

$$\mathbb{A}_{1d} = f_{A_1}(\mathbb{C}^\top, \mathbb{C}, \mathbb{C}(0)).$$

b. $\mathbb{A}_{1d\times}$

$\mathbb{A}_{1d\times}$ makes the relation between $\mathcal{D}_\times(\tau, \tau)$ and ϕ . From the Yule-Walker equation in lemma 3, we have

$$\mathbb{A}_{1d\times} = f_{A_1}(\mathcal{D}_\times, \tau \rightarrow \mathcal{D}_\times(-\tau), \mathcal{D}_\times(0)).$$

c. $A_{1\times}$

Similarly, $A_{1\times}$ encodes the calibration between $\mathcal{D}_\times(\tau, \tau)$ and ϕ_\times , from lemma 3, we have

$$A_{5\times} = f_{A_1}((\mathcal{D}_{(i\bar{i})i\bar{i}} - \bar{r}_i \bar{r}_{\bar{i}}^2)), \tau \rightarrow (\mathcal{D}_{(i\bar{i})i\bar{i}}(-\tau) - \bar{r}_i \bar{r}_{\bar{i}}^2), (\mathcal{D}_{(i\bar{i})i\bar{i}}(0) - \bar{r}_i \bar{r}_{\bar{i}}^2))$$

4. A_2 Structure

“ A_2 -like” matrices encode the calibration between the diagonal of the observable (\mathbb{C} and \mathcal{D}_\times) and the non diagonal of the kernels \mathbb{K}_d and \mathbf{K}_\times . The characteristic equation is

$$E_{A_2}(\tau) = 2 \sum_{k_2=\tau+1}^{+\infty} \sum_{k_1=k_2+1}^{+\infty} K(k_1, k_2) \mathcal{D}(k_1 - \tau, k_2 - \tau),$$

where K and \mathcal{D} are a random kernel and covariance structure.

Example – A_2 relation for $q = 3$

Up to the lag $q = 3$, we then have the following system

$$\begin{cases} E_{A_2}(1) &= K(1, 2)\mathcal{D}(0, 1) + K(1, 3)\mathcal{D}(0, 2) + K(2, 3)\mathcal{D}(1, 2) \\ E_{A_2}(2) &= K(1, 2)\mathcal{D}(-1, 0) + K(1, 3)\mathcal{D}(-1, 1) + K(2, 3)\mathcal{D}(0, 1) \\ E_{A_2}(3) &= K(1, 2)\mathcal{D}(-2, -1) + K(1, 3)\mathcal{D}(-2, 0) + K(2, 3)\mathcal{D}(-1, 0) \end{cases}.$$

Hence, we have the following matrix

$$\begin{pmatrix} E(1) \\ E(2) \\ E(3) \end{pmatrix} = \begin{pmatrix} \mathcal{D}(0, 1) & \mathcal{D}(0, 2) & \mathcal{D}(1, 2) \\ \mathcal{D}(-1, 0) & \mathcal{D}(-1, 1) & \mathcal{D}(0, 1) \\ \mathcal{D}(-2, -1) & \mathcal{D}(-2, 0) & \mathcal{D}(-1, 0) \end{pmatrix} \begin{pmatrix} K(1, 2) \\ K(1, 3) \\ K(2, 3) \end{pmatrix}$$

and if \mathcal{D} is causal,

$$\begin{pmatrix} E(1) \\ E(2) \\ E(3) \end{pmatrix} = \begin{pmatrix} \mathcal{D}(0, 1) & \mathcal{D}(0, 2) & \mathcal{D}(1, 2) \\ 0 & 0 & \mathcal{D}(0, 1) \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} K(1, 2) \\ K(1, 3) \\ K(2, 3) \end{pmatrix}$$

Generalizing for any lag q and for general structure \mathcal{D} , we can define $f_{A_2}(\mathcal{D})$ as following: if we consider the matrix $(\mathcal{D}(i, j))_{i, j \in \llbracket -(q-1), (q-1) \rrbracket}$ then, the n^{th} row of A_2 -like matrices is two times the upper triangle of $(\mathcal{D}(i, j))_{i, j \in \llbracket -(q-1), (q-1) \rrbracket} [q-n : 2q-n, q-n : 2q-n]$. Calling f_{A_2} the function performing the transformation $A_2 = f_{A_2}(\mathcal{D})$, we now proceed to the definition of the A_2 -like matrices.

a. \mathbb{A}_{2d}

For \mathbb{A}_{2d} , we use the link between \mathbb{C} and the quadratic kernel \mathbb{K} so

$$\mathbb{A}_{2d} = f_{A_2}(\mathbb{D}_d^\top).$$

b. $\mathbb{A}_{2d\times}$

For $\mathbb{A}_{2d\times}$, we use the link between the time diagonal of \mathcal{D}_\times and the off time-diagonal kernel $\mathbb{K}_d^{\text{off-diagonal}}$ from lemma 3, and we obtain

$$\mathbb{A}_{2d\times} = f_{A_2}\left(\begin{pmatrix} \mathcal{D}_{i\bar{i}ii} \\ \mathcal{D}_{i\bar{i}i\bar{i}} \end{pmatrix}\right).$$

c. $\mathbb{A}_{2\times}$

$\mathbb{A}_{2\times}$ encodes the calibration between the time diagonal of \mathcal{D}_\times and the kernel \mathbf{k}_\times . Since \mathbf{k}_\times is not time-symmetric (conversely to \mathbb{K}) we need to considered \mathbf{k}_\times for both $\tau_1 < \tau_2$ and $\tau_2 < \tau_1$. This phenomenon intervenes in several calibration matrices, as $\mathbb{A}_{2\times}$. We set the left bloc being the one encoding the calibration of $\mathbf{k}_\times(\tau_1, \tau_2)_{\tau_1 < \tau_2}$ and the right bloc for $\mathbf{k}_\times(\tau_1, \tau_2)_{\tau_1 > \tau_2}$ (consistently with the framework developped above).

Then, from lemma 3, and we obtain

$$\mathbb{A}_{2\times}^{\text{left bloc}} = f_{A_2}(\mathcal{D}_{p(i\bar{i})i\bar{i}}) \quad \& \quad \mathbb{A}_{2\times}^{\text{right bloc}} = f_{A_2}(\mathcal{D}_{p(i\bar{i})\bar{i}i}).$$

5. A_3 structure

“ A_3 -like” matrices are the one linking the off-time-diagonal of observable covariances (\mathbb{D}_d and \mathcal{D}_x with $\tau_1 \neq \tau_2$) with the time-diagonal kernels (\emptyset and ϕ_x). The characteristic equation is

$$E_{A_3}(\tau_1, \tau_2) = \sum_{u=1}^{+\infty} K(u) \mathcal{D}(\tau_1 - u, \tau_2 - u)$$

where K and \mathcal{D} are a random kernel and covariance structure.

Example – A_3 relation for $q = 3$

Up to the lag $q = 3$, we then have the following system

$$\begin{cases} \mathcal{D}_{A_3}(1, 2) = K(3)\mathcal{D}(-2, -1) + K(2)\mathcal{D}(-1, 0) + K(1)\mathcal{D}(0, 1) \\ \mathcal{D}_{A_3}(1, 3) = K(3)\mathcal{D}(-2, 0) + K(2)\mathcal{D}(-1, 1) + K(1)\mathcal{D}(0, 2) \\ \mathcal{D}_{A_3}(2, 3) = K(3)\mathcal{D}(-1, 0) + K(2)\mathcal{D}(0, 1) + K(1)\mathcal{D}(1, 2) \end{cases} .$$

Hence, we obtain the system:

$$\begin{pmatrix} \mathcal{D}_{A_3}(1, 2) \\ \mathcal{D}_{A_3}(1, 3) \\ \mathcal{D}_{A_3}(2, 3) \end{pmatrix} = \begin{pmatrix} \mathcal{D}(0, 1) & \mathcal{D}(-1, 0) & \mathcal{D}(-2, -1) \\ \mathcal{D}(0, 2) & \mathcal{D}(-1, 1) & \mathcal{D}(-2, 0) \\ \mathcal{D}(1, 2) & \mathcal{D}(0, 1) & \mathcal{D}(-1, 0) \end{pmatrix} \begin{pmatrix} K(1) \\ K(2) \\ K(3) \end{pmatrix}$$

and if \mathcal{D} is causal, we have

$$\begin{pmatrix} \mathcal{D}_{A_3}(1, 2) \\ \mathcal{D}_{A_3}(1, 3) \\ \mathcal{D}_{A_3}(2, 3) \end{pmatrix} = \begin{pmatrix} \mathcal{D}(0, 1) & 0 & 0 \\ \mathcal{D}(0, 2) & 0 & 0 \\ \mathcal{D}(1, 2) & \mathcal{D}(0, 1) & 0 \end{pmatrix} \begin{pmatrix} K(1, 1) \\ K(2, 2) \\ K(3, 3) \end{pmatrix} .$$

The function to create “ A_3 -like” matrices f_{A_3} , has the same structure as f_{A_2} , only transposed. Hence, we define $f_{A_3}(\mathcal{D}) = f_{A_2}(\mathcal{D})^\top / 2$. We now detailed the definition of each “ A_3 -like” matrix.

a. \mathbb{A}_{3d}

\mathbb{A}_{3d} encodes the feedback of \emptyset on the time-off-diagonal of \mathbb{D}_d . Using lemma 2, we have

$$\mathbb{A}_{3d} = f_{A_3}(\mathbb{D}_d)$$

In python, one can use `(A3=A2.transpose(1,0,3,2)/2)`.

b. \mathbb{A}_{3dx}

\mathbb{A}_{3dx} encodes the feedback of \emptyset on the off-time-diagonal of \mathcal{D}_x . \mathbb{A}_{3dx} is then composed of two blocs to capture the time asymmetry of \mathcal{D}_x . The upper bloc characterises the feedback of ϕ_x on $\mathcal{D}_x(\tau_1, \tau_2)$ for $\tau_1 < \tau_2$ and the lower bloc for $\tau_1 > \tau_2$.

Using lemmas 4 and 5, we can write

$$\mathbb{A}_{3dx}^{\text{upper}} = f_{A_3}\left(\begin{pmatrix} \mathcal{D}_{p,(ii)\bar{i}\bar{i}} \\ \mathcal{D}_{p,(\bar{i}\bar{i})i\bar{i}} \end{pmatrix}\right) \quad \& \quad \mathbb{A}_{3dx}^{\text{lower}} = f_{A_3}\left(\begin{pmatrix} \mathcal{D}_{p,(ii)\bar{i}i} \\ \mathcal{D}_{p,(\bar{i}\bar{i})\bar{i}i} \end{pmatrix}\right).$$

The second level is done with the symmetric of the first level.

c. \mathbb{A}_{3x}

\mathbb{A}_{3x} encodes the feedback of ϕ_x on the off-time-diagonal of \mathcal{D}_x . As for \mathbb{A}_{3dx} , we account for the time asymmetry of \mathcal{D}_x by allocating the upper bloc of \mathbb{A}_{3x} to $\mathcal{D}_x(\tau_1, \tau_2)$ for $\tau_1 < \tau_2$ and the lower bloc for $\tau_1 > \tau_2$. Using lemmas 4 and 5, we have

$$\mathbb{A}_{3x}^{\text{upper}} = f_{A_3}((\mathcal{D}_{(i\bar{i})\bar{i}\bar{i}})) \quad \& \quad \mathbb{A}_{3x}^{\text{lower}} = f_{A_3}((\mathcal{D}_{(\bar{i}\bar{i})\bar{i}i})).$$

6. A_4 structure

“ A_4 -like” matrices are the one linking off-time-diagonal of observable (\mathbb{D}_d and \mathcal{D}_\times with $\tau_1 \neq \tau_2$) and the off-time-diagonal of kernels (\mathbb{K} and \mathbf{k}_\times). The characteristic equation is

$$E_{A_4}(\tau_1, \tau_2) = 2 \sum_{u=\tau_1+1}^{+\infty} K(u, \tau_1) \mathcal{D}(u - \tau_1, \tau_2 - \tau_1)$$

where K and \mathcal{D} are a random kernel and covariance structure.

Example – A_4 relation for $q = 3$

Up to the lag $q = 3$, we then have the following system for $q = 3$,

$$\begin{cases} \mathcal{D}_{A_4}(1, 2) = 2(K(2, 1)\mathcal{D}(1, 1) + K(3, 1)\mathcal{D}(2, 1)) \\ \mathcal{D}_{A_4}(1, 3) = 2(K(2, 1)\mathcal{D}(1, 2) + K(3, 1)\mathcal{D}(2, 2)) \\ \mathcal{D}_{A_4}(2, 3) = 2K(3, 2)\mathcal{D}(1, 1) \end{cases}.$$

Hence, we obtain the system

$$\begin{pmatrix} \mathcal{D}_{A_4}(1, 2) \\ \mathcal{D}_{A_4}(1, 3) \\ \mathcal{D}_{A_4}(2, 3) \end{pmatrix} = \begin{pmatrix} \mathcal{D}(1, 1) & \mathcal{D}(2, 1) & 0 \\ \mathcal{D}(1, 2) & \mathcal{D}(2, 2) & 0 \\ 0 & 0 & \mathcal{D}(1, 1) \end{pmatrix} \begin{pmatrix} K(1, 2) \\ K(1, 3) \\ K(2, 3) \end{pmatrix}.$$

The function to create “ A_4 -like” matrices from the correlation \mathcal{D} can be deduced from the example. It creates blocs of the matrix $(\mathcal{D}(i, j))_{i,j}$ along the diagonal. We now detail the structure of the A_4 -like matrices.

a. \mathbb{A}_{4d}

\mathbb{A}_{4d} encodes the link between \mathbb{K} and \mathbb{D}_d . From lemma 2, we have

$$\mathbb{A}_{4d} = f_{A_4}((\mathcal{D}_{(ij)ij})_{ij}).$$

b. $\mathbb{A}_{4d\times}$

$\mathbb{A}_{4d\times}$ encodes the feedback of \mathbb{K}_d on the off-time-diagonal of \mathcal{D}_\times . To account for the asymmetry of \mathcal{D}_\times , we build $\mathbb{A}_{4d\times}$ with two blocs: the upper bloc being for $\mathcal{D}_\times(\tau_1, \tau_2)$ with $\tau_1 < \tau_2$ and the lower bloc for $\tau_2 < \tau_1$. From lemmas 4 and 5, we have

$$\mathbb{A}_{4d\times}^{\text{upper}} = f_{A_4}(\begin{pmatrix} \mathcal{D}_{p(ii)i\bar{i}} \\ \mathcal{D}_{p(i\bar{i})\bar{i}i} \end{pmatrix}) \quad \& \quad \mathbb{A}_{4d\times}^{\text{lower}} = f_{A_4}(\begin{pmatrix} \mathcal{D}_{p(\bar{i}i)i\bar{i}} \\ \mathcal{D}_{p(\bar{i}i)\bar{i}i} \end{pmatrix})$$

c. $\mathbb{A}_{4\times}$

$\mathbb{A}_{4\times}$ encodes the feedback of \mathbf{k}_\times on the off-time-diagonal of \mathcal{D}_\times . For this particular matrix, one has to account for the asymmetry of \mathcal{D}_\times and of \mathbf{k}_\times . As before, we thus build blocs for each case. From lemmas 4 and 5, we have

- upper left, $\mathcal{D}_\times(\tau_1 < \tau_2)$ VS $\mathbf{k}_\times(\tau_1 < \tau_2)$

$$\mathbb{A}_{4\times}^{\text{upper left}} = f_{A_4}((\mathcal{D}_{p(ii)\bar{i}\bar{i}}))$$

- upper right, $\mathcal{D}_\times(\tau_1 < \tau_2)$ VS $\mathbf{k}_\times(\tau_1 > \tau_2)$

$$\mathbb{A}_{4\times}^{\text{upper right}} = f_{A_4}((\mathcal{D}_{p(i\bar{i})\bar{i}i}))$$

- lower left, $\mathcal{D}_x(\tau_1 > \tau_2)$ VS $\mathbf{k}_x(\tau_1 < \tau_2)$

$$\mathbb{A}_{4 \times}^{\text{lower left}} = f_{A_4}((\mathcal{D}_{p(i\bar{i})\bar{i}i}))$$

- lower right, $\mathcal{D}_x(\tau_1 > \tau_2)$ VS $\mathbf{k}_x(\tau_1 > \tau_2)$

$$\mathbb{A}_{4 \times}^{\text{lower left}} = f_{A_4}((\mathcal{D}_{p(\bar{i}\bar{i})ii}))$$

7. Calibration matrices for the leverage kernels

a. Leverage calibration - framework

We now explicate a method to determine the leverage kernels $(L_j^i)_{i,j \in \{1,2\}}$. For those last kernels, we aim at building a linear system such that:

$$\begin{pmatrix} \mathbb{V}(1) \\ \mathbb{V}(2) \\ \vdots \\ \mathbb{V}(q) \end{pmatrix} = \mathbb{A}_{\mathbb{L}} \begin{pmatrix} \mathbb{L}(1) \\ \mathbb{L}(2) \\ \vdots \\ \mathbb{L}(q) \end{pmatrix}$$

where

$$\mathbb{L}(\tau) = \begin{pmatrix} L_1^1(\tau) & L_2^1(\tau) \\ L_1^2(\tau) & L_2^2(\tau) \end{pmatrix} \quad \text{and} \quad \mathbb{V}(\tau) = \begin{pmatrix} \mathcal{V}_{11}(\tau) & \mathcal{V}_{12}(\tau) \\ \mathcal{V}_{21}(\tau) & \mathcal{V}_{22}(\tau) \end{pmatrix}.$$

As before, we rely on the Yule-Walker type equation to construct $\mathbb{A}_{\mathbb{L}}$. As a reminder, lemma 6 gives

$$\begin{aligned} \mathcal{V}_{ji}(\tau) = & L_i^j(\tau) \bar{r}_i r_l \\ & + \sum_{i=1}^N \sum_{k=1}^{\tau} \phi_i^j(k) \mathcal{V}_{il}^r(\tau - k) \\ & + 2 \sum_{k_2=\tau+1}^{+\infty} K_i^j(\tau, k_2) \mathcal{V}_{(il)i}^r(k_2 - \tau) \\ & + \sum_{k=1}^{\tau} \phi_{\times}^j(k) \mathcal{V}_{(i\bar{i})l}^r(\tau - k) \\ & + \sum_{k_2=\tau+1}^{+\infty} k_{\times}^j(\tau, k_2) \mathcal{V}_{(jl)\bar{j}}^r(k_2 - \tau) \\ & + \sum_{k_1=\tau+1}^{+\infty} k_{\times}^j(k_1, \tau) \mathcal{V}_{(\bar{j}l)j}^r(k_1 - \tau). \end{aligned}$$

The equation above demonstrates that \mathcal{V}_{jl} not only depends on the leverage kernels $(L_j^i)_{i,j \in \{1,2\}}$ but also on the linear kernels $(\phi_j^i)_{i,j \in \{1,2\}}$, the quadratic kernels $(K_j^i)_{i,j \in \{1,2\}}$ and the cross kernels $(\phi_{\times}^i)_{i \in \{1,2\}}$ and $(k_{\times}^i)_{i \in \{1,2\}}$ that we determined in steps 1, 2 and 3. Hence, the equation system becomes:

$$\begin{pmatrix} \tilde{\mathbb{V}}(1) \\ \tilde{\mathbb{V}}(2) \\ \vdots \\ \tilde{\mathbb{V}}(q) \end{pmatrix} = \begin{pmatrix} \left(\frac{\bar{r}_1^2}{\bar{r}_1 r_2} \frac{\bar{r}_1 r_2}{\bar{r}_2^2} \right) & 0 \\ \ddots & \ddots \\ 0 & \left(\frac{\bar{r}_1^2}{\bar{r}_1 r_2} \frac{\bar{r}_1 r_2}{\bar{r}_2^2} \right) \end{pmatrix} \begin{pmatrix} \mathbb{L}(1) \\ \mathbb{L}(2) \\ \vdots \\ \mathbb{L}(q) \end{pmatrix}$$

where

$$\begin{pmatrix} \tilde{\mathbb{V}}(1) \\ \tilde{\mathbb{V}}(2) \\ \vdots \\ \tilde{\mathbb{V}}(q) \end{pmatrix} = \begin{pmatrix} \mathbb{V}(1) \\ \mathbb{V}(2) \\ \vdots \\ \mathbb{V}(q) \end{pmatrix} - \mathbb{A}_{1L} \begin{pmatrix} \phi(1) \\ \phi(2) \\ \vdots \\ \phi(q) \end{pmatrix} - \mathbb{A}_{1\times L} \begin{pmatrix} \phi_{\times}(1) \\ \phi_{\times}(2) \\ \vdots \\ \phi_{\times}(q) \end{pmatrix} - \mathbb{A}_{vL} \begin{pmatrix} \mathbb{K}_{\times}(1, 2) \\ \mathbb{K}_{\times}(1, 3) \\ \vdots \\ \mathbb{K}_{\times}(q-1, q) \end{pmatrix} - \mathbb{A}_{v\times L} \begin{pmatrix} \mathbf{K}_{\times}(1, 2) \\ \mathbf{K}_{\times}(1, 3) \\ \vdots \\ \mathbf{K}_{\times}(q-1, q) \end{pmatrix},$$

and \mathbb{A}_{1L} , $\mathbb{A}_{1\times L}$, \mathbb{A}_{vL} and $\mathbb{A}_{v\times L}$ are to be determined with the Yule-Walker equation of lemma 6. Notably, we have

$$\begin{aligned} \mathcal{V}_{jl}(\tau) &= L_i^j(\tau) \bar{r}_i \bar{r}_l \\ &\quad + \underbrace{\sum_{i=1}^N \sum_{k=1}^{\tau} \phi_i^j(k) \mathcal{V}_{il}^r(\tau-k)}_{\mathbb{A}_{1L}} \\ &\quad + \underbrace{2 \sum_{k_2=\tau+1}^{+\infty} K_i^j(\tau, k_2) \mathcal{V}_{(il)i}^r(k_2-\tau)}_{\mathbb{A}_{vL}} \\ &\quad + \underbrace{\sum_{k=1}^{\tau} \phi_{\times}^j(k) \mathcal{V}_{(ii)\bar{j}}^r(\tau-k)}_{\mathbb{A}_{1\times L}} \\ &\quad + \underbrace{\sum_{k_2=\tau+1}^{+\infty} k_{\times}^j(\tau, k_2) \mathcal{V}_{(jl)\bar{j}}^r(k_2-\tau)}_{\mathbb{A}_{v\times L} \text{ left bloc}} \\ &\quad + \underbrace{\sum_{k_1=\tau+1}^{+\infty} k_{\times}^j(k_1, \tau) \mathcal{V}_{(\bar{j}l)j}^r(k_1-\tau)}_{\mathbb{A}_{v\times L} \text{ right bloc}}. \end{aligned}$$

Using the work above, we can already determine \mathbb{A}_{1L} and $\mathbb{A}_{1\times L}$ as they are “ A_1 -like” matrices. Hence, we have

$$\mathbb{A}_{1L} = f_{A_1}(\mathbf{0}_{2\times 2}, (\mathbb{V}^r), \begin{pmatrix} \overline{\sigma_1^2 r_1} & 0 \\ 0 & \overline{\sigma_2^2 r_2} \end{pmatrix}) \quad \text{and} \quad \mathbb{A}_{1\times L} = f_{A_1}(\mathbf{0}_{2\times 2}, \begin{pmatrix} \mathcal{V}_{(ii)i}^r \\ \mathcal{V}_{(ii)\bar{i}}^r \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}).$$

b. A_5 structure

Additionally, we need to define another type of matrices for \mathbb{A}_{vL} and $\mathbb{A}_{v\times L}$. Both have the same structure which we call “ A_5 -like”. The characteristic equation is as follows

$$E_{A_5}(\tau) = \sum_{k=\tau+1}^{+\infty} K(\tau, k) \mathcal{D}(k-\tau).$$

where K and \mathcal{D} are a random kernel and covariance structure.

Example – A_5 relation for $q = 3$

Up to the lag $q = 3$, we then have the following system

$$\begin{cases} E_{A_5}(1) = K(1, 2)\mathcal{D}(1) + K(1, 3)\mathcal{D}(2) \\ E_{A_5}(2) = K(2, 3)\mathcal{D}(1) \\ E_{A_5}(3) = 0 \end{cases}$$

so the Yule-Walker matrix is

$$A_5 = \begin{pmatrix} \mathcal{D}(1) & \mathcal{D}(2) & 0 \\ 0 & 0 & \mathcal{D}(1) \\ 0 & 0 & 0 \end{pmatrix}$$

Hence, the generalisation of “ A_5 -like” matrices comes down to writing each row i of A_5 as $A_5[i, q - i : 2q - 2i - 1] + = \mathcal{D}[:, q - i - 1]$ and taking the other entries as zeros. Then, calling f_{A_5} the function to build “ A_5 -like” matrices such like $A_5 = f_{A_5}(\mathcal{D})$ we can define the last Yule-Walker matrices as

$$\mathbb{A}_{vL} = f_{A_5}\left(\begin{pmatrix} \mathcal{V}_{(ii)i}^r & \mathcal{V}_{(i\bar{i})i}^r \\ \mathcal{V}_{(i\bar{i})\bar{i}}^r & \mathcal{V}_{(\bar{i}\bar{i})\bar{i}}^r \end{pmatrix}\right), \quad \mathbb{A}_{v \times L}^{\text{left bloc}} = f_{A_5}\left(\begin{pmatrix} \mathcal{V}_{(ii)\bar{i}}^r \\ \mathcal{V}_{(i\bar{i})\bar{i}}^r \end{pmatrix}\right) \quad \text{and} \quad \mathbb{A}_{v \times L}^{\text{right bloc}} = f_{A_5}\left(\begin{pmatrix} \mathcal{V}_{(\bar{i}\bar{i})\bar{i}}^r(-) \\ \mathcal{V}_{(\bar{i}\bar{i})\bar{i}}^r(-) \end{pmatrix}\right).$$

Empirically, we observe that the contribution of these kernels is not significant, and so, in practice for the sake of simplicity, we only consider \mathbb{A}_{1L} and $\mathbb{A}_{1 \times L}$ in the calibration.

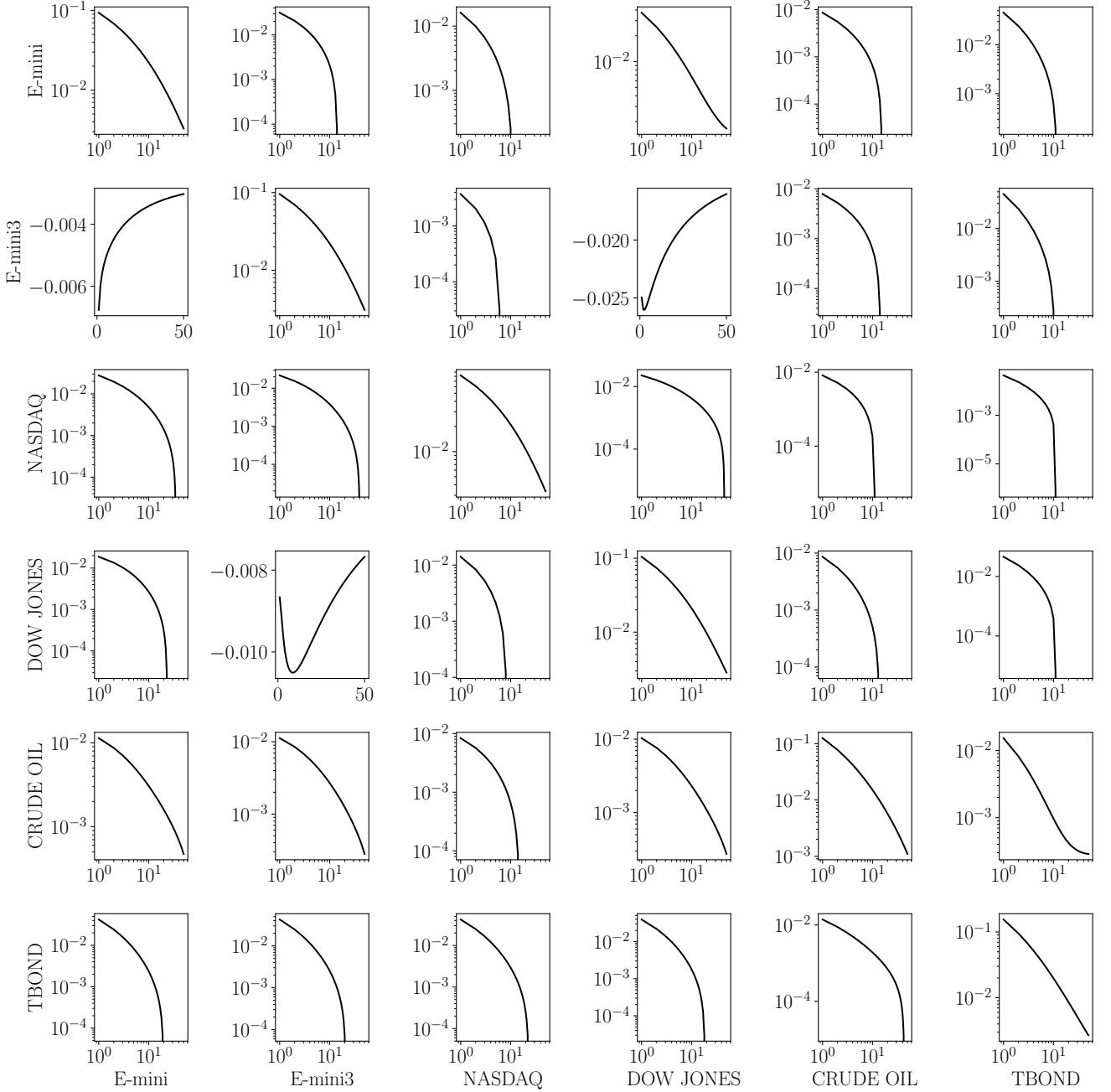
Appendix E: Kernels profiles from 2D-QGARCH Calibration on futures on indices


Figure 18: (ϕ_i^j) profiles – the x-labels determine the index providing feedback and the y-labels determine the index receiving the feedback, i.e., for a kernel ϕ_i^j , i is labelled on the x-axis while j is labelled on the y-axis.

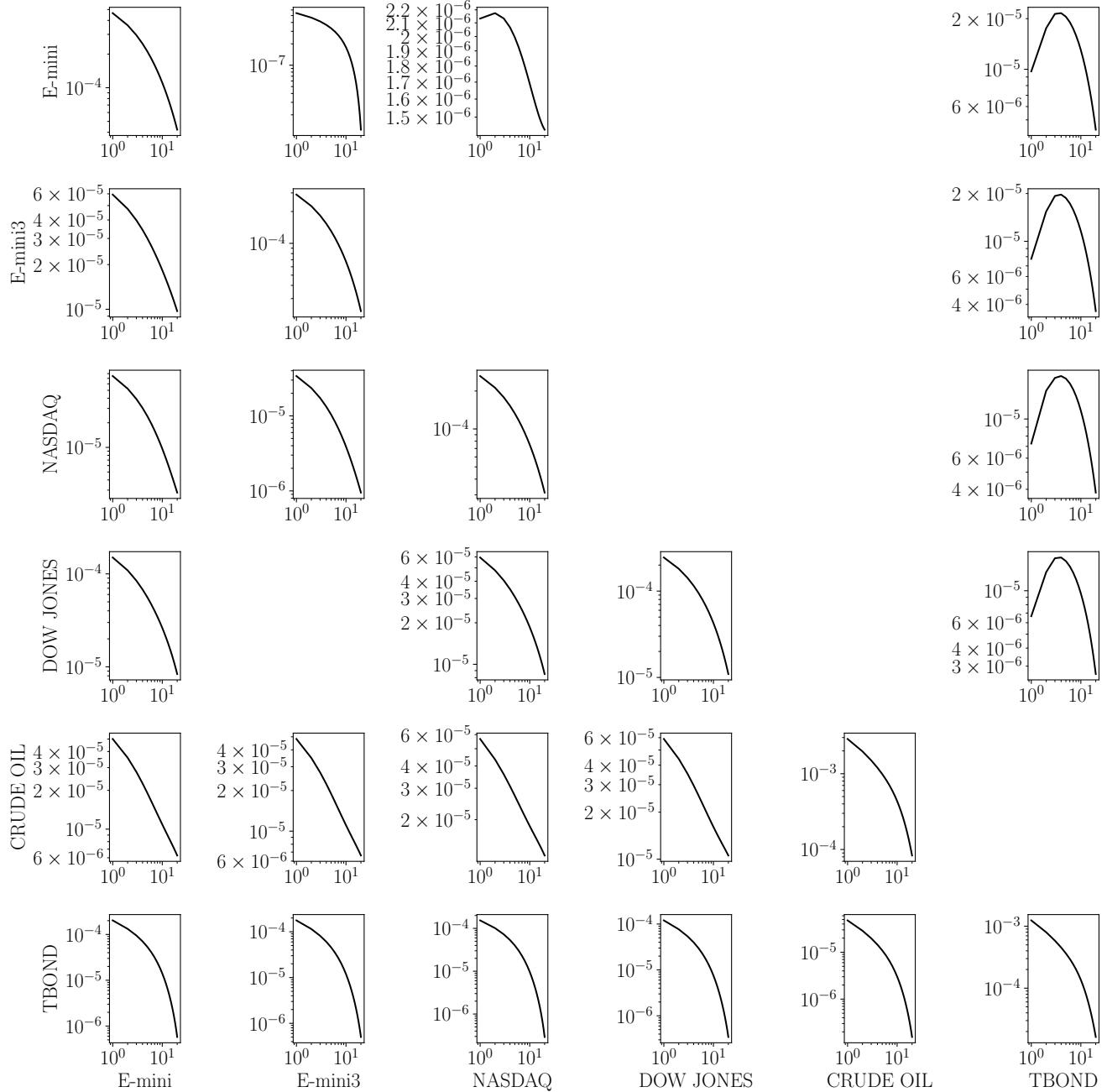


Figure 19: $(k_i^j)^2$ profiles – the x-labels determine the index providing feedback and the y-labels determine the index receiving the feedback, i.e., for a kernel $(k_i^j)^2$, i is labelled on the x-axis while j is labelled on the y-axis.

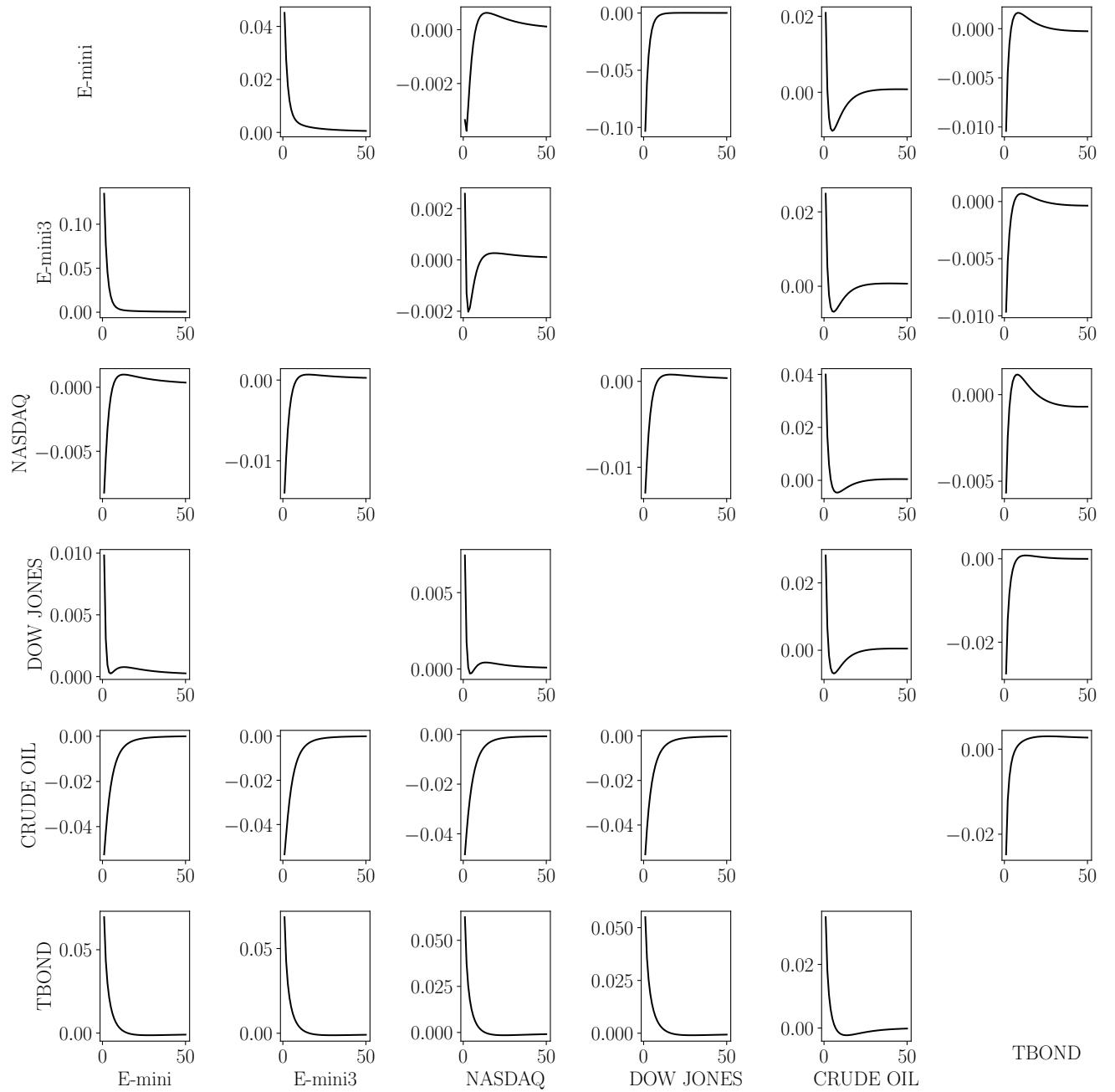


Figure 20: ϕ_x^j profiles – the y-labels determine the index receiving the feedback while the x-labels determine the index used for the correlation, i.e., the subplot (i, j) show the profile of the kernel (ϕ_x^i) for the correlation between asset i and asset j .

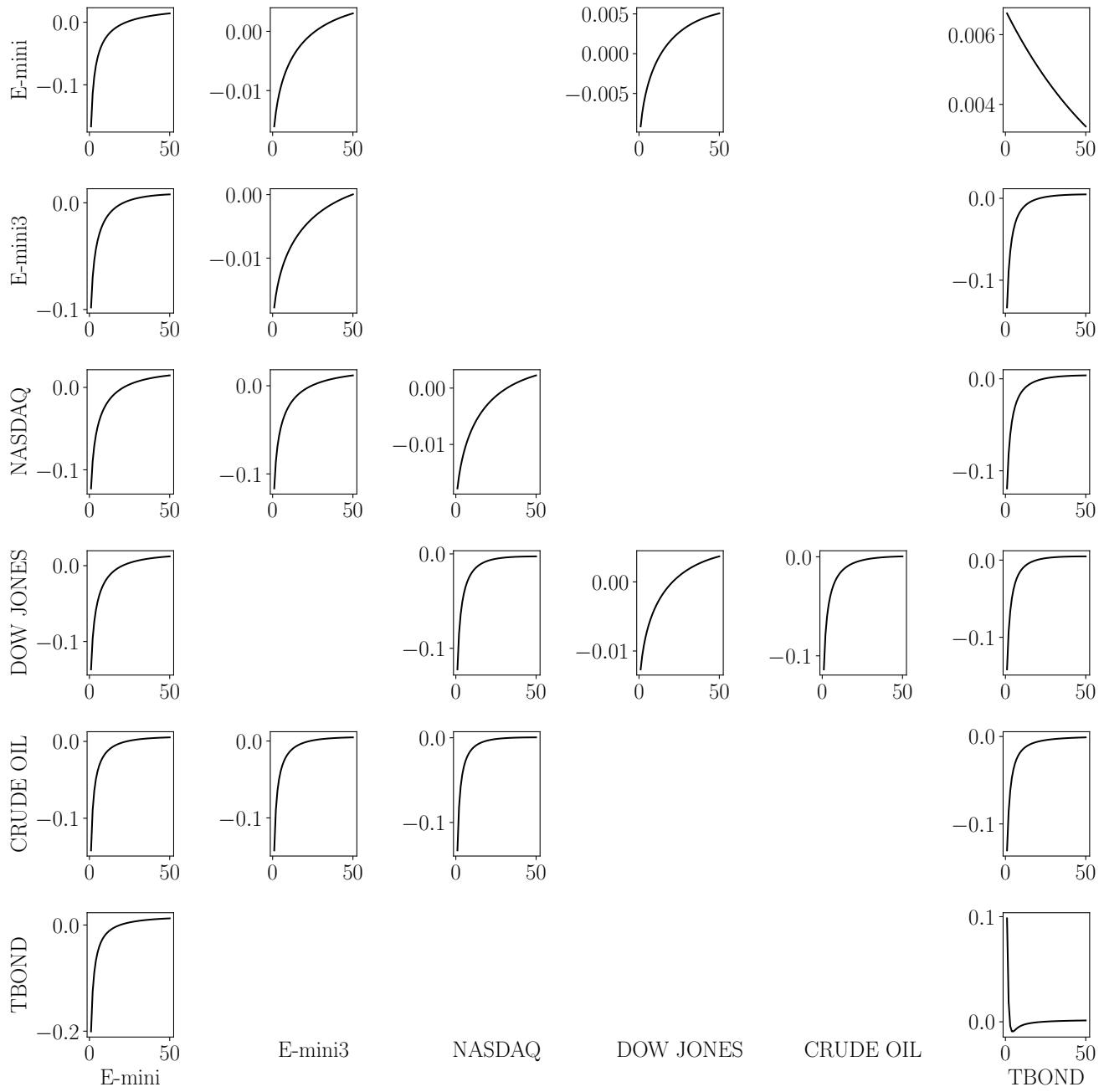


Figure 21: L_i^j profiles – the x-labels determine the index providing feedback and the y-labels determine the index receiving the feedback, i.e., for a kernel L_i^j , i is labelled on the x-axis while j is labelled on the y-axis.

Appendix F: Fits and noise reduction parameters

	α	β	γ	n
\mathcal{C}_E	0.740	1e-8	10.00	3.0845
$\mathcal{C}_{E,T}$	0.1184	1e-8	10.00	1.8810
$\mathcal{C}_{T,E}$	0.1151	1e-8	10.00	2.0408
\mathcal{C}_T	0.2088	1e-8	10.00	3.5990

Table IV: Fits parameters of the 2-points correlation \mathcal{C}_{ij} , as defined by Equation (18) with the fit of $\tau \rightarrow n \exp(-\beta\tau)(1 + \gamma\tau)^{-\alpha}$, for the pair E-MINI VS TBOND

	α	β	γ	n
$\mathcal{D}_{\times E}$	0.0911	1e-8	10.00	-0.6743
$\mathcal{D}_{\times T}$	0.1062	1e-8	10.00	-0.3118

Table V: Fits parameters of the 3-points correlation $\mathcal{D}_{\times j}$, as defined by Equation (20) with the fit of $\tau \rightarrow n \exp(-\beta\tau)(1 + \gamma\tau)^{-\alpha}$, for the pair E-MINI VS TBOND

Appendix G: Parametric Calibration–Maximum Likelihood

In this appendix, we describe a widely used parametric calibration method for (Q)Hawkes processes, with intensity λ . The method relies on maximising the likelihood of jump times. We present here two applications:

- maximum likelihood when exact jump times are available. Note, that it is the method used in the package `tick` of Python to calibrate (multidimensional) linear Hawkes.
- maximum likelihood with proxies, when event-by-event data is not available.

1. With exact time of events

a. Log likelihood for point process

For this first method of likelihood maximisation, one needs to have access to the event times $(t_i)_i$. The likelihood is then defined as the joint distribution of those time of events. Considering a Hawkes process of intensity λ , we first give the probability of event occurring between time t_i and $t_i + dt$, without any event happening between time t_{i-1} and t_i .

The probability of event between time t_i and $t_i + dt$ but no event between t_{i-1} and t_i is given by

$$\mathbb{P}(\text{one event in } [t_i, t_i + dt] \text{ but no event in } [t_{i-1}, t_i]) = \exp\left\{\left(-\int_{t_{i-1}}^{t_i} du \lambda(u|\theta)\right)\right\} \lambda(t_i|\theta) dt$$

One can then write the joint distribution of N events happening in $[0, T]$ at arrival times $(t_i)_i$, and deduce the log-likelihood.

For N events in $[0, T]$ at arrival times $(t_i)_i$, then, the log-likelihood is:

$$\ln(\mathbf{L}(t_1, t_2, \dots, t_N, \theta)) = \sum_{i=1}^N \ln(\lambda(t_i|\theta)) - \int_0^T du \lambda(u|\theta) \quad (\text{G1})$$

where λ is the intensity and θ the parameters we are looking to estimate.

Proof. From definition of the probability of event, we write the joint distribution of N independent events happening in $[0, T]$ at arrival times $(t_i)_i$

$$\begin{aligned} \mathbf{L}(t_1, t_2, \dots, t_N, \theta) &= \left(\prod_{i=1}^N \exp\left\{\left(-\int_{t_{i-1}}^{t_i} du \lambda(u|\theta)\right)\right\} \lambda(t_i|\theta) \right) \exp\left\{\left(-\int_{t_N}^T du \lambda(u|\theta)\right)\right\} \\ &= \left(\prod_{i=1}^N \lambda(t_i|\theta) \right) \exp\left\{\left(-\int_0^T du \lambda(u|\theta)\right)\right\}. \end{aligned}$$

Taking the log gives the result. □

b. Log likelihood for linear Hawkes process

For linear Hawkes processes, the intensity is defined as

$$\lambda(t|\theta) = \lambda_\infty + \int_{-\infty}^t h_\theta(t-s) dN_s = \lambda_\infty + \sum_{t_i < t} h(t-t_i)$$

where $(t_i)_i$ are the times of the jumps.

Particularly, with exponential kernels, we have $\theta = (\lambda_\infty, n_h, \beta)$, $h_\theta(t) = n_h \beta \exp(-\beta t)$. Then the expression of the log-likelihood evolves to:

$$\ln(\mathbf{L}(t_1, t_2, \dots, t_N, \theta)) = -T\lambda_\infty + \sum_{i=1}^N \left(\ln(\phi_0 + n_H \beta H_i) - n_H (\exp(-\beta(T-t_i)) - 1) \right)$$

with

$$H_i = \sum_{k=1}^{i-1} \exp(-\beta(t_i - t_k))$$

Indeed, we have $\lambda(t_i/\theta) = \lambda_\infty + n_H \beta H_i$. And

$$\begin{aligned} \int_0^T du \lambda(u/\theta) &= \lambda_\infty T + n_H \beta \int_0^T \int_{-\infty}^u \exp(-\beta(u-s)) dN_s du \\ &= \lambda_\infty T + n_H \beta \int_{-\infty}^T \int_s^T \exp(-\beta(u-s)) du dN_s \\ &= \lambda_\infty T + n_H \beta \int_{-\infty}^T \frac{1}{\beta} (1 - \exp(-\beta(T-s))) dN_s \\ &= \lambda_\infty T + n_H \sum_{t_s < T} (1 - \exp(-\beta(T-t_s))) \end{aligned}$$

c. Log likelihood for ZHawkes process

For ZHawkes processes, one needs to add the quadratic term to the intensity, which then becomes:

$$\begin{aligned} \lambda(t/\theta) &= \lambda_\infty + \int_{-\infty}^t h_\theta(t-s) dN_s + \left(\int_{-\infty}^t k_\theta(t-s) dP_s \right)^2 \\ &= \lambda_\infty + \sum_{t_i < t} h(t-t_i) + \left(\sum_{t_i < t} k(t-t_i) m_{t_i} \right)^2 \end{aligned}$$

where $(t_i)_i$ are the times of the jumps and $m(t_i)_t$ the marks of the jump at time t_i . Again, in the particular case of exponential kernels, we have $\theta = (\lambda_\infty, n_H, \beta, n_Z, \omega)$, $h_\theta(t) = n_H \beta \exp(-\beta t)$, $k_\theta(t) = \sqrt{2n_Z \omega} \exp(-\omega t)$. The log-likelihood can then entirely be defined by:

$$\ln(\mathbf{L}(t_1, t_2, \dots, t_N, \theta)) = -T \lambda_\infty + \sum_{i=1}^N \left(\ln(\lambda_\infty + n_H \beta H_i + 2n_Z \omega Z_i^2) - n_H (\exp(-\beta(T-t_i)) - 1) - 2n_Z \omega Z_i^2 \right)$$

with

$$H_i = \sum_{k=1}^{i-1} \exp(-\beta(t_i - t_k)) \quad \text{and} \quad Z_i = \sum_{k=1}^{i-1} \exp(-\omega(t_i - t_k)) m_{t_k}$$

d. Log likelihood for multivariate point process (2D)

We consider a 2D process, with intensities (λ_1, λ_2) and events arrival times $((t_i^1)_{i \in \llbracket 1, N \rrbracket}, (t_i^2)_{i \in \llbracket 1, N \rrbracket})$, where $(t_i^1)_{i \in \llbracket 1, N \rrbracket}$, respectfully $(t_i^2)_{i \in \llbracket 1, N \rrbracket}$, are the times of events of process 1, respectfully process 2. The log likelihood consider the joint probability of the 2 jump processes, and thus results in:

[Log likelihood in 2D]

Log likelihood of two point processes with intensities (λ_1, λ_2) , the log-likelihood of N events from process 1 at times $(t_i^1)_{i \in \llbracket 1, N \rrbracket}$, and N events of process 2 at times $(t_i^2)_{i \in \llbracket 1, N \rrbracket}$, between 0 and time T is given by:

$$\ln(\mathbf{L}((t_i^1)_{i \in \llbracket 1, N \rrbracket}, (t_i^2)_{i \in \llbracket 1, N \rrbracket}, \theta)) = \sum_{i=1}^N \ln(\lambda_1(t_i^1/\theta)) + \sum_{i=1}^N \ln(\lambda_2(t_i^2/\theta)) - \int_0^T du (\lambda_1(u/\theta) + \lambda_2(u/\theta))$$

where θ is the vector of the parameters we are looking for.

2. Without exact times of events - Using an intensity estimate

a. Defining the (log) likelihood

It is sometimes impossible to access event times $(t_i)_i$. However, we can estimate λ_t for each time bin t . We can consider λ_t constant over the bin t , then the probability to observe n events in the bin t is:

$$\mathbb{P}(\text{nb of events in bin } t = n) = \frac{\lambda_t^n}{n!} \exp(-\lambda_t)$$

The number of events we observe in bin t is our $\hat{\lambda}$ ($n = \hat{\lambda}$).

In the QHawkes framework, λ_t is defined by:

$$\lambda_t = \lambda_0 + \int_{-\infty}^t h(t-u)dN_u + \left(\int_{-\infty}^t k(t-u)dP_u \right)^2$$

Using the approximation $dN_u \equiv \mathbb{E}(dN_u) = \hat{\lambda}_u dt$, we obtain:

$$\lambda_t = \lambda_0 + \sum_{u=0}^t h(t-u)\hat{\lambda}_u du + \left(\sum_{u=0}^t k(t-u)dP_u \right)^2$$

where $u = 0$ represent the first bin of the day. Note that dP_t is the returns over the bin t , which is directly observable.

We can then define the log likelihood. We consider the joint density of observing the events $((\hat{\lambda}^i)_{i \in \llbracket 1, N \rrbracket})$, where N represent the number of bins we have, then the log likelihood is given by:

$$\begin{aligned} \ln(\mathbf{L}((\hat{\lambda}^i)_{i \in \llbracket 1, N \rrbracket}, \theta)) &= \sum_{i=1}^N \ln \left(\frac{\lambda_i(\theta)^{\hat{\lambda}_i}}{\hat{\lambda}_i!} \exp(-\lambda_i(\theta)) \right) = \sum_{i=1}^N \hat{\lambda}_i \ln(\lambda_i(\theta)) - \ln(\hat{\lambda}_i!) - \lambda_i(\theta) \\ &\sim \sum_{i=1}^N \hat{\lambda}_i \ln(\lambda_i(\theta)) - (\hat{\lambda}_i \ln(\hat{\lambda}_i) - \hat{\lambda}_i) - \lambda_i(\theta) \\ \ln(\mathbf{L}((\hat{\lambda}^i)_{i \in \llbracket 1, N \rrbracket}, \theta)) &= \sum_{i=1}^N \hat{\lambda}_i \left(\ln(\lambda_i(\theta)) - \ln(\hat{\lambda}_i) \right) + \hat{\lambda}_i - \lambda_i(\theta) \end{aligned}$$

b. Defining the (log) likelihood for multidimensionnal process

[likelihood of observing $(\hat{\lambda}^1, \hat{\lambda}^2)$]

We consider 2 processes, with intensities (λ^1, λ^2) . We observe those two processes over N bins, and the number of events from process j in bin i is $\hat{\lambda}_i^j$. Then, we can define the joint probability (likelihood), to observe $(\hat{\lambda}_i^1, \hat{\lambda}_i^2)_{i \in \llbracket 1, N \rrbracket}$ over the N bins:

$$\begin{aligned} \mathbf{L}((\hat{\lambda}_i^1, \hat{\lambda}_i^2)_i, \theta) &= \prod_{i=1}^N \mathbb{P}(n_1 \text{ events from process 1 in bin } i) \mathbb{P}(n_2 \text{ events from process 2 in bin } i) \\ &= \prod_{i=1}^N \frac{(\lambda_i^1(\theta))^{n_1}}{n_1!} \exp(-\lambda_i^1(\theta)) \frac{(\lambda_i^2(\theta))^{n_2}}{n_2!} \exp(-\lambda_i^2(\theta)) \\ &= \prod_{i=1}^N \frac{(\lambda_i^1(\theta))^{\hat{\lambda}_i^1}}{\hat{\lambda}_i^1!} \exp(-\lambda_i^1(\theta)) \frac{(\lambda_i^2(\theta))^{\hat{\lambda}_i^2}}{\hat{\lambda}_i^2!} \exp(-\lambda_i^2(\theta)) \end{aligned}$$

where θ represent the vector of parameters we are trying to estimate.

Consequently, the log likelihood writes:

$$\ln(\mathbf{L}((\hat{\lambda}_i^1)_i, (\hat{\lambda}_i^2)_i, \theta)) = \sum_{i=1}^N \hat{\lambda}_i^1 \ln(\lambda_i^1(\theta)) - \ln(\hat{\lambda}_i^1!) - \lambda_i^1(\theta) + \hat{\lambda}_i^2 \ln(\lambda_i^2(\theta)) - \ln(\hat{\lambda}_i^2!) - \lambda_i^2(\theta),$$

which we rewrite as:

$$\ln(\mathbf{L}((\hat{\lambda}_i^1)_i, (\hat{\lambda}_i^2)_i, \theta)) = \sum_{i=1}^N \hat{\lambda}_i^1 \ln(\lambda_i^1(\theta)) - (\hat{\lambda}_i^1 \ln(\hat{\lambda}_i^1) - \hat{\lambda}_i^1) - \lambda_i^1(\theta) + \hat{\lambda}_i^2 \ln(\lambda_i^2(\theta)) - (\hat{\lambda}_i^2 \ln(\hat{\lambda}_i^2) - \hat{\lambda}_i^2) - \lambda_i^2(\theta) \quad (\text{G2})$$

where:

- $\hat{\lambda}_i^{1,2}$: the estimated intensities of processes 1,2 in bin i
- $\lambda_i^{1,2}$: the theoretical intensities of processes 1 & 2 which depend on both the observation of $\hat{\lambda}$ and the parameters θ we are trying to estimate. In the case of a two dimensional ZHawkes, intensities are defined such that:

$$\lambda_t^j = \lambda_\infty^j + \int_{-\infty}^t h_1^j(t-u) dN_u^1 + \int_{-\infty}^t h_2^j(t-u) dN_u^2 + (\int_{-\infty}^t k_1^j(t-u) dP_u^1)^2 + (\int_{-\infty}^t k_2^j(t-u) dP_u^2)^2.$$

In our discrete environment, we have ordered bins $(t_i)_{i \in [1, N]}$, and thus we can write the intensity of process j in bin M such that

$$\lambda_{t_M}^j = \lambda_\infty^j + \sum_{i=1}^{M-1} h_1^j(t_M - t_i) \hat{\lambda}_i^1 dt + \sum_{i=1}^{M-1} h_2^j(t_M - t_i) \hat{\lambda}_i^2 dt + (\sum_{i=1}^{M-1} k_1^j(t_M - t_i) dP_i^1)^2 + (\sum_{i=1}^{M-1} k_2^j(t_M - t_i) dP_i^2)^2.$$

In the case of exponential kernels, we write $h_i^j(t) = n_i^j \exp(-\kappa_i^j t)$ and $k_i^j(t) = n_{iq}^j \exp(-\kappa_{iq}^j t)$, and the intensity becomes:

$$\begin{aligned} \lambda_{t_M}^j &= \lambda_\infty^j \\ &+ \sum_{i=1}^{M-1} n_i^j \exp(-\kappa_1^j(t_M - t_i)) \hat{\lambda}_i^1 dt + \sum_{i=1}^{M-1} n_i^j \exp(-\kappa_2^j(t_M - t_i)) \hat{\lambda}_i^2 dt \\ &+ (\sum_{i=1}^{M-1} n_{1q}^j \exp(-\kappa_{1q}^j(t_M - t_i)) dP_i^1)^2 + (\sum_{i=1}^{M-1} n_{2q}^j \exp(-\kappa_{2q}^j(t_M - t_i)) dP_i^2)^2. \end{aligned}$$

The parameters we want to estimate in this case are $(n_i^j)_{i,j \in [1, 2]}$, $(n_{i,q}^j)_{i,j \in [1, 2]}$, $(\kappa_i^j)_{i,j \in [1, 2]}$ and $(\kappa_{i,q}^j)_{i,j \in [1, 2]}$.

c. Minimize the log-likelihood

To minimise $-\ln(\mathbf{L}((\hat{\lambda}_i^1)_i, (\hat{\lambda}_i^2)_i, \theta))$, we look at gradient descent methods. We need to calculate the gradient of $-\ln(\mathbf{L}((\hat{\lambda}_i^1)_i, (\hat{\lambda}_i^2)_i, \theta))$ according to the parameters.

Then, with expression (G2), we can write the gradient:

$$\begin{aligned} \nabla(-\ln(\mathbf{L}((\hat{\lambda}_i^1)_i, (\hat{\lambda}_i^2)_i, \theta))) &= \sum_{i=1}^N -\nabla(\hat{\lambda}_i^1 \ln(\lambda_i^1(\theta))) + \nabla(\lambda_i^1(\theta)) - \nabla(\hat{\lambda}_i^2 \ln(\lambda_i^2(\theta))) + \nabla(\lambda_i^2(\theta)) \\ &= \sum_{i=1}^N -\hat{\lambda}_i^1 \nabla(\ln(\lambda_i^1(\theta))) + \nabla(\lambda_i^1(\theta)) - \hat{\lambda}_i^2 \nabla(\ln(\lambda_i^2(\theta))) + \nabla(\lambda_i^2(\theta)) \\ &= \sum_{i=1}^N -\hat{\lambda}_i^1 \frac{\nabla(\lambda_i^1(\theta))}{\lambda_i^1(\theta)} + \nabla(\lambda_i^1(\theta)) - \hat{\lambda}_i^2 \frac{\nabla(\lambda_i^2(\theta))}{\lambda_i^2(\theta)} + \nabla(\lambda_i^2(\theta)) \\ &= \sum_{i=1}^N \nabla \lambda_i^1(\theta) \left(-\frac{\hat{\lambda}_i^1}{\lambda_i^1(\theta)} + 1 \right) + \nabla \lambda_i^2(\theta) \left(-\frac{\hat{\lambda}_i^2}{\lambda_i^2(\theta)} + 1 \right). \end{aligned}$$

In practice, for financial time series, $\hat{\lambda} \equiv \frac{\sigma^2}{dt}$ where σ is the volatility, and dP are the returns.