Adv. Appl. Prob. **53,** 425–462 (2021)
*doi:*10.1017/apr.2020.60

FROM MICROSCOPIC PRICE DYNAMICS TO MULTIDIMENSIONAL ROUGH VOLATILITY MODELS

MATHIEU ROSENBAUM* AND
MEHDI TOMAS** *École Polytechnique*

Abstract

Rough volatility is a well-established statistical stylized fact of financial assets. This property has led to the design and analysis of various new rough stochastic volatility models. However, most of these developments have been carried out in the mono-asset case. In this work, we show that some specific multivariate rough volatility models arise naturally from microstructural properties of the joint dynamics of asset prices. To do so, we use Hawkes processes to build microscopic models that accurately reproduce high-frequency cross-asset interactions and investigate their long-term scaling limits. We emphasize the relevance of our approach by providing insights on the role of microscopic features such as momentum and mean-reversion in the multidimensional price formation process. In particular, we recover classical properties of high-dimensional stock correlation matrices.

Keywords: Rough volatility; multidimensional processes; microstructure; Hawkes processes; limit theorems; high-dimensional correlation matrices

2010 Mathematics Subject Classification: Primary 60F05
Secondary 60F17; 60G55; 62P05

1. Introduction

1.1. A microstructural viewpoint on rough volatility

It is now widely accepted that volatility is rough (see [11] and among others [6, 24]): the log-volatility process is well-approximated by a fractional Brownian motion with small Hurst parameter $H \approx 0.1$, which corresponds to Hölder regularity of order $H - \epsilon, \epsilon > 0$. Furthermore, rough volatility models capture key features of the implied volatility surface and its dynamics (see [3, 9, 17]).

The macroscopic phenomenon of rough volatility is seemingly universal: it is observed for a large class of financial assets and across time periods. This universality may stem from fundamental properties such as market microstructure or no-arbitrage. This has raised interest in building *microscopic* models for market dynamics which reproduce rough volatility at a *macroscopic* scale. For us, the microscopic time scale is of the order of milliseconds, where asset prices are jump processes, while the macroscopic scale is approximately of the order of days, where asset prices appear essentially continuous.

Received 18 October 2019; revision received 28 August 2020.
* Postal address: CMAP, École Polytechnique, Route de Saclay, 91120 Palaiseau, France.
** Postal address: CMAP & LadHyX, École Polytechnique, Route de Saclay, 91120 Palaiseau, France. Email address: mehdi.tomas@polytechnique.edu

Hawkes processes, first introduced in [13, 14, 15] to model earthquake aftershocks, are nowadays very popular to model the high-frequency dynamics of prices of financial assets (see [2] for an overview of applications). In particular, the papers [8, 20, 21] successfully establish a link between rough volatility and history-dependent Hawkes-type point processes which reproduce the following properties:

(i) the no-statistical-arbitrage property (i.e. it is very hard to design strategies which are on average profitable at the high-frequency scale);

(ii) the long-memory property of order flow, due to the splitting of large orders (meta-orders) into smaller orders;

(iii) the high degree of endogeneity of financial markets (i.e. the large majority of market activity (including price moves, cancellations, and market and limit orders) occurs in response to previous market activity, as opposed to exogenous information such as news).

We refer to [8, 12] for details about these three stylized facts. This Hawkes-based microscopic framework can easily account for other features of markets: for example [22] examines the issue of permanent market impact, [10] studies how a bid/ask asymmetry creates a negative price/volatility correlation, and the so-called Zumbach effect is considered in [7].

Inspired by [8, 20, 21] the goal of this paper is to use Hawkes processes to find a micro-founded setting of multivariate rough volatility which

(i) enforces no statistical arbitrage between multiple assets;

(ii) is consistent with the long-memory property of the order flow and the high degree of endogeneity of financial markets; and

(iii) explains stylized facts from the microscopic price formation process, with a focus on the structure of high-dimensional stock correlation matrices.

This approach enables us to characterize the type of price dynamics arising from these constraints. Readers interested in multivariate rough volatility may consult [5] for a general construction of a class of affine multivariate rough covariance models. Our goal is more modest here: we are interested in finding macroscopic dynamics originating from microscopic insights, not in a full mathematical analysis of the class of possible models for multivariate rough volatility. Note also that in the concomitant work [18], the authors study weak solutions of stochastic Volterra equations in a very comprehensive framework. Some of our technical results can be derived from their general approach. In our setting, however, we provide simple and natural proofs inspired by [8, 20, 21] this allows us to emphasize financial interpretations of the results, which are the core of this work.

1.2. Modelling endogeneity of financial markets

For clarity, we first introduce the asymptotic framework which models the high endogeneity of financial markets in the mono-asset case (as [1, 8, 20, 21]), before moving to the multivariate setting of interest. At the high-frequency scale, the price is a piecewise constant process with upward and downward jumps captured by a bi-dimensional counting process $N = \left(N^{1+}, N^{1-}\right)$, with N^{1+} counting the number of upward price moves and N^{1-} the number of downward price moves. Assuming that all jumps are of the same size, the microscopic price of the asset is the difference between the number of upward and the number of downward jumps (where the

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

initial price is set to zero for simplicity) and therefore can be written

$$P_t = N_t^{1+} - N_t^{1-}.$$

Our assumption is that $\boldsymbol{N}$ is a Hawkes process with intensity $\boldsymbol{\lambda} = (\lambda^{1+}, \lambda^{1-})$ such that

$$\lambda_t^{1+} = \mu_t^{1+} + \int_0^t \phi_{1+,1+}(t-s)dN_s^{1+} + \int_0^t \phi_{1+,1-}(t-s)dN_s^{1-},$$
$$\lambda_t^{1-} = \mu_t^{1-} + \int_0^t \phi_{1-,1+}(t-s)dN_s^{1+} + \int_0^t \phi_{1-,1-}(t-s)dN_s^{1-},$$

where $\boldsymbol{\mu} = (\mu^{1+}, \mu^{1-})\colon \mathbb{R}_+ \to \mathbb{R}_+^2$ is called the *baseline* and $\boldsymbol{\phi}\colon \mathbb{R}_+ \to \mathcal{M}_2(\mathbb{R}_+)$ is called the *kernel*. Here we write vectors and matrices in bold, and $\mathcal{M}_{n,m}(X)$ (resp. $\mathcal{M}_n(X)$) denotes the set of X-valued $n \times m$ (resp. $n \times n$) matrices. We can easily interpret the different terms above from a financial perspective:

(i) On the one hand, μ^{1+} (resp. μ^{1-}) is an exogenous source of upward (resp. downward) price moves.

(ii) On the other hand, $\boldsymbol{\phi}$ is an endogenous source of price moves. For example, $\phi_{1+,1-}$ increases the intensity of upward price jumps after a downward price jump, creating a mean-reversion effect (while $\phi_{1+,1+}$ creates a trending effect).

To further encode the long-memory property of the order flow, [8] and [20] consider heavy-tailed kernels where, writing $\rho(\boldsymbol{M})$ for the spectral radius of a matrix $\boldsymbol{M}$, for some $c > 0$ and $\alpha \in (1/2, 1)$ we have

$$\rho\Big(\int_t^\infty \boldsymbol{\phi}(s)ds\Big) \underset{t\to\infty}{\sim} ct^{-\alpha}.$$

Such a model satisfies the stability property of Hawkes processes (see for example [20]) as long as $\rho(\|\boldsymbol{\phi}\|_1) < 1$ (where we write $\|\cdot\|_1$ for the L^1 norm). In fact, calibration of Hawkes processes on financial data suggests that this stability condition is almost violated. To account for this effect, the authors of [8] and [20] model the market up to time T with a Hawkes process $\boldsymbol{N^T}$ of baseline $\boldsymbol{\mu}^T$ and kernel $\boldsymbol{\phi}^T$. The microscopic price until time T is then

$$P_t^{T,1} = N_t^{T,1+} - N_t^{T,1-}.$$

In order to obtain macroscopic dynamics, the time horizon must be large; thus the sequence T_n tends to infinity (from now on, we write T for T_n). As T tends to infinity, $\boldsymbol{\phi}^T$ almost saturates the stability condition:

$$\rho\Big(\|\boldsymbol{\phi}^T\|_1\Big) \underset{T\to\infty}{\to} 1.$$

A macroscopic limit then requires scaling the processes appropriately to obtain a nontrivial limit. Details on the proper rescaling of the processes are given in Section 1.4.

1.3. Multivariate setting

Having described the asymptotic setting in the mono-asset case, we now model m different assets. The associated counting process is now a $2m$-dimensional process $\boldsymbol{N}^T = (N^{T,1+}, N^{T,1-}, N^{T,2+}, \ldots, N^{T,m-})$, and its intensity satisfies

$$\boldsymbol{\lambda}_t^T = \boldsymbol{\mu}_t^T + \int_0^t \boldsymbol{\phi}^T(t-s)d\boldsymbol{N}_s^T.$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

The counting process $\boldsymbol{N}$ includes the upward and downward price jumps of m different assets, and the microscopic price of Asset i, where $1 \leq i \leq m$, is simply

$$P_t^{T,i} = N_t^{T,i+} - N_t^{T,i-}.$$

This allows us to capture correlations between assets, since, focusing for example on Asset 1, we have

$$\lambda_t^{T,1+} = \mu_t^{T,1+} + \int_0^t \phi^T_{1+,1+}(t-s)dN_s^{T,1+} + \int_0^t \phi^T_{1+,1-}(t-s)dN_s^{T,1-} + \int_0^t \phi^T_{1+,2+}(t-s)dN_s^{T,2+} + \int_0^t \phi^T_{1+,2-}(t-s)dN_s^{T,2-} + \cdots .$$

Therefore $\phi^T_{1+,2+}$ increases the intensity of upward jumps on Asset 1 after an upward jump of Asset 2, while $\phi^T_{1+,2+}$ increases the intensity of upward jumps on Asset 1 after a downward jump of Asset 2, etc.

We now need to adapt the nearly-unstable setting to the multidimensional case. Thus we have to find how to saturate the stability condition and to translate the long-memory property of the order flow. In [8], $\boldsymbol{\phi}^T(t)$ is taken diagonalizable (in a basis independent of T and t) with a maximum eigenvalue $\xi^T(t)$ such that

$$\|\xi^T\|_1 \underset{T\to\infty}{\to} 1.$$

However, this structure leads to the same volatility for all assets and thus cannot be a satisfying solution for realistic market dynamics. We take here a sequence of trigonalizable (in a basis $\boldsymbol{O}$ independent of T and t) kernels $\boldsymbol{\phi}^T(t)$ with $n_c > 0$ eigenvalues almost saturating the stability condition. Thus the Hawkes kernel is taken to be of the form

$$\boldsymbol{\phi}^T(t) = \boldsymbol{O}\begin{pmatrix} \boldsymbol{A}^T(t) & \boldsymbol{0} \\ \boldsymbol{B}^T(t) & \boldsymbol{C}^T(t) \end{pmatrix}\boldsymbol{O}^{-1}$$

(using block matrix notation that will be in force throughout the paper), where $\boldsymbol{A}^T : \mathbb{R}_+ \to \mathcal{M}_{n_c}(\mathbb{R})$, $\boldsymbol{B}^T : \mathbb{R}_+ \to \mathcal{M}_{2m-n_c,n_c}(\mathbb{R})$ and $\boldsymbol{C}^T : \mathbb{R}_+ \to \mathcal{M}_{2m-n_c}(\mathbb{R})$. Note that we will see that in the limit, macroscopic volatilities and prices are independent of the chosen basis. We assume that the stability condition is saturated at the speed $T^{-\alpha}$, where $\alpha \in (1/2, 1)$ is again related to the tail of the matrix kernel (see below). The saturation condition translates to

$$T^\alpha\left(\boldsymbol{I} - \int_0^\infty \boldsymbol{A}^T(s)ds\right) \underset{T\to\infty}{\to} \boldsymbol{K},$$

where $\boldsymbol{K}$ is an invertible matrix.

We now need to encode the long-memory property of the order flow. We can expect orders to be sent jointly on different assets (this can be due, for example, to portfolio rebalancing, risk management, or optimal trading) and split under different time scales depending on idiosyncratic components (such as daily traded volume or volatility). Empirically, the approximation that, despite idiosyncrasies, a common time scale for order splitting exists is partially justified: for example [4] shows that market impact, which is directly related to the order flow, is

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

well-approximated by a single time scale for many stocks. Finally, this property is encoded by imposing a heavy-tail condition for $\boldsymbol{A} := \lim_{T\to\infty} \boldsymbol{A}^T$ with the previous exponent α:

$$\alpha x^{\alpha} \int_x^{\infty} \boldsymbol{A}(s)ds \underset{x\to\infty}{\to} \boldsymbol{M},$$

with $\boldsymbol{M}$ an invertible matrix.

1.4. Main results and organization of the paper

In the framework described above, we show that the macroscopic limit of prices is a multivariate version of the rough Heston model introduced in [9, 10], where the volatility process is a solution of a multivariate rough stochastic Volterra equation. Thus we derive a natural multivariate setting for rough volatility using nearly-unstable Hawkes processes.

More precisely, we define the following rescaled processes (see [20] for details), for $t \in [0, 1]$:

$$\boldsymbol{X}_t^T := \frac{1}{T^{2\alpha}} \boldsymbol{N}_{tT}^T, \tag{1}$$

$$\boldsymbol{Y}_t^T := \frac{1}{T^{2\alpha}} \int_0^{tT} \boldsymbol{\lambda}_s ds, \tag{2}$$

$$\boldsymbol{Z}_t^T := T^{\alpha}\left(\boldsymbol{X}_t^T - \boldsymbol{Y}_t^T\right) = \frac{1}{T^{\alpha}} \boldsymbol{M}_{tT}^T, \tag{3}$$

$$\boldsymbol{P}_t^T = \frac{1}{T^{2\alpha}}\left(N_{tT}^{T,1+} - N_{tT}^{T,1-}, \cdots, N_{tT}^{T,m+} - N_{tT}^{T,m-}\right). \tag{4}$$

We refer to $\boldsymbol{P}^T$ as the (rescaled) microscopic price process. Under some additional technical and no-statistical-arbitrage assumptions, there exist an n_c-dimensional process $\tilde{\boldsymbol{V}}$, matrices $\boldsymbol{\Theta}^{\mathbf{1}} \in \mathcal{M}_{n_c}(\mathbb{R})$, $\boldsymbol{\Theta}^{\mathbf{2}} \in \mathcal{M}_{n-n_c}(\mathbb{R})$, $\boldsymbol{\Lambda}_{\mathbf{0}} \in \mathcal{M}_{n_c}(\mathbb{R})$, $\boldsymbol{\Lambda}_{\mathbf{1}} \in \mathcal{M}_{n_c}(\mathbb{R})$, $\boldsymbol{\Lambda}_{\mathbf{2}} \in \mathcal{M}_{n_c, n-n_c}(\mathbb{R})$, $\boldsymbol{\theta}_{\mathbf{0}} \in \mathbb{R}^{n_c}$, and a Brownian motion $\boldsymbol{B}$ such that the following hold:

(i) Any macroscopic limit point $\boldsymbol{P}$ of the sequence $\boldsymbol{P}^T$ satisfies

$$\boldsymbol{P}_t = (\boldsymbol{I} + \boldsymbol{\Delta})^{\top} \boldsymbol{Q} \int_0^t \operatorname{diag}\left(\sqrt{\boldsymbol{V}_s}\right) d\boldsymbol{B}_s,$$

where $\boldsymbol{Q} := (\boldsymbol{e}_1 - \boldsymbol{e}_2 \mid \cdots \mid \boldsymbol{e}_{2m-1} - \boldsymbol{e}_{2m})$, we write ${}^{\top}\boldsymbol{Q}$ for the transpose of $\boldsymbol{Q}$ and $(\boldsymbol{e}_i)_{1\le i\le 2m}$ for the canonical basis of $\mathbb{R}^{2m}$, $\boldsymbol{\Delta} = (\Delta_{ij})_{1\le i,j\le m} \in \mathcal{M}_m(\mathbb{R})$ is defined in Section 3, and $\boldsymbol{V}$ is defined below.

(ii) We have $\boldsymbol{\Theta}^{\mathbf{1}}\tilde{\boldsymbol{V}} = (V^1, \cdots, V^{n_c})$ and $\boldsymbol{\Theta}^{\mathbf{2}}\tilde{\boldsymbol{V}} = (V^{n_c+1}, \cdots, V^n)$.

(iii) Every component of $\tilde{\boldsymbol{V}}$ has pathwise Hölder regularity $\alpha - 1/2 - \epsilon$ for any $\epsilon > 0$.

(iv) For any t in [0,1], $\tilde{\boldsymbol{V}}$ satisfies

$$\tilde{\boldsymbol{V}}_t = \int_0^t (t-s)^{\alpha-1} \boldsymbol{\Lambda}_{\mathbf{1}} \operatorname{diag}\left(\sqrt{\boldsymbol{\Theta}^{\mathbf{1}}\tilde{\boldsymbol{V}}_s}\right) d\boldsymbol{W}_s + \int_0^t (t-s)^{\alpha-1} \boldsymbol{\Lambda}_{\mathbf{2}} \operatorname{diag}\left(\sqrt{\boldsymbol{\Theta}^{\mathbf{2}}\tilde{\boldsymbol{V}}_s}\right) d\boldsymbol{Z}_s + \int_0^t (t-s)^{\alpha-1}\left(\boldsymbol{\theta}_{\mathbf{0}} - \boldsymbol{\Lambda}_{\mathbf{0}}\tilde{\boldsymbol{V}}_s\right) ds,$$

where $\boldsymbol{W} := (B^1, \cdots, B^{n_c})$, $\boldsymbol{Z} := (B^{n_c+1}, \cdots, B^n)$ and we write $\sqrt{\boldsymbol{x}}$ for the componentwise square root of vectors of nonnegative entries.

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Thus the volatility process $\boldsymbol{V}$ is driven by $\tilde{\boldsymbol{V}}$, which represents volatility factors, of which there are as many as there are critical directions.

We can use this result to provide microstructural foundations for some empirical properties of correlation matrices. Informally, considering that our assets have similar self-exciting features in their microscopic dynamics, we show that any macroscopic limit point $\boldsymbol{P}$ of the sequence $\boldsymbol{P}^T$ satisfies

$$\boldsymbol{P}_t = \boldsymbol{\Sigma} \int_0^t \operatorname{diag}\Big(\sqrt{\boldsymbol{V}_s}\Big) d\boldsymbol{B}_s,$$

where $\boldsymbol{W}$ is a Brownian motion, $\boldsymbol{V}$ satisfies a stochastic Volterra equation, and $\boldsymbol{\Sigma}$ has one very large eigenvalue, followed by smaller eigenvalues that we can interpret as due to the presence of sectors, and a bulk of eigenvalues much smaller than the others. This is typical of actual stock correlation matrices (see for example [23] for an empirical study).

The paper is organized as follows. Section 2 rigorously introduces the technical framework sketched in the introduction. In Section 3 we present and discuss the main results, which are then applied in examples developed in Section 4. Proofs can be found in Section 5, while some technical results, including proofs of the various applications, are available in an appendix.

2. Assumptions

Before presenting the main results, we make precise the framework sketched out in the introduction. Examples of Hawkes processes satisfying our assumptions are given in Section 4.

Consider a sequence of measurable functions $\boldsymbol{\phi}^T : \mathbb{R}_+ \to \mathcal{M}_{2m}(\mathbb{R}_+)$ and $\boldsymbol{\mu}^T : \mathbb{R}_+ \to \mathbb{R}_+^{2m}$, where the pair $(\boldsymbol{\mu}^T, \boldsymbol{\phi}^T)$ will be used to model the market dynamics until time T via a Hawkes process $\boldsymbol{N}^T$ of baseline $\boldsymbol{\mu}^T$ and kernel $\boldsymbol{\phi}^T$. Each kernel $\boldsymbol{\phi}^T$ is stable ($\rho\big(\|\boldsymbol{\phi}^T\|_1\big) < 1$).

Assumption 1. *There exists an invertible matrix $\boldsymbol{O}$ such that each $\boldsymbol{\phi}^T$ can be written as*

$$\boldsymbol{\phi}^T = \boldsymbol{O} \begin{pmatrix} \boldsymbol{A}^T & \boldsymbol{0} \\ \boldsymbol{B}^T & \boldsymbol{C}^T \end{pmatrix} \boldsymbol{O}^{-1},$$

where $\boldsymbol{A}^T : \mathbb{R}_+ \to \mathcal{M}_{n_c}(\mathbb{R})$, $\boldsymbol{B}^T : \mathbb{R}_+ \to \mathcal{M}_{2m-n_c,n_c}(\mathbb{R})$, $\boldsymbol{C}^T : \mathbb{R}_+ \to \mathcal{M}_{2m-n_c}(\mathbb{R})$. Furthermore, the sequence $\boldsymbol{\phi}^T$ converges to $\boldsymbol{\phi} : \mathbb{R}_+ \to \mathcal{M}_{2m}(\mathbb{R}_+)$ as T tends to infinity, and, writing $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ for the limits of $\boldsymbol{A}^T, \boldsymbol{B}^T, \boldsymbol{C}^T$ as T tends to infinity, we have $\rho(\int_0^\infty \boldsymbol{C}(s)ds) < 1$.

Additionally, there exist $\alpha \in (1/2, 1)$, invertible matrices $\boldsymbol{K}$ and $\boldsymbol{M}$, and $\boldsymbol{\mu} : [0, 1] \to \mathbb{R}_+$ such that, for all $t \in [0, 1]$, we have

$$T^\alpha \Big(\boldsymbol{I} - \int_0^\infty \boldsymbol{A}^T(s)ds\Big) \underset{T\to\infty}{\to} \boldsymbol{K}, \tag{5}$$

$$\alpha x^\alpha \int_x^\infty \boldsymbol{A}(s)ds \underset{x\to\infty}{\to} \boldsymbol{M}, \tag{6}$$

$$T^{1-\alpha} \boldsymbol{\mu}_{tT}^T \underset{T\to\infty}{\to} \boldsymbol{\mu}_t, \tag{7}$$

where $\boldsymbol{K}\boldsymbol{M}^{-1}$ has strictly positive eigenvalues.

Realistic market dynamics require enforcing no-statistical-arbitrage conditions on the kernels, in the spirit of [20]. To determine which conditions need to be satisfied to prevent such arbitrage, we write the intensity of the counting process $\boldsymbol{\lambda}^T$ using the compensator process $\boldsymbol{M}_t^T := \boldsymbol{N}_t^T - \boldsymbol{\lambda}_t^T$. Writing $*k$ for the convolution product iterated k times (which is defined as

$$\boldsymbol{f}^{*k}(t) = \int_0^t \boldsymbol{f}(s)\boldsymbol{f}^{*(k-1)}(t-s)ds$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

for $k \geq 2$, with $\boldsymbol{f}^{*1} = \boldsymbol{f}$), we have $\boldsymbol{\psi}^T = \sum_{k\geq 1} (\boldsymbol{\phi}^T)^{*k}$ (see for example Proposition 2.1 in [20]). For any $t \in [0, T]$, we have

$$\boldsymbol{\lambda}_t^T = \boldsymbol{\mu}_t^T + \int_0^t \boldsymbol{\psi}^T(t-s)\boldsymbol{\mu}_s^T ds + \int_0^t \boldsymbol{\psi}^T(t-s) d\boldsymbol{M}_s^T. \tag{8}$$

Thus, the expected intensities of upward and downward price jumps of Asset i are

$$\mathbb{E}\big[\lambda_t^{T,i+}\big] = \mu_t^{T,i+} + \sum_{1\leq j\leq 2m} \int_0^t \psi_{i+,j-}^T(t-s)\mu_s^{T,j-} ds + \sum_{1\leq j\leq 2m} \int_0^t \psi_{i+,j+}^T(t-s)\mu_s^{T,j+} ds,$$

$$\mathbb{E}\big[\lambda_t^{T,i-}\big] = \mu_t^{T,i-} + \sum_{1\leq j\leq 2m} \int_0^t \psi_{i-,j-}^T(t-s)\mu_s^{T,j-} ds + \sum_{1\leq j\leq 2m} \int_0^t \psi_{i-,j+}^T(t-s)\mu_s^{T,j+} ds.$$

The above leads us to the following assumption.

Assumption 2. *For any $1 \leq i, j \leq m$, the following hold:*

(i) *No pair-trading arbitrage:* $\psi_{i+,j+}^T + \psi_{i+,j-}^T = \psi_{i-,j+}^T + \psi_{i-,j-}^T$.

(ii) *Suitable asymptotic behaviour of the intensities:*

$$\lim_{T\to\infty}\left(\int_0^\infty \psi_{i+,j+}^T - \int_0^\infty \psi_{i+,j-}^T\right) < \infty.$$

Under the above conditions, if $\mu^{T,i+} = \mu^{T,i-}$ for all $1 \leq i \leq m$, then $\mathbb{E}[\lambda_t^{T,i+}] = \mathbb{E}[\lambda_t^{T,i-}]$ and there are on average as many upward as downward jumps, which we interpret as a no-statistical-arbitrage property.

Define, for any $1 \leq i, j \leq m$,

$$\delta_{ji}^T := \psi_{j+,i+}^T - \psi_{j-,i+}^T, \tag{9}$$

$$\Delta_{ji} := \lim_{T\to\infty} \left\|\psi_{j+,i+}^T\right\|_1 - \left\|\psi_{j-,i+}^T\right\|_1. \tag{10}$$

We can make the following remark.

Remark 1. For any $1 \leq k \leq m$, define $\boldsymbol{e}_{k+} := \boldsymbol{e}_{2k-1}$, $\boldsymbol{e}_{k-} := \boldsymbol{e}_{2k}$, and $\boldsymbol{v}_k := \boldsymbol{e}_{k+} - \boldsymbol{e}_{k-}$. Using Part (i) of Assumption 2 and recalling that $\boldsymbol{\psi}^T : t \mapsto \boldsymbol{\psi}^T(t) \in \mathcal{M}_{2m}(\mathbb{R})$, we have

$$\begin{aligned}
{}^\top\boldsymbol{\psi}^T\boldsymbol{v}_k &= {}^\top\boldsymbol{\psi}^T(\boldsymbol{e}_{k+} - \boldsymbol{e}_{k-}) \\
&= \sum_{i=1}^m \big(\psi_{k+,i+}^T - \psi_{k-,i+}^T\big)\boldsymbol{e}_{i+} + \big(\psi_{k+,i-}^T - \psi_{k-,i-}^T\big)\boldsymbol{e}_{i-} \\
&= \sum_{i=1}^m \big(\psi_{k+,i+}^T - \psi_{k-,i+}^T\big)\boldsymbol{e}_{i+} - \big(\psi_{k+,i+}^T - \psi_{k-,i+}^T\big)\boldsymbol{e}_{i-} \\
&= \sum_{i=1}^m \big(\psi_{k+,i+}^T - \psi_{k-,i+}^T\big)\boldsymbol{v}_i = \sum_{i=1}^m \delta_{ki}^T\boldsymbol{v}_i.
\end{aligned}$$

A sufficient condition for the no-pair-trading-arbitrage condition in Part (i) of Assumption 2 to hold is that, for all $1 \leq i \leq m$,

$${}^\top\boldsymbol{\phi}^T\boldsymbol{v}_i = \sum_{1\leq j\leq m} \Big({}^\top\boldsymbol{\phi}^T\boldsymbol{v}_i \cdot \boldsymbol{v}_j\Big)\boldsymbol{v}_j,$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

since then we have, for any $1 \le k \le m$,

$$\sum_{1\le l\le m} (\psi^T_{k+,l+} - \psi^T_{k-,l+})\boldsymbol{e}_{l+} - (\psi^T_{k+,l+} - \psi^T_{k-,l+})\boldsymbol{e}_{l-} = \sum_{1\le l\le m} (\psi^T_{k+,l+} - \psi^T_{k-,l+})\boldsymbol{e}_{l+} - (\psi^T_{k+,l-} - \psi^T_{k-,l-})\boldsymbol{e}_{l-}.$$

In our applications in Section 4 we will use this condition, as it is easier to check assumptions on $\boldsymbol{\phi}$ than on $\boldsymbol{\psi}$.

3. Main results

We are now in the position to rigorously state the main results of this paper. We use the processes $\boldsymbol{X}^T$, $\boldsymbol{Y}^T$, and $\boldsymbol{Z}^T$ defined in the introduction (see Equations (1), (2) (3)) and write

$$\boldsymbol{O}^{-1} = \begin{pmatrix} \boldsymbol{O}_{11}^{(-1)} & \boldsymbol{O}_{12}^{(-1)} \\ \boldsymbol{O}_{21}^{(-1)} & \boldsymbol{O}_{22}^{(-1)} \end{pmatrix}, \boldsymbol{O} = \begin{pmatrix} \boldsymbol{O}_{11} & \boldsymbol{O}_{12} \\ \boldsymbol{O}_{21} & \boldsymbol{O}_{22} \end{pmatrix}.$$

We set

$$\boldsymbol{\Theta}^1 := \left(\boldsymbol{O}_{11} + \boldsymbol{O}_{12}\left(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds\right)^{-1}\int_0^\infty \boldsymbol{B}(s)ds\right)\boldsymbol{K}^{-1},$$

$$\boldsymbol{\Theta}^2 := \left(\boldsymbol{O}_{21} + \boldsymbol{O}_{22}\left(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds\right)^{-1}\int_0^\infty \boldsymbol{B}(s)ds\right)\boldsymbol{K}^{-1},$$

$$\boldsymbol{\theta}_0 := \begin{pmatrix} \boldsymbol{O}_{11}^{(-1)} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{O}_{12}^{(-1)} \end{pmatrix}\boldsymbol{\mu},$$

$$\boldsymbol{\Lambda} := \frac{\alpha}{\Gamma(1-\alpha)}\boldsymbol{K}\boldsymbol{M}^{-1}.$$

We have the following theorem.

Theorem 1. *The sequence $(\boldsymbol{X}^T, \boldsymbol{Y}^T, \boldsymbol{Z}^T)$ is C-tight (see for example [19]) for the Skorokhod topology. Furthermore, for every limit point $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$ of the sequence, there exist a positive process $\boldsymbol{V}$ and a 2m-dimensional Brownian motion $\boldsymbol{B}$ such that the following hold:*

(i) *We have $\boldsymbol{X}_t = \int_0^t \boldsymbol{V}_s ds$, $\boldsymbol{Z}_t = \int_0^t \operatorname{diag}\left(\sqrt{\boldsymbol{V}_s}\right)d\boldsymbol{B}_s$.*

(ii) *There exists $\tilde{\boldsymbol{V}}$, a process of Hölder regularity $\alpha - 1/2 - \varepsilon$ for any $\varepsilon > 0$, such that $\boldsymbol{\Theta}^{\boldsymbol{1}}\tilde{\boldsymbol{V}} = (V^1, \cdots, V^{n_c})$, $\boldsymbol{\Theta}^{\boldsymbol{2}}\tilde{\boldsymbol{V}} = (V^{n_c+1}, \cdots, V^{2m})$, and $\tilde{\boldsymbol{V}}$ is solution of the following stochastic Volterra equation:*

$$\forall t \in [0, 1], \tilde{\boldsymbol{V}}_t = \frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}(\boldsymbol{\theta}_0 - \tilde{\boldsymbol{V}}_s)ds + \frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\boldsymbol{O}_{11}^{(-1)}\operatorname{diag}\left(\sqrt{\boldsymbol{\Theta}^1\tilde{\boldsymbol{V}}_s}\right)d\boldsymbol{W}_s^{\boldsymbol{1}} + \frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\boldsymbol{O}_{12}^{(-1)}\operatorname{diag}\left(\sqrt{\boldsymbol{\Theta}^2\tilde{\boldsymbol{V}}_s}\right)d\boldsymbol{W}_s^{\boldsymbol{2}}, \quad (11)$$

where $\boldsymbol{W}^{\boldsymbol{1}} := (B^1, \cdots, B^{n_c})$, $\boldsymbol{W}^{\boldsymbol{2}} := (B^{n_c+1}, \cdots, B^{2m})$, $\boldsymbol{\Theta}^{\boldsymbol{1}}$, $\boldsymbol{\Theta}^{\boldsymbol{2}}$, $\boldsymbol{O}_{11}^{(-1)}$, $\boldsymbol{O}_{12}^{(-1)}$, $\boldsymbol{\theta}_{\boldsymbol{0}}$ do not depend on the chosen basis.

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Finally, any limit point $\boldsymbol{P}$ of the rescaled price processes $\boldsymbol{P}^T$ satisfies

$$\boldsymbol{P}_t = (\boldsymbol{I} + \boldsymbol{\Delta})^\top \boldsymbol{Q}\bigg(\int_0^t \text{diag}\Big(\sqrt{\boldsymbol{V}_s}\Big) d\boldsymbol{B}_s + \int_0^t \boldsymbol{\mu}_s ds\bigg),$$

where $\boldsymbol{\Delta}$ is defined in Equation (10).

Theorem 1 links multivariate nearly-unstable Hawkes processes and multivariate rough volatility. We note the following:

(i) The resulting stochastic Volterra equation has nontrivial solutions, as the examples in Section 4 will show.

(ii) From a financial perspective, Theorem 1 shows that the limiting volatility process for a given asset is a sum of several volatility factors. The matrix $\boldsymbol{\Delta}$ mixes them and is therefore responsible for correlations between asset prices. Remarks and comments on $\boldsymbol{I} + \boldsymbol{\Delta}$ are developed in Section 4.

(iii) The theorem implies that adding/removing an asset to/from a market has an impact on the individual volatility of other assets. We can estimate the magnitude of such volatility modifications by calibrating Hawkes processes on price changes.

(iv) Since there is a one-to-one correspondence between the Hurst exponent H and the long-memory parameter of the order flow α, our model yields the same roughness for all assets. Extensions to allow for different exponents to coexist, for example by introducing an asset-dependent scaling through $\boldsymbol{D} = (\alpha_1, \cdots, \alpha_m)$ and studying $\boldsymbol{T}^{-\boldsymbol{D}} \boldsymbol{\lambda}^T_{tT}$, are more intricate. In particular, one needs to use a special function extending the Mittag-Leffler matrix function so that its Laplace transform is of the form $(\boldsymbol{I} + \boldsymbol{\Lambda} t^D)^{-1}$.

4. Applications

In this section, we give two examples of processes obtained through Theorem 3 under different assumptions on the microscopic parameters. In the first example we study the influence of microscopic parameters on the limiting price and volatility processes when modelling two assets. In the second example, we model many different assets to reproduce realistic high-dimensional correlation matrices.

4.1. Influence of microscopic properties on the price dynamics of two correlated assets

Our first model to understand the price formation process focuses on two assets. Let $\mu^1, \mu^2 > 0$, $\alpha \in (1/2, 1)$, $\gamma_1, \gamma_2 \in [0, 1]$, and $H^c_{21}, H^a_{21}, H^c_{12}, H^a_{12} \in [0, 1]$ such that the following hold (here $\sqrt{\cdot}$ denotes the principal square root, so that if $x < 0$, then $\sqrt{x} = i\sqrt{-x}$):

$$0 \le \big(H^c_{12} + H^a_{12}\big)\big(H^c_{21} + H^a_{21}\big) < 1,$$

$$0 \le |\, 1 - (\gamma_1 + \gamma_2) - \sqrt{(H^c_{12} - H^a_{12})(H^c_{21} - H^a_{21}) + (\gamma_1 - \gamma_2)^2}\, | < 1,$$

$$0 \le |\, 1 - (\gamma_1 + \gamma_2) + \sqrt{(H^c_{12} - H^a_{12})(H^c_{21} - H^a_{21}) + (\gamma_1 - \gamma_2)^2}\, | < 1.$$

In the above, the superscript c (resp. a) stands for continuation (resp. alternation) to describe that after a price move in a given direction, H^c (resp. H^a) encodes the tendency to trigger other

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

price moves in the same (resp. the opposite) direction. We now have to choose a kernel which satisfies the various assumptions of Section 2 to model the interactions between our two assets. Theorem 1 states that the only relevant parameters for the macroscopic price are $\boldsymbol{K}$ and $\boldsymbol{M}$. For simplicity we choose the kernel so that $\boldsymbol{M}=\alpha\boldsymbol{I}$. This leads us to define, for $t\geq 0$,

$$\begin{aligned}
\phi_1^T(t)&=(1-\gamma_1)\alpha(1-T^{-\alpha})\mathbb{1}_{t\geq 1}t^{-(\alpha+1)}, & \phi_{21}^{T,c}(t)&=\alpha T^{-\alpha}H_{21}^c\mathbb{1}_{t\geq 1}t^{-(\alpha+1)},\\
\phi_2^T(t)&=\gamma_1\alpha(1-T^{-\alpha})\mathbb{1}_{t\geq 1}t^{-(\alpha+1)}, & \phi_{21}^{T,a}(t)&=\alpha T^{-\alpha}H_{21}^a\mathbb{1}_{t\geq 1}t^{-(\alpha+1)},\\
\tilde\phi_1^T(t)&=(1-\gamma_2)\alpha(1-T^{-\alpha})\mathbb{1}_{t\geq 1}t^{-(\alpha+1)}, & \phi_{12}^{T,c}(t)&=\alpha T^{-\alpha}H_{12}^c\mathbb{1}_{t\geq 1}t^{-(\alpha+1)},\\
\tilde\phi_2^T(t)&=\gamma_2\alpha(1-T^{-\alpha})\mathbb{1}_{t\geq 1}t^{-(\alpha+1)}, & \phi_{12}^{T,a}(t)&=\alpha T^{-\alpha}H_{12}^a\mathbb{1}_{t\geq 1}t^{-(\alpha+1)}.
\end{aligned}$$

For a realistic model, we require the exogenous sources of upward and downward price moves to be equal: $\mu^{1+}=\mu^{1-}$ and $\mu^{2+}=\mu^{2-}$. Thus, the sequences of baselines and kernels are chosen as

$$\boldsymbol{\mu}^T=T^{\alpha-1}\begin{pmatrix}\mu^1\\ \mu^1\\ \mu^2\\ \mu^2\end{pmatrix},\quad \boldsymbol{\phi}^T=\begin{pmatrix}\phi_1^T & \phi_2^T & \phi_{12}^{T,c} & \phi_{12}^{T,a}\\ \phi_2^T & \phi_1^T & \phi_{12}^{T,a} & \phi_{12}^{T,c}\\ \phi_{21}^{T,c} & \phi_{21}^{T,a} & \tilde\phi_1^T & \tilde\phi_2^T\\ \phi_{21}^{T,a} & \phi_{21}^{T,c} & \tilde\phi_2^T & \tilde\phi_1^T\end{pmatrix}.$$

We set

$$\boldsymbol{\chi}:=\frac{\sqrt{2}}{4\gamma_1\gamma_2-\big(H_{12}^c-H_{12}^a\big)\big(H_{21}^c-H_{21}^a\big)}\begin{pmatrix}2\gamma_2 & H_{21}^c-H_{21}^a\\ H_{12}^c-H_{12}^a & 2\gamma_1\end{pmatrix},$$

$$\boldsymbol{\Gamma}:=\frac{1}{1-\big(H_{12}^c+H_{12}^a\big)\big(H_{21}^c+H_{21}^a\big)}\begin{pmatrix}1 & H_{21}^c+H_{21}^a\\ H_{12}^c+H_{12}^a & 1\end{pmatrix}.$$

Applying Theorem 1 yields the following result.

Corollary 1. *Consider any limit point $\boldsymbol{P}$ of $\boldsymbol{P}^T$. Under the above assumptions, it satisfies*

$$\boldsymbol{P}_t=\boldsymbol{\chi}\int_0^t\begin{pmatrix}\sqrt{V_s^1}\,dW_s^1\\ \sqrt{V_s^2}\,dW_s^2\end{pmatrix}, \tag{12}$$

with

$$\begin{aligned}
\begin{pmatrix}V_t^1\\ V_t^2\end{pmatrix}&=\frac{\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)}\int_0^t(t-s)^{\alpha-1}\left(\begin{pmatrix}\mu^1\\ \mu^2\end{pmatrix}-\boldsymbol{\Gamma}\begin{pmatrix}V_s^1\\ V_s^2\end{pmatrix}\right)ds\\
&\quad+\sqrt{2}\frac{\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)}\int_0^t(t-s)^{\alpha-1}\begin{pmatrix}\sqrt{V_s^1}\,dZ_s^1\\ \sqrt{V_s^2}\,dZ_s^2\end{pmatrix},
\end{aligned} \tag{13}$$

where **W** *and* **Z** *are bi-dimensional independent Brownian motions.*

This model helps us understand how microscopic parameters drive the price formation process to generate a macroscopic price and volatility.

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

We begin our remarks with some definitions. We define *momentum* as the trend (i.e., the imbalance between the number of upward and downward jumps) created by jumps of one asset on itself. For example, momentum is strong when the next price jump after an upward price jump on an asset is more likely to be upward than downward. The opposite effect is referred to as *mean-reversion*. For example, the parameter γ_1 controls the intensity of self-induced bid–ask bounce on Asset 1: γ_1 close to zero corresponds to a strong momentum while γ_1 close to one corresponds to a strong mean-reversion.

We define *cross-asset momentum* as the trend created by jumps of one asset on another. For example, cross-asset momentum from Asset 2 to Asset 1 (resp. Asset 1 to Asset 2) appears via $H^c_{21} - H^a_{21}$ (resp. $H^c_{12} - H^a_{12}$): when both $H^c_{21} - H^a_{21}$ and $H^c_{12} - H^a_{12}$ are nil, the prices of Asset 1 and Asset 2 are uncorrelated.

We now turn to comments on the volatility process. Because of its role in the single-asset case, we refer to $\boldsymbol{V}$ as the *fundamental variance*: for example, V^1 is the fundamental variance of Asset 1. The equation satisfied by $\boldsymbol{V}$ depends only on the sum of the feedback effects between the two assets through $H^c_{12} + H^a_{12}$: from a volatility viewpoint, upward and downward jumps have the same impact. Furthermore, we can compute the expected fundamental variance using Mittag-Leffler functions (see Section 5).

Mean-reversion drives down volatility while cross-asset momentum increases it. Indeed, computing $\mathbb{E}[(P^1_t)^2]$, for example, we get

$$\mathbb{E}\Big[\big(P^1_t\big)^2\Big] = 2\frac{4\gamma_2^2 \int_0^t \mathbb{E}\big[V^1_s\big]ds + \big(H^c_{12} - H^a_{12}\big)\big(H^c_{21} - H^a_{21}\big)\int_0^t \mathbb{E}\big[V^2_s\big]ds}{\big[4\gamma_1\gamma_2 - \big(H^c_{12} - H^a_{12}\big)\big(H^c_{21} - H^a_{21}\big)\big]^2}.$$

In particular, increasing γ_1 does not change $\boldsymbol{V}$ but reduces $\mathbb{E}[(P^1_t)^2]$. This example may be particularly relevant to understanding the contribution of Asset 2 to the volatility of Asset 1 through calibration to market data, since if Asset 2 were removed from the market, we would have

$$\mathbb{E}\Big[\big(P^1_t\big)^2\Big] = \frac{1}{2\gamma_1}.$$

Focusing now on the price formation process, we see that it results from a combination of momentum, mean-reversion, and cross-asset momentum. We illustrate this in two extreme cases: when there is no cross-asset momentum and when cross-asset momentum is strong.

(i) When there is no cross-asset momentum (i.e. $H^c_{12} = H^a_{12}$ and $H^c_{21} = H^a_{21}$), at the microscopic scale, a price move on Asset 2 has the same impact on the intensity of upward and downward price moves of Asset 1. Thus the difference between the expected number of upward and downward jumps does not change after a price move on Asset 2: the expected microscopic price of Asset 1 is unaffected, and price moves of Asset 2 generate no trend on Asset 1. This results in macroscopic prices being uncorrelated (see Equation (12)).

(ii) On the other hand, when cross-asset momentum is strong (i.e. $(H^c_{12} - H^a_{12})(H^c_{21} - H^a_{21}) \approx 4\gamma_1\gamma_2$, for example if $H^c_{12} - H^a_{12} = 2\gamma_1\sqrt{1-\epsilon}$ and $H^c_{12} - H^a_{12} = 2\gamma_2\sqrt{1-\epsilon}$ for some small $\epsilon > 0$), at the microscopic scale, a price move on Asset 2 significantly increases the probability of a future price move on Asset 1 in the same direction (and vice versa). In this context we have

$$\boldsymbol{\Delta} + \boldsymbol{I} = \frac{1}{2\gamma_1\gamma_2\epsilon}\begin{pmatrix} \gamma_2 & \gamma_2\sqrt{1-\epsilon} \\ \gamma_1\sqrt{1-\epsilon} & \gamma_1 \end{pmatrix}.$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Using Equation (12) we can check that

$$\frac{\mathbb{E}\left[P_t^1 P_t^2\right]}{\sqrt{\mathbb{E}\left[\left(P_t^1\right)^2\right]\mathbb{E}\left[\left(P_t^2\right)^2\right]}} \underset{\epsilon\to 0}{\to} 1,$$

and prices evolve in unison.

This example underlines that in our approach (thanks to our no-arbitrage constraint) microscopic features transfer to macroscopic properties in an intuitive way.

4.2. Reproducing realistic correlation matrices of a large number of assets using microscopic properties

It is well known that the correlation matrix of stocks has few large eigenvalues outside of a 'bulk' of eigenvalues attributable to noise (see for example [23]). The largest eigenvalue is referred to as the market mode (because the associated eigenvector places a roughly equal weight on each asset) and is much larger than other eigenvalues. Other significant eigenvalues can be related to the presence of sectors: groups of companies with similar characteristics.

How can we provide microstructural foundations for this stylized fact? The large eigenvalue associated to the market mode implies that, in a first approximation, stock prices move together: a price increase on one asset is likely followed by a price increase on all other assets. Translating this into our framework, an upward (resp. downward) jump on an asset increases the probability of an upward (resp. downward) jump on all other assets. We further expect that an upward price move on an asset increases this probability much more on an asset from the same sector than on an unrelated one.

The above remarks lead us to consider a model with the following properties:

(i) All stocks share some fundamental high-frequency properties because they have similar self-excitement parameters in the kernel.

(ii) Stocks have a stronger influence on price changes of stocks within the same sector.

(iii) Within the same sector, all stocks have the same microscopic parameters.

The technical details of our setting are presented in Appendix A.4; here we provide only the elements essential to understanding the framework. Let $\mu^1, \ldots, \mu^m > 0$ be the baselines of each asset, where we assume $\mu^{i+} = \mu^{i-}$ for all $1 \le i \le m$. Using the same notation as before, take $\gamma \in [0, 1]$, $\alpha \in (1/2, 1)$ and $H^c, H^a > 0$. We consider $R > 0$ different sectors, Sector r having m_r stocks. For a pair of stocks which we dub 1 and 2 in analogy to the previous example, we have the following:

(i) The self-excitement parameters are equal: $\gamma_1 = \gamma_2 = \gamma$, where γ is the same for all stocks.

(ii) If Stock 1 and Stock 2 do not belong to the same sector, then $H^c_{21} = H^c_{12} = H^c$ and $H^a_{21} = H^a_{12} = H^a$, where H^c, H^a are the same for all stocks.

(iii) If Stock 1 and Stock 2 belong to the same sector r, $H^c_{21} = H^c_{12} = H^c + H^c_r$, $H^a_{21} = H^a_{12} = H^a + H^a_r$ where H^c_r, H^a_r are the same for all stocks belonging to Sector r.

The asymptotic framework is built as in the previous example, with the details given in the proof of Corollary 2 in Appendix A.4. We write $i_r := m_0 + m_1 + \cdots + m_{r-1}$ for $1 \le r \le R$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

(with the convention $m_0 = 1$), so that stocks from Sector r are indexed from i_r to $i_{r+1} - 1$, and define the following vectors:

$$\boldsymbol{w} := \frac{1}{\sqrt{m}}(\boldsymbol{e}_1 + \cdots + \boldsymbol{e}_m),$$

$$\boldsymbol{w}_r := \frac{1}{\sqrt{m_r}} \sum_{i_r \le i < i_{r+1}} \boldsymbol{e}_i,$$

$$\boldsymbol{\theta} := \sum_{1 \le i \le m} \mu^i \boldsymbol{e}_i.$$

We consider an asymptotic framework where the number of assets will eventually grow to infinity. As will become clear in the proof, the only nontrivial regime appears when

$$H^c, H^a, H_r^c, H_r^a \underset{m\to\infty}{=} \mathcal{O}(m^{-1}).$$

Thus we assume that mH^c, mH^a, mH_r^c, mH_r^a converge to $\bar{H}^c$, $\bar{H}^a$, $\bar{H}_r^c$, $\bar{H}_r^a$ as m tends to infinity. We also assume that the proportion of stocks in a given sector relative to the total number of stocks does not vanish: for each $1 \le r \le R$,

$$\frac{m_r}{m} \underset{m\to\infty}{\to} \eta_r > 0.$$

We define the following constants, which will appear in the price and volatility processes: $\lambda^+ := \bar{H}^c + \bar{H}^a$, $\lambda_r^+ := \bar{H}_r^c + \bar{H}_r^c$, $\lambda^- := \bar{H}^c - \bar{H}^a$, $\lambda_r^- := \bar{H}_r^c - \bar{H}_r^a$. Applying Theorem 1 yields the following result.

Corollary 2. *Consider any limit point $\boldsymbol{P}$ of $\boldsymbol{P}^T$. Under the above assumptions, it satisfies*

$$\boldsymbol{P}_t = \sqrt{2}\boldsymbol{\Sigma}_\varepsilon \int_0^t \operatorname{diag}\big(\sqrt{\boldsymbol{V}_s}\big) d\boldsymbol{B}_s,$$

where $\boldsymbol{B}$ is a Brownian motion;

$$\boldsymbol{\Sigma}_\varepsilon := \left(2\gamma \boldsymbol{I} - \lambda^- \boldsymbol{v}^\top \boldsymbol{v} - \sum_{1\le r\le R} \eta_r \lambda_r^- \boldsymbol{v_r}^\top \boldsymbol{v_r} + \boldsymbol{\varepsilon}\right)^{-1}$$

with $\boldsymbol{\epsilon}$ a deterministic $m \times m$ matrix such that

$$\rho(\boldsymbol{\epsilon}) \underset{m\to\infty}{=} o\big(m^{-1}\big);$$

and $\boldsymbol{V}$ satisfies the stochastic Volterra equation

$$\boldsymbol{V}_t = \frac{\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^t (t-s)^{\alpha-1}(\boldsymbol{\theta} - \boldsymbol{\mathcal{V}}_\epsilon \boldsymbol{V}_s) ds + \frac{\sqrt{2}\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)} \int_0^t (t-s)^{\alpha-1} \operatorname{diag}\big(\sqrt{\boldsymbol{V}_s}\big) d\boldsymbol{Z}_s,$$

with $\boldsymbol{Z}$ a Brownian motion independent from $\boldsymbol{W}$, and

$$\boldsymbol{\mathcal{V}}_\epsilon := \left(\boldsymbol{I} - \lambda^+ \boldsymbol{v}^\top \boldsymbol{v} - \sum_{1\le r\le R} \eta_r \lambda_r^+ \boldsymbol{v_r}^\top \boldsymbol{v_r} + \boldsymbol{\epsilon}\right)^{-1}$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

where $\boldsymbol{\varepsilon}$ is a deterministic $m \times m$ matrix such that

$$\rho(\boldsymbol{\varepsilon}) \underset{m\to\infty}{=} o(m^{-1}).$$

Under the previous corollary, using $\propto$ to denote equality up to a multiplicative constant, the expected fundamental variance can be written as follows using the cumulative Mittag-Leffler function (see Definition 4 in Appendix A.2):

$$\mathbb{E}[\boldsymbol{V}_t] \propto \boldsymbol{F}^{\alpha, \boldsymbol{\mathcal{V}}_\epsilon(t)} \boldsymbol{\theta}.$$

Since

$$\rho(\boldsymbol{\epsilon}) \underset{m\to\infty}{=} o(m^{-1}),$$

we neglect it in further comments and use the approximation $\boldsymbol{\mathcal{V}}_\epsilon \approx \boldsymbol{\mathcal{V}}_0$. Writing ξ for the largest eigenvalue of $\boldsymbol{\mathcal{V}}_0$ and $\boldsymbol{z}$ for the associated eigenvector, and neglecting other eigenvalues (which is reasonable if $\lambda^+ + \sum_{1\le r\le R} \eta_r \lambda_r^+ \approx 1$), from the definition of the Mittag-Leffler function we have

$$\mathbb{E}[\boldsymbol{V}_t] \propto F^{\alpha,\xi}(t)\big({}^\top\boldsymbol{\theta}\boldsymbol{z}\big)\boldsymbol{z}.$$

Making the further approximation that $\eta_r\lambda_r^+$ is independent of r, we have $\boldsymbol{z} \propto (1, \cdots, 1)$ and

$$\begin{aligned}\mathbb{E}\big[\boldsymbol{P}_t{}^\top \boldsymbol{P}_t\big] &\propto \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \mathrm{diag}(\mathbb{E}[\boldsymbol{V}_t])^\top \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \\ &\propto \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \mathrm{diag}(\boldsymbol{z})^\top \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \\ &\propto \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}{}^\top \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \propto \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}{}^2.\end{aligned}$$

Therefore the eigenvectors of $\mathbb{E}[\boldsymbol{P}_t{}^\top \boldsymbol{P}_t]$ are those of $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$. Now, as

$$\rho(\boldsymbol{\varepsilon}) \underset{m\to\infty}{=} o(m^{-1}),$$

we neglect it in further comments and use the approximation $\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}} \approx \boldsymbol{\Sigma}_{\boldsymbol{0}}$. When $\lambda^- + \sum_{1\le r\le R} \eta_r\lambda_r^- \approx 2\gamma$, $\boldsymbol{\Sigma}_{\boldsymbol{0}}$ has one large eigenvalue followed by $R-1$ smaller eigenvalues and much smaller eigenvalues. This is consistent with stylized facts about high-dimensional stock correlation matrices; we have thus built a microscopic model to explain the macroscopic structure of correlation matrices.

The conditions $\lambda^- + \sum_{1\le r\le R} \eta_r\lambda_r^- \approx 1$ and $\lambda^+ + \sum_{1\le r\le R} \eta_r\lambda_r^+ \approx 1$ correspond to the parameters being close to the point where all directions are critical: when $\lambda^- + \sum_{1\le r\le R} \eta_r\lambda_r^- \approx 2\gamma$ or $\lambda^- + \sum_{1\le r\le R} \eta_r\lambda_r^- \approx 1$, the spectral radius of $\int_0^\infty \boldsymbol{C}$ is equal to one and we cannot split the kernel into a critical and a non-critical component.

It would be interesting to study other implications of this model. In particular, we believe that encoding a negative price/volatility correlation into the microscopic parameters could explain the so-called index leverage effect (see [25] for a definition and empirical analysis of this stylized fact).

5. Proof of Theorem 1

We split the proof into four steps. Our approach is inspired by [8, 20, 21]. First, we show that the sequence $\big(\boldsymbol{X}^T, \boldsymbol{Y}^T, \boldsymbol{Z}^T\big)$ is C-tight. Second, we use tightness and representation theorems to find equations satisfied by any limit point $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$ of $\big(\boldsymbol{X}^T, \boldsymbol{Y}^T, \boldsymbol{Z}^T\big)$. Third, properties of

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

the Mittag-Leffler function enable us to prove Equation (11). Fourth and finally, we derive the equation satisfied by any limit point $\boldsymbol{P}$ of $\boldsymbol{P}^T$.

Preliminary lemmas

We start with lemmas that will be useful in the proofs. Lemma A.1 from [8] yields

$$\frac{1}{T^\alpha}\boldsymbol{\lambda}^T_{tT} = \frac{\boldsymbol{\mu}^T_{tT}}{T^\alpha} + \frac{1}{T^\alpha}\int_0^{tT} \boldsymbol{\psi}^T(tT-s)\boldsymbol{\mu}^T_s\,ds + \frac{1}{T^\alpha}\int_0^{tT} \boldsymbol{\psi}^T(tT-s)d\boldsymbol{M}^T_s. \tag{14}$$

Thus to investigate the limit of

$$\frac{1}{T^\alpha}\boldsymbol{\lambda}^T_{\cdot T}$$

we need to study

$$\frac{1}{T^\alpha}\boldsymbol{\psi}^T(T\,\cdot\,),$$

which we will do through its Laplace transform. Given an $L^1(\mathbb{R}_+)$ function f, we write its Laplace transform as $\hat{f}(t) := \int_0^\infty f(x)e^{-tx}dx$ for $t \geq 0$ (and similarly for matrix-valued functions $\boldsymbol{F} = (F_{ij})$ where each $F_{ij} \in L^1(\mathbb{R}_+)$). Note that $\widehat{f^{*k}} = \hat{f}^k$. The following lemma holds.

Lemma 1. *Set, for any $t > 0$,*

$$\boldsymbol{\chi}(s) := \left(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds\right)^{-1}\int_0^\infty \boldsymbol{B}(s)ds.$$

We have the following convergence for any $t > 0$:

$$T^{-\alpha}\widehat{\boldsymbol{\psi}^T(T\,\cdot\,)}(t) \underset{T\to\infty}{\to} \boldsymbol{O}\begin{pmatrix} \left[\frac{\Gamma(1-\alpha)}{\alpha}t^\alpha\boldsymbol{M} + \boldsymbol{K}\right]^{-1} & 0 \\ \boldsymbol{\chi}(s)\left[\frac{\Gamma(1-\alpha)}{\alpha}t^\alpha\boldsymbol{M} + \boldsymbol{K}\right]^{-1} & 0 \end{pmatrix}\boldsymbol{O}^{-1}, \tag{15}$$

where $\boldsymbol{K}$ and $\boldsymbol{M}$ are defined in Equations (5) and (6).

Proof. Define $\boldsymbol{\varphi}^T := \boldsymbol{O}^{-1}\hat{\boldsymbol{\phi}}^T\boldsymbol{O}$. Then

$$\hat{\boldsymbol{\psi}}^T(t) = \sum_{k\geq 1}\hat{\boldsymbol{\phi}}^{T,*k} = \boldsymbol{O}\big(\boldsymbol{I} - \hat{\boldsymbol{\varphi}}^T\big)^{-1}\hat{\boldsymbol{\varphi}}^T\boldsymbol{O}^{-1}.$$

We can use the shape of $\boldsymbol{\varphi}^T$ and matrix block inversion to rewrite this expression. Doing so, we find

$$\hat{\boldsymbol{\psi}}^T(t) = \boldsymbol{O}\begin{pmatrix} \left(\boldsymbol{I} - \hat{A}^T(t)\right)^{-1}\hat{A}^T(t) & 0 \\ \left(\boldsymbol{I} - \hat{C}^T(t)\right)^{-1}\hat{B}^T(t)\left(\boldsymbol{I} - \hat{A}^T(t)\right)^{-1}\hat{A}^T(t) - \left(\boldsymbol{I} - \hat{C}^T(t)\right)^{-1}\hat{B}^T(t) & \left(\boldsymbol{I} - \hat{C}^T(t)\right)^{-1}\hat{C}^T(t) \end{pmatrix}\boldsymbol{O}^{-1}.$$

To derive the limiting process, we use Equations (5) and (6). Using integration by parts and a Tauberian theorem as in [8, 21], we have

$$\int_0^\infty \boldsymbol{A}^T(s)ds - \hat{A}^T(t/T) \underset{T\to\infty}{=} \frac{\Gamma(1-\alpha)}{\alpha}t^\alpha\boldsymbol{M}T^{-\alpha} + o(T^{-\alpha})$$

$$\boldsymbol{I} - \int_0^\infty \boldsymbol{A}^T(s)ds \underset{T\to\infty}{=} \boldsymbol{K}T^{-\alpha} + o(T^{-\alpha}).$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Therefore

$$T\Big(\boldsymbol{I}-\hat{A}^T(t/T)\Big)=T\Big(\int_0^\infty \boldsymbol{A}^T(s)ds-\hat{A}^T(t/T)\Big)+T\Big(\boldsymbol{I}-\int_0^\infty \boldsymbol{A}^T(s)ds\Big)$$

$$\underset{T\to\infty}{=}\left[\frac{\Gamma(1-\alpha)}{\alpha}t^\alpha\boldsymbol{M}+\boldsymbol{K}\right]T^{1-\alpha}+o(T^{1-\alpha}).$$

Consequently

$$T^{\alpha-1}T\Big(\boldsymbol{I}-\hat{A}^T(t/T)\Big)\underset{T\to\infty}{=}\frac{\Gamma(1-\alpha)}{\alpha}t^\alpha\boldsymbol{M}+\boldsymbol{K}+o(1).$$

By Assumption 1 $\boldsymbol{M}$ is invertible and $\boldsymbol{K}\boldsymbol{M}^{-1}$ has strictly positive eigenvalues. Thus $\boldsymbol{M}t+\boldsymbol{K}=(\boldsymbol{K}\boldsymbol{M}^{-1}+t\boldsymbol{I})\boldsymbol{M}$ is invertible for any $t\geq 0$. The Laplace transform of $T^{-\alpha}\boldsymbol{\psi}^T(T\,\cdot)$ being $T^{1-\alpha}\widehat{\boldsymbol{\psi}}^T(\cdot/T)$, we have proved that for any $t\geq 0$,

$$T^{-\alpha}\widehat{\boldsymbol{\psi}^T(T\,\cdot)}(t)\underset{T\to\infty}{\to}\boldsymbol{O}\begin{pmatrix}\left[\frac{\Gamma(1-\alpha)}{\alpha}t^\alpha\boldsymbol{M}+\boldsymbol{K}\right]^{-1} & 0\\ \Big(\boldsymbol{I}-\int_0^\infty \boldsymbol{C}(s)ds\Big)^{-1}\int_0^\infty \boldsymbol{B}(s)ds\left[\frac{\Gamma(1-\alpha)}{\alpha}t^\alpha\boldsymbol{M}+\boldsymbol{K}\right]^{-1} & 0\end{pmatrix}\boldsymbol{O}^{-1}.$$

□

We show in the technical appendix that the inverse Laplace transform of $\boldsymbol{\Lambda}(t^\alpha\boldsymbol{I}+\boldsymbol{\Lambda})^{-1}$, where $\boldsymbol{\Lambda}\in\mathcal{M}_n(\mathbb{R})$ has positive eigenvalues, is a simple extension of the Mittag-Leffler density function to matrices (see Definition 4 in the appendix), denoted by $\boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}$. Thus we define, for any $t\in[0,1]$,

$$\boldsymbol{f}(t):=\boldsymbol{O}\begin{pmatrix}\boldsymbol{K}^{-1}\boldsymbol{f}^{\alpha,\frac{\alpha}{\Gamma(1-\alpha)}\boldsymbol{K}\boldsymbol{M}^{-1}} & \boldsymbol{0}\\ \Big(\boldsymbol{I}-\int_0^\infty \boldsymbol{C}(s)ds\Big)^{-1}\int_0^\infty \boldsymbol{B}(s)ds\boldsymbol{K}^{-1}\boldsymbol{f}^{\alpha,\frac{\alpha}{\Gamma(1-\alpha)}\boldsymbol{K}\boldsymbol{M}^{-1}} & \boldsymbol{0}\end{pmatrix}\boldsymbol{O}^{-1}. \tag{16}$$

The following lemma shows the weak convergence of $\boldsymbol{\psi}^T$ to $\boldsymbol{f}$.

Lemma 2. *For any bounded measurable function g and $1\leq i,j\leq n$*

$$\int_{[0,1]}g(x)T^{-\alpha}\psi_{ij}^T(Tx)dx\underset{T\to\infty}{\to}\int_{[0,1]}g(x)f_{ij}(x)dx.$$

Proof. First note that when $\|f_{ij}\|_1=0$ (which implies $f_{ij}=0$), using Equation (15) with $t=0$ we have

$$\Big\|T^{1-\alpha}\psi_{ij}^T\Big\|_1\underset{T\to\infty}{\to}\|f_{ij}\|_1=0,$$

which implies, since $1-\alpha\geq 0$, that

$$\Big\|\psi_{ij}^T\Big\|_1\underset{T\to\infty}{\to}0.$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Therefore, as $\psi_{ij}^T \geq 0$, for any bounded measurable function g we have

$$\left| \int_{[0,1]} g(x) T^{-\alpha} \psi_{ij}^T(Tx) dx \right| \leq c \int_{[0,1]} T^{-\alpha} \psi_{ij}^T(Tx) dx \leq c \left\| T^{1-\alpha} \psi_{ij}^T \right\|_1,$$

and the result holds. Assume now that $\|f_{ij}\|_1 > 0$. It will be convenient for us to proceed with random variables, so define

$$\rho_{ij}^T := \frac{T^{-\alpha} \psi_{ij}^T(T\,\cdot\,)}{\left\| T^{1-\alpha} \psi_{ij}^T \right\|_1}.$$

We can view ρ_{ij}^T as the density of a random variable taking values in [0, 1], say S. Lemma 5 gives the convergence of the characteristic functions of S to

$$\hat{\rho}_{ij} := \frac{\hat{f}_{ij}}{\|f_{ij}\|_1}.$$

Since ρ_{ij} is continuous (as ψ_{ij}^T is continuous), Lévy's continuity theorem guarantees that ρ_{ij}^T converges weakly to ρ_{ij}. Therefore, for any bounded measurable function g,

$$\int_{[0,1]} g(x) \rho_{ij}^T(x) dx \underset{T\to\infty}{\to} \int_{[0,1]} g(x) \rho_{ij}(x) dx,$$

$$\int_{[0,1]} g(x) \frac{T^{-\alpha} \psi_{ij}^T(Tx)}{\left\| T^{1-\alpha} \psi_{ij}^T \right\|_1} dx \underset{T\to\infty}{\to} \int_{[0,1]} g(x) \frac{f_{ij}(x)}{\|f_{ij}\|_1} dx.$$

Equation (15) implies

$$\left\| T^{1-\alpha} \psi_{ij}^T \right\|_1 \underset{T\to\infty}{\to} \|f_{ij}\|_1,$$

so that together with the above we have

$$\int_{[0,1]} g(x) T^{-\alpha} \psi_{ij}^T(Tx) dx \underset{T\to\infty}{\to} \int_{[0,1]} g(x) f_{ij}(x) dx. \qquad \square$$

We introduce the cumulative functions

$$\boldsymbol{F}^T(t) = \int_0^t T^{-\alpha} \boldsymbol{\psi}^T(Ts) ds,$$

$$\boldsymbol{F}(t) = \int_0^t \boldsymbol{f}(s) ds.$$

We have just shown in particular that $\boldsymbol{F}^T$ converges pointwise to $\boldsymbol{F}$ and therefore, by Dini's theorem, converges uniformly to $\boldsymbol{F}$.

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

5.1. Step 1: C-tightness of $\left(\boldsymbol{X}^T, \boldsymbol{Y}^T, \boldsymbol{Z}^T\right)$

Recall the definition of the rescaled processes

$$\begin{aligned}
\boldsymbol{X}_t^T &:= \frac{1}{T^{2\alpha}} \boldsymbol{N}_{tT}^T, \\
\boldsymbol{Y}_t^T &:= \frac{1}{T^{2\alpha}} \int_0^{tT} \boldsymbol{\lambda}_s ds, \\
\boldsymbol{Z}_t^T &:= T^\alpha \left(\boldsymbol{X}_t^T - \boldsymbol{Y}_t^T\right) = \frac{1}{T^\alpha} \boldsymbol{M}_{tT}^T.
\end{aligned}$$

As in [8] and [21] we show that the limiting processes of $\boldsymbol{X}^T$ and $\boldsymbol{Y}^T$ are the same and that the limiting process of $\boldsymbol{Z}^T$ is the quadratic variation of the limiting process of $\boldsymbol{X}^T$. We have the following proposition.

Proposition 1. $\left(C\text{-tightness of } \left(\boldsymbol{X}^T, \boldsymbol{Y}^T, \boldsymbol{Z}^T\right)\right)$

The sequence $\left(\boldsymbol{X}^T, \boldsymbol{Y}^T, \boldsymbol{Z}^T\right)$ *is C-tight, and if* $(\boldsymbol{X}, \boldsymbol{Z})$ *is a limit point of* $\left(\boldsymbol{X}^T, \boldsymbol{Z}^T\right)$*, then* $\boldsymbol{Z}$ *is a continuous martingale with* $[\boldsymbol{Z}, \boldsymbol{Z}] = \operatorname{diag}(\boldsymbol{X})$*. Furthermore, we have the convergence in probability*

$$\sup_{t\in[0,1]} \left\| \boldsymbol{Y}_t^T - \boldsymbol{X}_t^T \right\|_2 \underset{T\to\infty}{\overset{\mathbb{P}}{\to}} 0.$$

Proof. The proof is essentially the same as in [8], adapted to our structure of Hawkes processes. Given $t \in [0, T]$, we have

$$\boldsymbol{\lambda}_t^T = \boldsymbol{\mu}_t^T + \int_0^t \boldsymbol{\psi}^T(t-s)\boldsymbol{\mu}_s^T ds + \int_0^t \boldsymbol{\psi}^T(t-s) d\boldsymbol{M}_s^T,$$

and therefore

$$\begin{aligned}
\mathbb{E}[\boldsymbol{N}_t^T] &= \mathbb{E}\Big[\int_0^T \boldsymbol{\lambda}_s^T ds\Big] \\
&= \int_0^T \boldsymbol{\mu}_t^T dt + \int_0^T \int_0^t \boldsymbol{\psi}^T(t-s)\boldsymbol{\mu}_s^T ds dt \le cT^{2\alpha} \|\boldsymbol{\mu}\|_\infty,
\end{aligned}$$

where we have used the convergence of $T^{1-\alpha}\boldsymbol{\mu}_{T\cdot}^T$ (see Equation (7)) together with the weak convergence of $T^{-\alpha}\boldsymbol{\psi}^T(T\,\cdot\,)$ (see Lemma 2). It follows then that

$$\mathbb{E}\left[\boldsymbol{X}_1^T\right] = \mathbb{E}\left[\boldsymbol{Y}_1^T\right] \le c,$$

and since the processes are increasing, $\boldsymbol{X}^T$ and $\boldsymbol{Y}^T$ are tight. As the maximum jump size of $\boldsymbol{X}^T$ and $\boldsymbol{Y}^T$ tends to 0, we have the C-tightness of $(\boldsymbol{X}^T, \boldsymbol{Y}^T)$. Since $\boldsymbol{N}^T$ is the quadratic variation of $\boldsymbol{M}^T$, $(M^{T,i})^2 - N^{T,i}$ is an L^2 martingale starting at 0, and Doob's inequality yields

$$\begin{aligned}
\sum_{1\le i\le n} \mathbb{E}\Big[\sup_{t\in[0,1]} \left(X_t^{T,i} - Y_t^{T,i}\right)^2\Big] &\le 4 \sum_{1\le i\le n} \mathbb{E}\Big[\left(X_1^{T,i} - Y_1^{T,i}\right)^2\Big] \\
&\le 4T^{-4\alpha} \sum_{1\le i\le n} \mathbb{E}\Big[\left(M_T^{T,i}\right)^2\Big] \\
&\le 4T^{-4\alpha} \sum_{1\le i\le n} \mathbb{E}\Big[N_T^{T,i}\Big] \\
&\le cT^{-2\alpha}.
\end{aligned}$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Using the same approach as in [8] we conclude that $\boldsymbol{Z}$ is a continuous martingale and $[\boldsymbol{Z}, \boldsymbol{Z}]$ is the limit of $[\boldsymbol{Z}^T, \boldsymbol{Z}^T]$. □

5.2. Step 2: Rewriting of limit points of $(\boldsymbol{X}^T, \boldsymbol{Y}^T, \boldsymbol{Z}^T)$

By Proposition 1, for any limit point $(\boldsymbol{X}, \boldsymbol{Y})$ of $(\boldsymbol{X}^T, \boldsymbol{Y}^T)$, we have $\boldsymbol{X} = \boldsymbol{Y}$ almost surely. We use $\boldsymbol{Y}^T$ to derive an equation for $\boldsymbol{X}$. As

$$\boldsymbol{Y}^T = \frac{1}{T^{2\alpha}} \int_0^{tT} \boldsymbol{\lambda}_s^T ds,$$

we first study $\boldsymbol{\lambda}_{sT}^T$. Using Equation (14), for any $t \in [0, T]$ we have

$$\begin{aligned}\int_0^t \boldsymbol{\lambda}_s^T ds &= \int_0^t \boldsymbol{\mu}_s^T ds + \int_0^t \int_0^u \boldsymbol{\psi}^T(s-u)\boldsymbol{\mu}_u^T du ds + \int_0^t \boldsymbol{\psi}^T(t-s)\boldsymbol{M}_s^T ds \\ &= \int_0^t \boldsymbol{\mu}_s^T ds + \int_0^t \boldsymbol{\psi}^T(t-s)\int_0^s \boldsymbol{\mu}_u^T du ds + \int_0^t \boldsymbol{\psi}^T(t-s)\boldsymbol{M}_s^T ds.\end{aligned}$$

Thus, for any $t \in [0, 1]$, a change of variables leads to

$$\begin{aligned}\int_0^{tT} \boldsymbol{\lambda}_s^T ds &= \int_0^{tT} \boldsymbol{\mu}_s^T ds + \int_0^{tT} \boldsymbol{\psi}^T(tT-s)\int_0^s \boldsymbol{\mu}_u^T du ds + \int_0^{tT} \boldsymbol{\psi}^T(tT-s)\boldsymbol{M}_s^T ds \\ &= \int_0^{tT} \boldsymbol{\mu}_s^T ds + T\int_0^t \boldsymbol{\psi}^T(tT-sT)\int_0^{sT} \boldsymbol{\mu}_u^T du ds + T\int_0^t \boldsymbol{\psi}^T(tT-sT)\boldsymbol{M}_{sT}^T ds \\ &= T\int_0^t \boldsymbol{\mu}_{sT}^T ds + T\int_0^t \boldsymbol{\psi}^T(T(t-s))\int_0^{sT} \boldsymbol{\mu}_u^T du ds + T\int_0^t \boldsymbol{\psi}^T(T(t-s))\boldsymbol{M}_{sT}^T ds.\end{aligned}$$

Therefore

$$T^{2\alpha}\boldsymbol{Y}_t^T = T\int_0^t \boldsymbol{\mu}_{sT}^T ds + T\int_0^t \boldsymbol{\psi}^T(T(t-s))\int_0^{sT} \boldsymbol{\mu}_u^T du ds + T\int_0^t \boldsymbol{\psi}^T(T(t-s))\boldsymbol{M}_{sT}^T ds \quad (17)$$

$$=: T^{2\alpha}\left(\boldsymbol{Y}_t^{T,1} + \boldsymbol{Y}_t^{T,2} + \boldsymbol{Y}_t^{T,3}\right), \quad (18)$$

with obvious notation. Thus, to obtain our limit we use the convergence properties of $\boldsymbol{F}^T$ which we derived earlier. We have the following proposition.

Proposition 2. *Let $(\boldsymbol{X}, \boldsymbol{Z})$ be a limit point of $(\boldsymbol{X}^T, \boldsymbol{Z}^T)$. Then, for any $t \in [0, 1]$, we have*

$$\boldsymbol{X}_t = \int_0^t \boldsymbol{F}(t-s)\boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{F}(t-s)d\boldsymbol{Z}_s.$$

Proof. Let $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$ be a limit point of $(\boldsymbol{X}^T, \boldsymbol{Y}^T, \boldsymbol{Z}^T)$. First, since

$$T^{1-\alpha}\boldsymbol{\mu}_{tT}^T \underset{T\to\infty}{\to} \boldsymbol{\mu}_t$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

(see Equation (7)), $\boldsymbol{Y}_t^{T,1}$ converges to 0 as T tends to infinity. Moving on to $\boldsymbol{Y}^{T,2}$, by integration by parts, for any $t \in [0, 1]$ we obtain

$$\begin{aligned}\boldsymbol{Y}_t^{T,2} &= \int_0^t T^{1-\alpha}\boldsymbol{\psi}^T(T(t-s))T^{-\alpha}\int_0^{sT}\boldsymbol{\mu}_u^T du ds \\ &= \left[\boldsymbol{F}^T(t-s)T^{-\alpha}\int_0^{sT}\boldsymbol{\mu}_u^T du\right]_0^t + \int_0^t \boldsymbol{F}^T(t-s)T^{1-\alpha}\boldsymbol{\mu}_{sT}^T ds \\ &= \int_0^t \boldsymbol{F}^T(t-s)T^{1-\alpha}\boldsymbol{\mu}_{sT}^T ds.\end{aligned}$$

Using Equation (7) again, together with the uniform convergence of $\boldsymbol{F}^T$ (see Lemma 2), we have the convergence

$$\boldsymbol{Y}_t^{T,2} \underset{T\to\infty}{\to} \int_0^t \boldsymbol{F}(t-s)\boldsymbol{\mu}_s ds.$$

Finally, $\boldsymbol{Y}_t^{T,3}$ can be written as

$$\begin{aligned}\boldsymbol{Y}_t^{T,3} = T^{1-2\alpha}\int_0^t \boldsymbol{\psi}^T(T(t-s))\boldsymbol{M}_{sT}^T ds = \int_0^t \boldsymbol{F}^T(t-s)d\boldsymbol{Z}_s^T \\ = \int_0^t \boldsymbol{F}(t-s)d\boldsymbol{Z}_s + \int_0^t \boldsymbol{F}(t-s)(d\boldsymbol{Z}_s^T - d\boldsymbol{Z}_s) + \int_0^t (\boldsymbol{F}^T(t-s) - \boldsymbol{F}(t-s))d\boldsymbol{Z}_s^T.\end{aligned}$$

The Skorokhod representation theorem applied to $(\boldsymbol{Z}^T, \boldsymbol{Z})$ yields the existence of copies in law $(\tilde{\boldsymbol{Z}}^T, \tilde{\boldsymbol{Z}})$, with $\tilde{\boldsymbol{Z}}^T$ converging almost surely to $\tilde{\boldsymbol{Z}}$. We proceed with $(\tilde{\boldsymbol{Z}}^T, \tilde{\boldsymbol{Z}})$ and keep the previous notation. The stochastic Fubini theorem [27] gives, almost surely,

$$\int_0^t \boldsymbol{F}(t-s)(d\boldsymbol{Z}_s^T - d\boldsymbol{Z}_s) = \int_0^t \boldsymbol{f}(s)(\boldsymbol{Z}_{t-s}^T - \boldsymbol{Z}_{t-s})ds.$$

From the dominated convergence theorem we obtain the almost sure convergence

$$\int_0^t \boldsymbol{f}(s)(\boldsymbol{Z}_{t-s}^T - \boldsymbol{Z}_{t-s})ds \underset{T\to\infty}{\to} 0.$$

Furthermore, since $[\boldsymbol{Z}^T, \boldsymbol{Z}^T] = \mathrm{diag}(\boldsymbol{X}^T)$ we have

$$\begin{aligned}\sum_{1\le i\le n}\mathbb{E}\left[\left(\int_0^t (\boldsymbol{F}^T(t-s) - \boldsymbol{F}(t-s))d\boldsymbol{Z}_s^T\right)_i^2\right] \le \sum_{1\le i,j\le n}\int_0^t \big(F_{ij}^T(t-s) \\ - F_{ij}(t-s)\big)^2 T^{1-\alpha}\mathbb{E}\big[\lambda_{sT}^{T,j}\big]ds.\end{aligned}$$

Using Equation (14) together with Lemma 1, we can bound $\mathbb{E}[\lambda_{sT}^{T,j}]$ independently of T, and

$$\sum_{1\le i\le n}\mathbb{E}\left[\left(\int_0^t (\boldsymbol{F}^T(t-s) - \boldsymbol{F}(t-s))d\boldsymbol{Z}_s^T\right)_i^2\right] \le c\sum_{1\le i,j\le n}\int_0^t \Big(F_{ij}^T(t-s) - F_{ij}(t-s)\Big)^2 ds.$$

The right-hand side converges to 0 by the dominated convergence theorem together with the uniform convergence of $\boldsymbol{F}^T$ to $\boldsymbol{F}$ (see Lemma 2). From Proposition 1 we know that $\boldsymbol{Y} = \boldsymbol{X}$ almost surely. Putting everything together, almost surely,

$$\boldsymbol{X}_t = \int_0^t \boldsymbol{F}(t-s)\boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{F}(t-s) d\boldsymbol{Z}_s.$$

This is valid for any limit point $(\boldsymbol{X}, \boldsymbol{Z})$ of $(\boldsymbol{X}^T, \boldsymbol{Z}^T)$, which concludes the proof. □

The previous proposition gives suitable martingale properties of limit points of $\boldsymbol{Z}^T$ to apply the martingale representation theorem, which is the topic of the following proposition.

Proposition 3. *Let $(\boldsymbol{X}, \boldsymbol{Z})$ be a limit point of $(\boldsymbol{X}^T, \boldsymbol{Z}^T)$. There exists, up to an extension of the original probability space, an n-dimensional Brownian motion $\boldsymbol{B}$ and a nonnegative process $\boldsymbol{V}$ such that, for any $t \in [0, 1]$, one has*

$$\boldsymbol{X}_t = \int_0^t \boldsymbol{V}_s ds,$$

$$\boldsymbol{Z}_t = \int_0^t \operatorname{diag}\left(\sqrt{\boldsymbol{V}_s}\right) d\boldsymbol{B}_s,$$

$$\boldsymbol{V}_t = \int_0^t \boldsymbol{f}(t-s)\boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{f}(t-s)\operatorname{diag}\left(\sqrt{\boldsymbol{V}_s}\right) d\boldsymbol{B}_s.$$

Proof. This proof relies on the martingale representation theorem applied to $\boldsymbol{Z}$. Consider a limit point $(\boldsymbol{X}, \boldsymbol{Z})$ of $(\boldsymbol{X}^T, \boldsymbol{Z}^T)$. Following the proof of Theorem 3.2 in [21] and using Proposition 2, $\boldsymbol{X}$ can be written as the integral of a process $\boldsymbol{V}$, i.e.

$$\boldsymbol{X}_t = \int_0^t \boldsymbol{V}_s ds,$$

with $\boldsymbol{V}$ satisfying the equation

$$\boldsymbol{V}_t = \int_0^t \boldsymbol{f}(t-s)\boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{f}(t-s) d\boldsymbol{Z}_s.$$

Therefore, as $[\boldsymbol{Z}, \boldsymbol{Z}]_t = \operatorname{diag}(\boldsymbol{X}_t) = \operatorname{diag}(\int_0^t \boldsymbol{V}_s ds)$ and $\boldsymbol{Z}$ is a continuous martingale, by the martingale representation theorem (see for example Theorem 3.9 from [26]), there exists (up to an enlargement of the probability space) a multivariate Brownian motion $\boldsymbol{B}$ and a predictable square-integrable process $\boldsymbol{H}$ such that

$$\boldsymbol{Z}_t = \int_0^t \boldsymbol{H}_s d\boldsymbol{B}_s.$$

Furthermore, note that $\boldsymbol{V}$ is a nonnegative process (as $\boldsymbol{X}$ is a nondecreasing process), and we have

$$\boldsymbol{Z}_t = \int_0^t \operatorname{diag}\left(\sqrt{\boldsymbol{V}_s}\right)\operatorname{diag}\left(\sqrt{\boldsymbol{V}_s}\right)^{-1}\boldsymbol{H}_s d\boldsymbol{B}_s.$$

A simple computation shows that, since $[\boldsymbol{Z}, \boldsymbol{Z}]_t = \int_0^t \boldsymbol{H}_s{}^{\top}\boldsymbol{H}_s ds = \boldsymbol{X}_t = \int_0^t \boldsymbol{V}_s ds$, the process $\tilde{\boldsymbol{B}}_t := \int_0^t \operatorname{diag}\left(\sqrt{\boldsymbol{V}_s}\right)^{-1}\boldsymbol{H}_s d\boldsymbol{B}_s$ is a Brownian motion. Finally,

$$\boldsymbol{V}_t = \int_0^t \boldsymbol{f}(t-s)\boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{f}(t-s)\operatorname{diag}\left(\sqrt{\boldsymbol{V}_s}\right) d\tilde{\boldsymbol{B}}_s.$$ □

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

A straightforward application of Lemma 4.4 and Lemma 4.5 in [21] yields the following lemma.

Lemma 3. *Consider a (weak) nonnegative solution* $\boldsymbol{V}$ *of the stochastic Volterra equation*

$$\boldsymbol{V}_t = \int_0^t \boldsymbol{f}(t-s)\boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{f}(t-s)\text{diag}\left(\sqrt{\boldsymbol{V}_s}\right)d\boldsymbol{B}_s,$$

where $\boldsymbol{B}$ *is a Brownian motion. Then every component of* $\boldsymbol{V}$ *has pathwise Hölder regularity* $\alpha - 1/2 - \epsilon$ *for any* $\epsilon > 0$.

5.3. Step 3: Proof of Equation (11)

Properties of the Mittag-Leffler function (as in [8]) enable us to rewrite the previous stochastic differential equation using power-law kernels, which is the subject of the next proposition. Let

$$\boldsymbol{\Theta}^1 := \left(\boldsymbol{O_{11}} + \boldsymbol{O_{12}}\left(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds\right)^{-1}\int_0^\infty \boldsymbol{B}(s)ds\right)\boldsymbol{K}^{-1},$$

$$\boldsymbol{\Theta}^2 := \left(\boldsymbol{O_{21}} + \boldsymbol{O_{22}}\left(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds\right)^{-1}\int_0^\infty \boldsymbol{B}(s)ds\right)\boldsymbol{K}^{-1},$$

$$\boldsymbol{\Lambda} := \frac{\alpha}{\Gamma(1-\alpha)}\boldsymbol{K}\boldsymbol{M}^{-1}.$$

Proposition 4. *Given an m-dimensional Brownian motion* $\boldsymbol{B}$, *a nonnegative process* $\boldsymbol{V}$ *is a solution of the stochastic differential equation*

$$\boldsymbol{V}_t = \int_0^t \boldsymbol{f}(t-s)\boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{f}(t-s)\text{diag}\left(\sqrt{\boldsymbol{V}_s}\right)d\boldsymbol{B}_s$$

if and only if there exists a process $\tilde{\boldsymbol{V}}$ *of Hölder regularity* $\alpha - 1/2 - \epsilon$ *for any* $\epsilon > 0$ *such that* $\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}_t = (V^1, \cdots, V^{n_c})$ *and* $\boldsymbol{\Theta^2}\tilde{\boldsymbol{V}}_t = (V^{n_c+1}, \cdots, V^{2m})$ *are nonnegative processes and* $\tilde{\boldsymbol{V}}$ *is solution of the following stochastic Volterra equation:*

$$\begin{aligned}\tilde{\boldsymbol{V}}_t = &\frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\left(\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\boldsymbol{\mu^1}_s + \boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\boldsymbol{\mu^2}_s - \tilde{\boldsymbol{V}}_s\right)ds\\ &+ \frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}_s}\right)d\boldsymbol{W}_s^{\boldsymbol{1}}\\ &+ \frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{\Theta}^2\tilde{\boldsymbol{V}}_s}\right)d\boldsymbol{W}_s^{\boldsymbol{2}},\end{aligned}$$

where $\boldsymbol{W^1} := (B^1, \cdots, B^{n_c})$ *and* $\boldsymbol{W^2} := (B^{n_c+1}, \cdots, B^{2m})$.

Proof. We begin by showing the first implication. Starting from Proposition 3 we have

$$\boldsymbol{V}_t = \int_0^t \boldsymbol{f}(t-s)\boldsymbol{\mu}_s ds + \int_0^t \boldsymbol{f}(t-s)\text{diag}\left(\sqrt{\boldsymbol{V}_s}\right)d\boldsymbol{B}_s.$$

Developing from the definition of $\boldsymbol{f}$ in Equation (16), for any $t \in [0, 1]$, $\boldsymbol{f}$ can be written

$$\boldsymbol{f}(t) = \begin{pmatrix}\big(\boldsymbol{O_{11}} + \boldsymbol{O_{12}}\big(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds\big)^{-1}\int_0^\infty \boldsymbol{B}(s)ds\big)\boldsymbol{K}^{-1}\boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t) & \boldsymbol{0}\\ \big(\boldsymbol{O_{21}} + \boldsymbol{O_{22}}(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds)^{-1}\int_0^\infty \boldsymbol{B}(s)ds\big)\boldsymbol{K}^{-1}\boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t) & \boldsymbol{0}\end{pmatrix}\begin{pmatrix}\boldsymbol{O}_{\boldsymbol{11}}^{(-1)} & \boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\\ \boldsymbol{O}_{\boldsymbol{21}}^{(-1)} & \boldsymbol{O}_{\boldsymbol{22}}^{(-1)}\end{pmatrix}.$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Defining $\boldsymbol{V^1} := (V^1, \cdots, V^{n_c})$ and $\boldsymbol{V^2} := (V^{n_c+1}, \cdots, V^{2m})$, we have

$$\boldsymbol{V^1_t} = \boldsymbol{\Theta^1} \int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\boldsymbol{\mu}_s^{\boldsymbol{1}} ds + \boldsymbol{\Theta^1} \int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\boldsymbol{\mu}_s^{\boldsymbol{2}} ds$$
$$+ \boldsymbol{\Theta^1} \int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{V^1_s}}\right)d\boldsymbol{W^1_s} + \boldsymbol{\Theta^1} \int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{V^2_s}}\right)d\boldsymbol{W^2_s}.$$

If $\boldsymbol{\Theta^1}$ were nonsingular, we could express $\boldsymbol{V^1}$ with power-law kernels using the same approach as in [8]. In general we define

$$\tilde{\boldsymbol{V}}_t := \int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)(\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\boldsymbol{\mu^1}_s + \boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\boldsymbol{\mu^2}_s)ds$$
$$+ \int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{V^1_s}}\right)d\boldsymbol{W^1_s} + \int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{V^2_s}}\right)d\boldsymbol{W^2_s}.$$

From the same arguments as in Lemma 3, Hölder regularity of $\boldsymbol{V}$ carries over to $\tilde{\boldsymbol{V}}$, and the components of $\tilde{\boldsymbol{V}}$ are of Hölder regularity $\alpha - 1/2 - \epsilon$ for any $\epsilon > 0$; hence Lemma 3 shows $\mathcal{K} := I^{1-\alpha}\tilde{\boldsymbol{V}}$ is well-defined, where $I^{1-\alpha}$ is the fractional integration operator of order $1-\alpha$ (see Definition 1 in Appendix A.2). Note that for any t in [0, 1], using Lemma 4 of Appendix A.2, we have

$$\begin{aligned}
\mathcal{K}_t =& \int_0^t \boldsymbol{\Lambda}(\boldsymbol{I} - \boldsymbol{F}^{\alpha,\boldsymbol{\Lambda}}(t-s))(\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\boldsymbol{\mu^1}_s + \boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\boldsymbol{\mu^2}_s)ds \\
&+ \int_0^t \boldsymbol{\Lambda}(\boldsymbol{I} - \boldsymbol{F}^{\alpha,\boldsymbol{\Lambda}}(t-s))\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{V^1_s}}\right)d\boldsymbol{W^1_s} \\
&+ \int_0^t \boldsymbol{\Lambda}(\boldsymbol{I} - \boldsymbol{F}^{\alpha,\boldsymbol{\Lambda}}(t-s))\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{V^2_s}}\right)d\boldsymbol{W^2_s} \\
=& \boldsymbol{\Lambda} \int_0^t (\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\boldsymbol{\mu^1}_s + \boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\boldsymbol{\mu^2}_s)ds + \int_0^t \boldsymbol{\Lambda}\boldsymbol{O}_{\boldsymbol{11}}\text{diag}\left(\sqrt{\boldsymbol{V^1_s}}\right)d\boldsymbol{W^1_s} \\
&+ \int_0^t \boldsymbol{\Lambda}\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{V^2_s}}\right)d\boldsymbol{W^2_s} \\
&- \boldsymbol{\Lambda} \int_0^t \left[\boldsymbol{F}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\boldsymbol{\mu^1}_s + \int_0^s \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(s-u)\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{V^1_u}}\right)d\boldsymbol{W^1_u}\right] ds \\
&- \boldsymbol{\Lambda} \int_0^t \left[\boldsymbol{F}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\boldsymbol{\mu^2}_s + \int_0^s \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(s-u)\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{V^2_u}}\right)d\boldsymbol{W^2_u}\right] ds.
\end{aligned}$$

The last two terms can be rewritten using the definition of $\tilde{\boldsymbol{V}}$, so that

$$\mathcal{K}_t = \boldsymbol{\Lambda} \int_0^t \left(\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\boldsymbol{\mu^1}_s + \boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\boldsymbol{\mu^2}_s - \tilde{\boldsymbol{V}}_s\right)ds + \boldsymbol{\Lambda} \int_0^t \boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}_s}\right)d\boldsymbol{W^1_s}$$
$$+ \boldsymbol{\Lambda} \int_0^t \boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\text{diag}\left(\sqrt{\boldsymbol{\Theta^2}\tilde{\boldsymbol{V}}_s}\right)d\boldsymbol{W^2_s}.$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Thanks to the Hölder regularity of $\tilde{\boldsymbol{V}}$, we can now apply the fractional differentiation operator of order $1-\alpha$ (see Definition 1 in Appendix A.2) together with the stochastic Fubini theorem to deduce that

$$\begin{aligned}\tilde{\boldsymbol{V}}_t = &\frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\Big(\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\boldsymbol{\mu^1}_s + \boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\boldsymbol{\mu^2}_s - \tilde{\boldsymbol{V}}_s\Big)ds \\ &+ \frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}_s}\Big)d\boldsymbol{W}_s^{\boldsymbol{1}} \\ &+ \frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta^2}\tilde{\boldsymbol{V}}_s}\Big)d\boldsymbol{W}_s^{\boldsymbol{2}}.\end{aligned}$$

This concludes the proof of the first implication. We now show the second implication. Suppose there exists $\tilde{\boldsymbol{V}}$ of Hölder regularity $\alpha - 1/2 - \epsilon$ for any $\epsilon > 0$ such that $\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}$ and $\boldsymbol{\Theta^2}\tilde{\boldsymbol{V}}$ are positive, solving the following stochastic Volterra equation:

$$\begin{aligned}\tilde{\boldsymbol{V}}_t = &\frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\Big(\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\boldsymbol{\mu^1}_s + \boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\boldsymbol{\mu^2}_s - \tilde{\boldsymbol{V}}_s\Big)ds \\ &+ \frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}_s}\Big)d\boldsymbol{W}_s^{\boldsymbol{1}} \\ &+ \frac{1}{\Gamma(\alpha)}\boldsymbol{\Lambda}\int_0^t (t-s)^{\alpha-1}\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta^2}\tilde{\boldsymbol{V}}_s}\Big)d\boldsymbol{W}_s^{\boldsymbol{2}}.\end{aligned}$$

For this proof, let us write

$$\boldsymbol{\theta} := \boldsymbol{\Lambda}\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}\boldsymbol{\mu^1} + \boldsymbol{\Lambda}\boldsymbol{O}_{\boldsymbol{12}}^{(-1)}\boldsymbol{\mu^2}, \qquad \boldsymbol{\Lambda}_1 := \boldsymbol{\Lambda}\boldsymbol{O}_{\boldsymbol{11}}^{(-1)}, \qquad \boldsymbol{\Lambda}_2 := \boldsymbol{\Lambda}\boldsymbol{O}_{\boldsymbol{12}}^{(-1)},$$

so that, for any t in [0,1],

$$\begin{aligned}\tilde{\boldsymbol{V}}_t = &\frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\big(\boldsymbol{\theta}_s - \boldsymbol{\Lambda}\tilde{\boldsymbol{V}}_s\big)ds \\ &+ \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\boldsymbol{\Lambda_1}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}_s}\Big)d\boldsymbol{W}_s^{\boldsymbol{1}} \\ &+ \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\boldsymbol{\Lambda_2}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta^2}\tilde{\boldsymbol{V}}_s}\Big)d\boldsymbol{W}_s^{\boldsymbol{2}}.\end{aligned}$$

Notice that the above can be written

$$\tilde{\boldsymbol{V}}_t = I^\alpha(\boldsymbol{\theta} - \boldsymbol{\Lambda}\tilde{\boldsymbol{V}})_t + I^\alpha_{\boldsymbol{B}^1}\Big(\boldsymbol{\Lambda_1}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}}\Big)\Big)_t + I^\alpha_{\boldsymbol{B}^2}(\boldsymbol{\Lambda_2}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta^2}\tilde{\boldsymbol{V}}}\Big)\Big)_t,$$

where $I^\alpha_{\boldsymbol{B}}$ is the fractional integration operator with respect to $\boldsymbol{B}$ (see Definition 2 in Appendix A.2). Iterating the application of I^α we find that, for any $N \geq 1$, $\tilde{\boldsymbol{V}}$ satisfies

$$\begin{aligned}\tilde{\boldsymbol{V}} = &\sum_{1\le k\le N}\boldsymbol{\Lambda}^{k-1}(-1)^{k-1}I^{(k-1)\alpha}\Big[I^\alpha\boldsymbol{\theta} + I^\alpha_{\boldsymbol{B}^1}\Big(\boldsymbol{\Lambda_1}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}}\Big)\Big) + I^\alpha_{\boldsymbol{B}^2}\Big(\boldsymbol{\Lambda_2}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta^2}\tilde{\boldsymbol{V}}}\big)\Big)\Big] \\ &+ \boldsymbol{\Lambda}^N(-1)^N I^{(N+1)\alpha}\tilde{\boldsymbol{V}}.\end{aligned}$$

Now, note that $\boldsymbol{\theta}$, $\mathrm{diag}(\sqrt{\boldsymbol{\Theta}^{\mathbf{1}}\tilde{\boldsymbol{V}}})$, $\mathrm{diag}(\sqrt{\boldsymbol{\Theta}^{\mathbf{2}}\tilde{\boldsymbol{V}}})$, and $\tilde{\boldsymbol{V}}$ are square-integrable processes, and Lemma 8 in Appendix A.2 shows that the sum converges almost surely to the series, while $\boldsymbol{\Lambda}^N(-1)^N I^{(N+1)\alpha}\tilde{\boldsymbol{V}}$ converges almost surely to zero, as N tends to infinity. Thus we have

$$\begin{aligned}
\tilde{\boldsymbol{V}} &= \sum_{k\geq 0} \boldsymbol{\Lambda}^k(-1)^k I^{k\alpha}\Big[I^\alpha\boldsymbol{\theta} + I^\alpha_{\boldsymbol{B}^1}\Big(\boldsymbol{\Lambda}_{\mathbf{1}}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta}^{\mathbf{1}}\tilde{\boldsymbol{V}}}\Big)\Big) + I^\alpha_{\boldsymbol{B}^2}\Big(\boldsymbol{\Lambda}_{\mathbf{2}}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta}^{\mathbf{2}}\tilde{\boldsymbol{V}}}\Big)\Big)\Big] \\
&= \sum_{k\geq 0} \boldsymbol{\Lambda}^k(-1)^k I^{k\alpha} I^\alpha\boldsymbol{\theta} + \sum_{k\geq 0} \boldsymbol{\Lambda}^k(-1)^k I^{k\alpha} I^\alpha_{\boldsymbol{B}^1}\Big(\boldsymbol{\Lambda}_{\mathbf{1}}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta}^{\mathbf{1}}\tilde{\boldsymbol{V}}}\Big)\Big) \\
&\quad + I^\alpha_{\boldsymbol{B}^2}\Big(\boldsymbol{\Lambda}_{\mathbf{2}}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta}^{\mathbf{2}}\tilde{\boldsymbol{V}}}\Big)\Big)\Big] \\
&= \boldsymbol{\Lambda}^{-1}\sum_{k\geq 0} \boldsymbol{\Lambda}^{k+1}(-1)^k I^{(k+1)\alpha}\boldsymbol{\theta} + \sum_{k\geq 0} \boldsymbol{\Lambda}^k(-1)^k I^{k\alpha} I^\alpha_{\boldsymbol{B}^1}\Big(\boldsymbol{\Lambda}_{\mathbf{1}}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta}^{\mathbf{1}}\tilde{\boldsymbol{V}}}\Big)\Big) \\
&\quad + I^\alpha_{\boldsymbol{B}^2}\Big(\boldsymbol{\Lambda}_{\mathbf{2}}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta}^{\mathbf{2}}\tilde{\boldsymbol{V}}}\Big)\Big)\Big].
\end{aligned}$$

Lemmas 5 and 7 in Appendix A.2 enable us to rewrite the above using the matrix Mittag-Leffler function. This yields, for any t in [0,1] and almost surely,

$$\begin{aligned}
\tilde{\boldsymbol{V}}_t &= \boldsymbol{\Lambda}^{-1}\int_0^t \boldsymbol{f}^{\alpha,\Lambda}(t-s)\boldsymbol{\theta}_s ds + \boldsymbol{\Lambda}^{-1}\int_0^t \boldsymbol{f}^{\alpha,\Lambda}(t-s)\boldsymbol{\Lambda}_{\mathbf{1}}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta}^{\mathbf{1}}\tilde{\boldsymbol{V}}_s}\Big)d\boldsymbol{W}^{\boldsymbol{1}}_s \\
&\quad + \boldsymbol{\Lambda}^{-1}\int_0^t \boldsymbol{f}^{\alpha,\Lambda}(t-s)\boldsymbol{\Lambda}_{\mathbf{2}}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta}^{\mathbf{2}}\tilde{\boldsymbol{V}}_s}\Big)d\boldsymbol{W}^{\boldsymbol{2}}_s.
\end{aligned}$$

Replacing $\boldsymbol{\theta}$, $\boldsymbol{\Lambda}_1$, $\boldsymbol{\Lambda}_2$ by their expressions, almost surely and for any t in [0,1], we have

$$\begin{aligned}
\tilde{\boldsymbol{V}}_t = &\int_0^t \boldsymbol{f}^{\alpha,\Lambda}(t-s)(\boldsymbol{O}^{(-1)}_{\boldsymbol{11}}\boldsymbol{\mu}^{\mathbf{1}}{}_s + \boldsymbol{\Lambda}\boldsymbol{O}^{(-1)}_{\boldsymbol{12}}\boldsymbol{\mu}^{\mathbf{2}}{}_s)ds \\
&+ \int_0^t \boldsymbol{f}^{\alpha,\Lambda}(t-s)\boldsymbol{O}^{(-1)}_{\boldsymbol{11}}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta}^{\mathbf{1}}\tilde{\boldsymbol{V}}_s}\Big)d\boldsymbol{W}^{\boldsymbol{1}}_s \\
&+ \int_0^t \boldsymbol{f}^{\alpha,\Lambda}(t-s)\boldsymbol{O}^{(-1)}_{\boldsymbol{12}}\mathrm{diag}\Big(\sqrt{\boldsymbol{\Theta}^{\mathbf{2}}\tilde{\boldsymbol{V}}_s}\Big)d\boldsymbol{W}^{\boldsymbol{2}}_s.
\end{aligned}$$

This concludes the second implication and the proof. □

5.4. Step 4: Equation satisfied by the limiting price process

The previous results on the convergence of the intensity process enable us to now turn to the question of the limiting price dynamics. Recall that the sequence of rescaled price processes $\boldsymbol{P}^T$ is defined as

$$\boldsymbol{P}^T := {}^\top\boldsymbol{Q}\boldsymbol{X}^T,$$

where $\boldsymbol{Q} = \big(\boldsymbol{e}_1 - \boldsymbol{e}_2 \,|\cdots|\, \boldsymbol{e}_{2m-1} - \boldsymbol{e}_{2m}\big)$. We have the following result.

Proposition 5. *Let* $(\boldsymbol{X}, \boldsymbol{Z})$ *be a limit point of* $(\boldsymbol{X}^T, \boldsymbol{Z}^T)$ *and* $\boldsymbol{P} = {}^\top\boldsymbol{Q}\boldsymbol{X}$. *Then*

$$\boldsymbol{P}_t = (\boldsymbol{I} + \boldsymbol{\Delta})^\top\boldsymbol{Q}\Big(\boldsymbol{Z}_t + \int_0^t \boldsymbol{\mu}_s ds\Big),$$

where $\boldsymbol{\Delta} = \Big(\int_0^\infty \delta^T_{ij}\Big)_{1\leq i,j\leq m}$.

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Proof. Let $(\boldsymbol{X}, \boldsymbol{Z})$ be a limit point of $(\boldsymbol{X}^T, \boldsymbol{Z}^T)$. For any $1 \leq i \leq m$ we can compute the difference between upward and downard jumps on Asset i as

$$\boldsymbol{v}_i \cdot \boldsymbol{N}_t^T = \boldsymbol{v}_i \cdot \boldsymbol{M}_t^T + \boldsymbol{v}_i \cdot \int_0^t \boldsymbol{\lambda}_s^T ds,$$

with the following expression for the integrated intensity:

$$\int_0^{tT} \boldsymbol{\lambda}_s^T ds = T \int_0^t \boldsymbol{\mu}_{sT}^T ds + T \int_0^t \int_0^{T(t-s)} \boldsymbol{\psi}^T(u) du \boldsymbol{\mu}_{sT}^T ds + \left\| \psi^{\boldsymbol{T}} \right\|_1 \boldsymbol{M}_{\boldsymbol{tT}}^{\boldsymbol{T}} - \int_0^{tT} \int_{tT-s}^{\infty} \boldsymbol{\psi}^T(u) du d\boldsymbol{M}_s^T.$$

Thus the microscopic price for Asset i satisfies

$$\begin{aligned}
T^{-\alpha} \boldsymbol{v}_i \cdot \boldsymbol{N}_{tT}^T &= T^{1-\alpha} \int_0^t \boldsymbol{v}_i \cdot \boldsymbol{\mu}_{sT}^T ds + T^{1-\alpha\,\top} \left\| \boldsymbol{\psi}^T \right\|_1 \boldsymbol{v}_i \cdot \int_0^t \boldsymbol{\mu}_{sT}^T ds + \boldsymbol{v}_i \cdot \boldsymbol{Z}_{\boldsymbol{t}}^{\boldsymbol{T}} + {}^\top \left\| \psi^{\boldsymbol{T}} \right\|_1 \boldsymbol{v}_i \cdot \boldsymbol{Z}_{\boldsymbol{t}}^{\boldsymbol{T}} \\
&\quad - T^{-\alpha} \int_0^t \int_{T(t-s)}^{\infty} {}^\top \boldsymbol{\psi}^T(u) \boldsymbol{v}_i \cdot \boldsymbol{\mu}_{sT}^T du ds - T^{-\alpha} \int_0^{tT} \int_{tT-s}^{\infty} \boldsymbol{\psi}^T(u) du d\boldsymbol{M}_s^T \\
&= \sum_{1 \leq k \leq m} \left(\mathbb{1}_{ik} + \int_0^{\infty} \delta_{ik}^T \right) \boldsymbol{v}_k \cdot \boldsymbol{Z}_{\boldsymbol{t}}^{\boldsymbol{T}} + \sum_{1 \leq k \leq m} \left(\mathbb{1}_{ik} + \int_0^{\infty} \delta_{ik}^T \right) T^{1-\alpha} \int_0^t \boldsymbol{v}_k \cdot \boldsymbol{\mu}_{sT}^T ds \\
&\quad - \int_0^t \int_{tT-s}^{\infty} {}^\top \boldsymbol{\psi}^T(u) \boldsymbol{v}_i du \cdot d\boldsymbol{Z}_{\boldsymbol{s}}^{\boldsymbol{T}} - T^{-\alpha} \int_0^t \int_{T(t-s)}^{\infty} {}^\top \boldsymbol{\psi}^T(u) \boldsymbol{v}_i \cdot \boldsymbol{\mu}_{sT}^T du ds.
\end{aligned}$$

It is straightforward to show that the last two terms converge to zero, and thus any limit point $\boldsymbol{P}$ of $\boldsymbol{P}^T = {}^\top \boldsymbol{Q} \boldsymbol{X}^T$ is such that

$$\boldsymbol{P}_t = (\boldsymbol{I} + \boldsymbol{\Delta})^\top \boldsymbol{Q} \left(\boldsymbol{Z}_t + \int_0^t \boldsymbol{\mu}_s ds \right). \qquad \square$$

Replacing $\boldsymbol{Z}$ by the expression obtained in Proposition 3 concludes the proof of Theorem 1, since

$$\boldsymbol{P}_t = (\boldsymbol{I} + \boldsymbol{\Delta})^\top \boldsymbol{Q} \left(\int_0^t \mathrm{diag}\left(\sqrt{\boldsymbol{V}_s} \right) d\boldsymbol{B}_s + \int_0^t \boldsymbol{\mu}_s ds \right).$$

Appendix A. Technical appendix

A.1 Independence of Equation (11) from chosen basis

We consider two representations which satisfy Assumption 1. Let $\boldsymbol{P}, \tilde{\boldsymbol{P}}$ be invertible matrices, $0 \leq n_c, n_{c'} \leq n$, and let

$$\boldsymbol{A}^T \in \mathcal{F}(\mathcal{M}_{n_c}(\mathbb{R})), \qquad \boldsymbol{C}^T \in \mathcal{F}(\mathcal{M}_{n-n_c}(\mathbb{R})), \qquad \boldsymbol{B}^T \in \mathcal{F}(\mathcal{M}_{n-n_c,n_c}(\mathbb{R})),$$
$$\tilde{\boldsymbol{A}}^T \in \mathcal{F}(\mathcal{M}_{n_{c'}}(\mathbb{R})), \qquad \tilde{\boldsymbol{C}}^T \in \mathcal{F}(\mathcal{M}_{n-n_{c'}}(\mathbb{R})), \qquad \tilde{\boldsymbol{B}}^T \in \mathcal{F}(\mathcal{M}_{n-n_{c'},n_{c'}}(\mathbb{R}))$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

be such that

$$\boldsymbol{\phi}^T = \boldsymbol{P}\begin{pmatrix} \boldsymbol{A}^T & \boldsymbol{0} \\ \boldsymbol{B}^T & \boldsymbol{C}^T \end{pmatrix}\boldsymbol{P}^{-1} = \tilde{\boldsymbol{P}}\begin{pmatrix} \tilde{\boldsymbol{A}}^T & \boldsymbol{0} \\ \tilde{\boldsymbol{B}}^T & \tilde{\boldsymbol{C}}^T \end{pmatrix}\tilde{\boldsymbol{P}}^{-1}.$$

We write $\boldsymbol{A}$ for the limit of $\boldsymbol{A}^T$ (and similarly for $\boldsymbol{B}^T$, $\boldsymbol{C}^T$, etc.). First, notice that we must have $n_c = n_{c'}$. Indeed, since $\rho(\int_0^\infty \boldsymbol{C}) < 1$ and $\rho(\int_0^\infty \tilde{\boldsymbol{C}}) < 1$, 1 is neither an eigenvalue of $\int_0^\infty \boldsymbol{C}$ nor of $\int_0^\infty \tilde{\boldsymbol{C}}$. Yet, since $\boldsymbol{A} = \boldsymbol{I}$ and $\tilde{\boldsymbol{A}} = \boldsymbol{I}$, 1 is an eigenvalue of $\boldsymbol{\phi}$ with multiplicity n_c and $n_{c'}$. Therefore $n_c = n_{c'}$.

Writing $\boldsymbol{L} = \boldsymbol{P}^{-1}\tilde{\boldsymbol{P}}$, we have

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{B} & \boldsymbol{C} \end{pmatrix} = \boldsymbol{L}\begin{pmatrix} \tilde{\boldsymbol{A}} & \boldsymbol{0} \\ \tilde{\boldsymbol{B}} & \tilde{\boldsymbol{C}} \end{pmatrix}\boldsymbol{L}^{-1}.$$

Since $\boldsymbol{A} = \tilde{\boldsymbol{A}} = \boldsymbol{I}$ because of Equation (5), writing the blockwise matrix product and using the assumption that $\boldsymbol{I} - \boldsymbol{C}$ is invertible, we get

$$\begin{aligned} \boldsymbol{L}_{11} &= \boldsymbol{0}, \\ (\boldsymbol{I} - \boldsymbol{C})\boldsymbol{L}_{21} &= \boldsymbol{B}\boldsymbol{L}_{11} - \boldsymbol{L}_{22}\tilde{\boldsymbol{B}}, \\ \boldsymbol{C}\boldsymbol{L}_{22} &= \boldsymbol{L}_{22}\tilde{\boldsymbol{C}}. \end{aligned}$$

Since $\boldsymbol{L}\boldsymbol{L}^{-1} = \boldsymbol{I}$, $\boldsymbol{L}_{11} = \boldsymbol{I}$, $\boldsymbol{L}_{22} = \boldsymbol{I}$, and $\boldsymbol{L}_{21} = -\boldsymbol{L}_{21}^{(-1)}$, we deduce that

$$\boldsymbol{L}_{11} = \boldsymbol{I}, \qquad \boldsymbol{L}_{22} = \boldsymbol{I}, \qquad \boldsymbol{L}_{12} = \boldsymbol{0}, \qquad (\boldsymbol{I} - \boldsymbol{C})\boldsymbol{L}_{21} = \boldsymbol{B} - \tilde{\boldsymbol{B}}, \qquad \boldsymbol{C} = \tilde{\boldsymbol{C}}.$$

As $\boldsymbol{L} = \boldsymbol{P}^{-1}\tilde{\boldsymbol{P}}$, we have

$$\begin{aligned} \boldsymbol{P}^{-1} &= \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ (\boldsymbol{I} - \boldsymbol{C})^{-1}(\boldsymbol{B} - \tilde{\boldsymbol{B}}) & \boldsymbol{I} \end{pmatrix}\tilde{\boldsymbol{P}}^{-1} \\ &= \begin{pmatrix} \tilde{\boldsymbol{P}}_{11}^{(-1)} & \tilde{\boldsymbol{P}}_{12}^{(-1)} \\ (\boldsymbol{I} - \boldsymbol{C})^{-1}(\boldsymbol{B} - \tilde{\boldsymbol{B}})\tilde{\boldsymbol{P}}_{11}^{(-1)} + \tilde{\boldsymbol{P}}_{21}^{(-1)} & (\boldsymbol{I} - \boldsymbol{C})^{-1}(\boldsymbol{B} - \tilde{\boldsymbol{B}})\tilde{\boldsymbol{P}}_{12}^{(-1)} + \tilde{\boldsymbol{P}}_{22}^{(-1)} \end{pmatrix}. \end{aligned}$$

Computing the matrix product $\tilde{\boldsymbol{P}} = \boldsymbol{P}\boldsymbol{L}$ blockwise and using the above, we find

$$\begin{aligned} &\tilde{\boldsymbol{P}}_{11}^{(-1)} = \boldsymbol{P}_{11}^{(-1)}, \qquad \tilde{\boldsymbol{P}}_{12}^{(-1)} = \boldsymbol{P}_{12}^{(-1)}, \qquad \tilde{\boldsymbol{P}}_{12} = \boldsymbol{P}_{12}, \qquad \tilde{\boldsymbol{P}}_{22} = \boldsymbol{P}_{22}, \\ &\tilde{\boldsymbol{P}}_{11} = \boldsymbol{P}_{11} + \boldsymbol{P}_{12}(\boldsymbol{I} - \boldsymbol{C})^{-1}(\boldsymbol{B} - \tilde{\boldsymbol{B}}), \\ &\tilde{\boldsymbol{P}}_{21} = \boldsymbol{P}_{21} + \boldsymbol{P}_{22}(\boldsymbol{I} - \boldsymbol{C})^{-1}(\boldsymbol{B} - \tilde{\boldsymbol{B}}). \end{aligned}$$

Thus

$$\begin{aligned} &\tilde{\boldsymbol{P}}_{11}^{(-1)} = \boldsymbol{P}_{11}^{(-1)}, \qquad \tilde{\boldsymbol{P}}_{12}^{(-1)} = \boldsymbol{P}_{12}^{(-1)}, \\ &\tilde{\boldsymbol{P}}_{11} + \tilde{\boldsymbol{P}}_{12}(\boldsymbol{I} - \boldsymbol{C})^{-1}\tilde{\boldsymbol{B}} = \boldsymbol{P}_{11} + \boldsymbol{P}_{12}(\boldsymbol{I} - \boldsymbol{C})^{-1}\boldsymbol{B}, \\ &\tilde{\boldsymbol{P}}_{21} + \tilde{\boldsymbol{P}}_{22}(\boldsymbol{I} - \boldsymbol{C})^{-1}\tilde{\boldsymbol{B}} = \boldsymbol{P}_{21} + \boldsymbol{P}_{22}(\boldsymbol{I} - \boldsymbol{C})^{-1}\boldsymbol{B}. \end{aligned}$$

Therefore, regardless of the chosen basis, Equation (11) is the same, which concludes the proof.

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

A.2 Fractional operators

This section is a brief reminder about fractional operators, which are used in the proofs. We also introduce the matrix-extended Mittag-Leffler function.

Definition 1. *(Fractional differentiation and integration operators.)* For $\alpha \in (0, 1)$, the fractional differentiation operator, denoted by D^{α} is defined as

$$D^{\alpha} f(t) := \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t (t-s)^{-\alpha} f(s) ds,$$

where f is a measurable, Hölder continuous function of order strictly greater than α. The fractional integration operator, denoted by I^{α}, is defined as

$$I^{\alpha} f(t) := \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s) ds,$$

where f is a measurable function.

It will be convenient for us to define fractional integration with respect to a Brownian motion.

Definition 2. *(Fractional integration operator with respect to a Brownian motion.)* Given a Brownian motion B and $\alpha \in (1/2, 1)$, the fractional integration operator with respect to B, denoted by I_B^{α}, is defined as

$$I_B^{\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{1-\alpha} f(s) dB_s,$$

for f a measurable, square-integrable stochastic process.

Remark 2. The fractional integration of a matrix-valued stochastic process $\boldsymbol{f}$ with respect to a multivariate Brownian motion $\boldsymbol{B}$ is

$$I_{\boldsymbol{B}}^{\alpha} \boldsymbol{f}(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{1-\alpha} \boldsymbol{f}(s) d\boldsymbol{B}_s.$$

We now extend the Mittag-Leffler function to matrices (for a theory of matrix-valued functions, see for example [16]). We have the following definition.

Definition 3. *(Matrix-extended Mittag-Leffler function.)* Let $\alpha, \beta \in \mathbb{C}$ such that $\mathrm{Re}(\alpha)$, $\mathrm{Re}(\beta) > 0$, and let $\boldsymbol{\Lambda} \in \mathcal{M}_n(\mathbb{R})$. Then the matrix Mittag-Leffler function is defined as

$$\boldsymbol{E}_{\alpha,\beta}(\boldsymbol{\Lambda}) := \sum_{n \geq 0} \frac{\boldsymbol{\Lambda}^n}{\Gamma(\alpha n + \beta)}.$$

We also extend the Mittag-Leffler density function for matrices.

Definition 4. *(Mittag-Leffler density for matrices.)* Let $\alpha \in \mathbb{C}$ such that $\mathrm{Re}(\alpha) > 0$, $\boldsymbol{\Lambda} \in \mathcal{M}_n(\mathbb{R})$. Then the matrix Mittag-Leffler density function $\boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}$ is defined as

$$\boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t) := \boldsymbol{\Lambda} t^{\alpha-1} \boldsymbol{E}_{\alpha,\alpha}(-\boldsymbol{\Lambda} t^{\alpha}).$$

We write $\boldsymbol{F}^{\alpha,\Lambda}$ for the cumulative matrix Mittag-Leffler density function,

$$\boldsymbol{F}^{\alpha,\Lambda}(t) := \int_0^t \boldsymbol{f}^{\alpha,\Lambda}(s) ds.$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Using Definition 3, it is easy to prove the following lemma.

Lemma 4. *Let* $\alpha \in \mathbb{C}$ *such that* $\mathrm{Re}(\alpha) > 0$, *and let* $\boldsymbol{\Lambda} \in \mathcal{M}_n(\mathbb{R})$. *Then*

$$I^{1-\alpha} \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}} = \boldsymbol{\Lambda}(\boldsymbol{I} - \boldsymbol{F}^{\alpha,\boldsymbol{\Lambda}}).$$

Furthermore, if $\alpha \in (1/2, 1)$, *then*

$$\widehat{\boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}}(z) = \boldsymbol{\Lambda}(\boldsymbol{I} z^{\alpha} + \boldsymbol{\Lambda})^{-1}.$$

We need another important property relating Mittag-Leffler functions to fractional integration operators.

Lemma 5. *Let* $\alpha > 0$ *and* $\boldsymbol{\Lambda} \in \mathcal{M}_m(\mathbb{R})$. *Then*

$$I^1 \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}} = \sum_{n \geq 1} (-1)^{n-1} \boldsymbol{\Lambda}^n I^{n\alpha}(1).$$

Proof. Using Lemma 4 and repeated applications of I^{α}, for all $N \geq 1$ we have

$$I\boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}} = \sum_{1 \leq n \leq N} (-1)^{n-1} \boldsymbol{\Lambda}^n I^{n\alpha}(1) + (-1)^{N-1} \boldsymbol{\Lambda}^N I^{N\alpha} I \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}.$$

Therefore, if we show that

$$(-1)^{N-1} \boldsymbol{\Lambda}^N I^{N\alpha} I \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}} \underset{N \to \infty}{\to} 0,$$

the result will follow. To prove this we make use of the series expansion of $I^{N\alpha}\boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}$ to deduce bounds which will converge to zero. Writing C for a constant independent of t and N which may change from line to line, $N_\alpha = \lfloor \frac{1}{\alpha} \rfloor$, and $\|\cdot\|_{\mathrm{op}}$ for the operator norm, we have

$$\begin{aligned} \left\| \boldsymbol{\Lambda}^N \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t) \right\|_{\mathrm{op}} &= \left\| \boldsymbol{\Lambda}^{N+1} \sum_{n \geq 0} (-1)^n \frac{t^{(n+1)\alpha - 1}}{\Gamma((n+1)\alpha)} \right\|_{\mathrm{op}} \\ &\leq \left\| \boldsymbol{\Lambda}^{N+1} \sum_{0 \leq n \leq N_\alpha} (-1)^n \frac{t^{(n+1)\alpha - 1}}{\Gamma((n+1)\alpha)} + \boldsymbol{\Lambda}^{N+1} C \right\|_{\mathrm{op}}. \end{aligned}$$

Therefore, applying the fractional integration operator of order $N\alpha$, and writing $g_n : t \mapsto t^{(n+1)\alpha - 1}$, we have

$$\begin{aligned} I^{N\alpha} \left\| \boldsymbol{\Lambda}^N \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t) \right\|_{\mathrm{op}} &\leq \left\| \boldsymbol{\Lambda}^{N+1} I^{N\alpha} \left(\sum_{0 \leq n \leq N_\alpha} (-1)^n \frac{g_n}{\Gamma((n+1)\alpha)} \right) + \boldsymbol{\Lambda}^{N+1} I^{N\alpha}(C) \right\|_{\mathrm{op}} \\ &\leq \sum_{0 \leq n \leq N_\alpha} \frac{1}{\Gamma((n+1)\alpha)} \left\| \boldsymbol{\Lambda}^{N+1} I^{N\alpha}(g_n) \right\|_{\mathrm{op}} + \left\| \boldsymbol{\Lambda}^{N+1} I^{N\alpha}(C) \right\|_{\mathrm{op}}. \end{aligned}$$

An explicit computation of $I^{N\alpha}(g_n)$ shows the convergence to zero of the right-hand side as N tends to infinity, which concludes the proof. □

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Finally, we need to combine fractional integration I^{α} with I^{α}_{B}. We have the following lemma.

Lemma 6. *Let* $m \geq 1$, $\boldsymbol{B}$ *an* m*-dimensional Brownian motion,* $\boldsymbol{X}$ *an* $m \times m$ *matrix-valued adapted square-integrable stochastic process, and* $\alpha, \beta > 0$. *Then we have*

$$I^{\alpha} I^{\beta}_{\boldsymbol{B}}(\boldsymbol{X}) = I^{\alpha+\beta}_{\boldsymbol{B}}(\boldsymbol{X}).$$

Proof. The proof is a straightforward application of the definitions of the operators together with the stochastic Fubini theorem. □

The next lemma is useful for transforming stochastic convolutions of stochastic processes with the Mittag-Leffler density function into series of repeated applications of $I^{\alpha}_{\boldsymbol{B}}$.

Lemma 7. *Let* $m \geq 1$, $\boldsymbol{B}$ *an* m*-dimensional Brownian motion,* $\boldsymbol{X}$ *an* $m \times m$ *matrix-valued adapted and square-integrable stochastic process,* $\alpha > 0$, *and* $\boldsymbol{\Lambda} \in \mathcal{M}_m(\mathbb{R})$. *Then, for all* $t \geq 0$ *and almost surely,*

$$\int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{X}_s d\boldsymbol{B}_s = \sum_{n\geq 1} (-1)^{n-1} \boldsymbol{\Lambda}^n I^{n\alpha}_{\boldsymbol{B}}(\boldsymbol{X}),$$

where the series converges almost surely.

Proof. Using Lemma 5, we can write the integral using a series of fractional integration operators and apply the stochastic Fubini theorem (as $\boldsymbol{X}$ is square-integrable) to obtain

$$\begin{aligned}
\int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{X}_s d\boldsymbol{B}_s &= \int_0^t \sum_{n\geq 1} (-1)^{n-1} \boldsymbol{\Lambda}^n I^{n\alpha-1}(1)_{t-s} \boldsymbol{X}_s d\boldsymbol{B}_s \\
&= \sum_{n\geq 1} \int_0^t (-1)^{n-1} \boldsymbol{\Lambda}^n I^{n\alpha-1}(1)_{t-s} \boldsymbol{X}_s d\boldsymbol{B}_s \\
&= \sum_{n\geq 1} (-1)^{n-1} \boldsymbol{\Lambda}^n \int_0^t I^{n\alpha-1}(1)_{t-s} \boldsymbol{X}_s d\boldsymbol{B}_s \\
&= \sum_{n\geq 1} \frac{(-1)^{n-1}}{\Gamma(n\alpha-1)} \boldsymbol{\Lambda}^n \int_0^t \int_0^{t-s} (t-s-\tau)^{n\alpha-2} d\tau \boldsymbol{X}_s d\boldsymbol{B}_s.
\end{aligned}$$

After a change of variables and using the stochastic Fubini theorem (see for example [27]), we deduce the simpler expression

$$\int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{X}_s d\boldsymbol{B}_s = \sum_{n\geq 1} \frac{(-1)^{n-1}}{\Gamma(n\alpha-1)} \boldsymbol{\Lambda}^n \int_0^t (t-\tau)^{n\alpha-2} \int_0^{\tau} \boldsymbol{X}_s d\boldsymbol{B}_s d\tau.$$

Integrating by parts, we finally obtain the result:

$$\begin{aligned}
\int_0^t \boldsymbol{f}^{\alpha,\boldsymbol{\Lambda}}(t-s)\boldsymbol{X}_s d\boldsymbol{B}_s &= \sum_{n\geq 1} \frac{(-1)^{n-1}}{\Gamma(n\alpha-1)(n\alpha-1)} \boldsymbol{\Lambda}^n \int_0^t (t-\tau)^{n\alpha-1} \boldsymbol{X}_\tau d\boldsymbol{B}_\tau, \\
&= \sum_{n\geq 1} \frac{(-1)^{n-1}}{\Gamma(n\alpha)} \boldsymbol{\Lambda}^n \int_0^t (t-\tau)^{n\alpha-1} \boldsymbol{X}_\tau d\boldsymbol{B}_\tau, \\
&= \sum_{n\geq 1} (-1)^{n-1} \boldsymbol{\Lambda}^n I^{n\alpha}_{\boldsymbol{B}}(\boldsymbol{X}).
\end{aligned}$$

□

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

The last lemma gives convergence for terms of a series of repeated iterations of I^α.

Lemma 8. *Let* $\alpha > 0$, $\mathbf{\Lambda} \in \mathcal{M}_m(\mathbb{R})$, $\boldsymbol{B}$ *an m-dimensional Brownian motion, and* $\boldsymbol{X}$ *an m-dimensional vector-valued square-integrable stochastic process. Then, almost surely and for all* $t \in [0, 1]$,

$$(-1)^{N-1}\mathbf{\Lambda}^N I^{N\alpha}(\boldsymbol{X})_t \underset{N\to\infty}{\to} 0,$$

$$\sum_{n\geq N}(-1)^{n-1}\mathbf{\Lambda}^n I^{n\alpha}_{\boldsymbol{B}}(\mathrm{diag}(\boldsymbol{X}))_t \underset{N\to\infty}{\to} 0.$$

Proof. Let $N^* > N_\alpha := \left\lfloor \frac{1}{\alpha} \right\rfloor$. Since $\boldsymbol{X}$ is square-integrable, we have

$$\mathbb{E}\left[\left\| \sum_{N>N_*} \mathbf{\Lambda}^N I_{\boldsymbol{B}}^{(N+1)\alpha}(\mathrm{diag}(\boldsymbol{X}))_t \right\|^2\right] \leq \sum_{N_1,N_2>N_*} \mathbb{E}\left[{}^\top\left(\mathbf{\Lambda}^{N_1} I_{\boldsymbol{B}}^{(N_1+1)\alpha}(\mathrm{diag}(\boldsymbol{X}))_t\right)\left(\mathbf{\Lambda}^{N_2} I_{\boldsymbol{B}}^{(N_2+1)\alpha}(\mathrm{diag}(\boldsymbol{X}))_t\right)\right].$$

Using the Cauchy–Schwartz inequality and writing $\|\cdot\|_{\mathrm{op}}$ for the operator norm associated to the Euclidean norm, we find

$$\begin{aligned}
\mathbb{E}\left[\left\| \sum_{N>N_*} \mathbf{\Lambda}^N I_{\boldsymbol{B}}^{(N+1)\alpha}(\mathrm{diag}(\boldsymbol{X}))_t \right\|^2\right]
&\leq \sum_{N_1,N_2>N_*} \|\mathbf{\Lambda}\|_{\mathrm{op}}^{N_1+N_2} \sum_{1\leq k,l\leq m} \mathbb{E}\left[I_{B^k}^{(N_1+1)\alpha}\left(X^k\right)_t I_{B^l}^{(N_2+1)\alpha}\left(X^l\right)_t\right] \\
&\leq \sum_{N_1,N_2>N_*} \frac{\|\mathbf{\Lambda}\|_{\mathrm{op}}^{N_1+N_2}}{\Gamma((N_1+1)\alpha)\Gamma((N_2+1)\alpha)} \sum_{1\leq i\leq m} \int_0^t (t-s)^{(N_1+N_2)\alpha-2}\mathbb{E}\left[\left(X_s^i\right)^2\right]ds \\
&\leq c \sum_{N_1,N_2>N_*} \frac{\|\mathbf{\Lambda}\|_{\mathrm{op}}^{N_1+N_2}}{\Gamma((N_1+1)\alpha)\Gamma((N_2+1)\alpha)} \\
&\leq c\left(\sum_{N>N_*} \frac{\|\mathbf{\Lambda}\|_{\mathrm{op}}^{N}}{\Gamma((N+1)\alpha)}\right)^2.
\end{aligned}$$

Thus, by comparison of functions (for example by application of Stirling's formula), for all $\epsilon > 0$,

$$\sum_{N>N_\alpha} \mathbb{P}\left(\left\| \sum_{N>N_*} \mathbf{\Lambda}^N I_{\boldsymbol{B}}^{(N+1)\alpha}(\mathrm{diag}(\boldsymbol{X}))_t \right\| > \epsilon\right) \leq \frac{1}{\epsilon^2} \sum_{N_*\geq N_\alpha} \mathbb{E}\left[\left\| \sum_{N>N_*} \mathbf{\Lambda}^N I_{\boldsymbol{B}}^{(N+1)\alpha}(\mathrm{diag}(\boldsymbol{X}))_t \right\|^2\right] < \infty.$$

The Borel–Cantelli lemma yields the almost sure convergence to zero of $\mathbf{\Lambda}^N I_{\boldsymbol{B}}^{(N+1)\alpha}(\mathrm{diag}(\boldsymbol{X}))$ as $N \to \infty$. The same approach yields the almost sure convergence to zero of $(-1)^{N-1}\mathbf{\Lambda}^N I^{N\alpha}(\boldsymbol{X})$ as $N \to \infty$. □

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

A.3 Proof of Corollary 1

We split the proof into two steps. First, we show that the structure of the kernel satisfies the assumptions of Section 2. Then, we compute the equations satisfied by variance and prices.

Checking for the assumptions of Theorem 1. We write

$$\boldsymbol{O}_1 := \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{O}_2 := \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{O}_3 := \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{O}_4 := \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

Then, setting $\boldsymbol{O} := \big(\boldsymbol{O}_1 \mid \boldsymbol{O}_2 \mid \boldsymbol{O}_3 \mid \boldsymbol{O}_4\big)$, we have

$$\boldsymbol{\phi}^T = \boldsymbol{O} \begin{pmatrix} \phi_1^T + \phi_2^T & \phi_{12}^{T,c} + \phi_{12}^{T,a} & 0 & 0 \\ \phi_{21}^b + \phi_{21}^s & \tilde{\phi}_1^T + \tilde{\phi}_2^T & 0 & 0 \\ 0 & 0 & \phi_1^T - \phi_2^T & \phi_{12}^{T,c} - \phi_{12}^{T,a} \\ 0 & 0 & \phi_{21}^b - \phi_{21}^s & \tilde{\phi}_1^T - \tilde{\phi}_2^T \end{pmatrix} \boldsymbol{O}^{-1}.$$

It is straightforward to check that the assumptions are satisfied if

$$\begin{aligned} &0 \le (H_{12}^c + H_{12}^a)(H_{21}^c + H_{21}^a) < 1, \\ &0 \le |\, 1 - (\gamma_1 + \gamma_2) - \sqrt{(H_{12}^c - H_{12}^a)(H_{21}^c - H_{21}^a) + (\gamma_1 - \gamma_2)^2}\, | < 1, \\ &0 \le |\, 1 - (\gamma_1 + \gamma_2) + \sqrt{(H_{12}^c - H_{12}^a)(H_{21}^c - H_{21}^a) + (\gamma_1 - \gamma_2)^2}\, | < 1. \end{aligned}$$

Under those conditions, $\boldsymbol{K} = \boldsymbol{I} - \boldsymbol{H}$ has positive eigenvalues, and therefore $\boldsymbol{K}\boldsymbol{M}^{-1} = \dfrac{1}{\alpha}\boldsymbol{K}$ has positive eigenvalues. Therefore all the assumptions of Theorem 1 are satisfied.

Limiting variance process. Since we can apply Theorem 1, we now compute the relevant quantities. As the blockwise matrix $\boldsymbol{B}$ is equal to zero, writing $H^{12} := H_{12}^a + H_{12}^c$ and $H^{21} := H_{21}^a + H_{21}^c$, we have

$$\boldsymbol{O}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \qquad \boldsymbol{K}^{-1} = \frac{1}{1 - H_{12}H_{21}} \begin{pmatrix} 1 & H_{12} \\ H_{21} & 1 \end{pmatrix},$$

$$\boldsymbol{\Theta}^{\mathbf{1}} = \frac{1}{1 - H_{12}H_{21}} \begin{pmatrix} 1 & H_{12} \\ 1 & H_{12} \end{pmatrix}, \qquad \boldsymbol{\Theta}^{\mathbf{2}} = \frac{1}{1 - H_{12}H_{21}} \begin{pmatrix} H_{21} & 1 \\ H_{21} & 1 \end{pmatrix}.$$

One can check that the equations satisfied by $\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}$ and $\boldsymbol{\Theta^2}\tilde{\boldsymbol{V}}$ are the following, where $\boldsymbol{B}$ is a Brownian motion:

$$\begin{aligned}\boldsymbol{\Theta^1}\tilde{\boldsymbol{V}}_t &= \frac{\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)}\int_0^t (t-s)^{\alpha-1}\left[\begin{pmatrix}\mu_1\\ \mu_1\end{pmatrix} - \begin{pmatrix}\tilde{V}_s^1\\ \tilde{V}_s^1\end{pmatrix}\right] ds \\ &\quad + \frac{\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)}\int_0^t (t-s)^{\alpha-1}\sqrt{\tilde{V}_s^1 + H_{12}\tilde{V}_s^2}\begin{pmatrix}dB_s^1 + dB_s^2\\ dB_s^1 + dB_s^2\end{pmatrix}, \\ \boldsymbol{\Theta^2}\tilde{\boldsymbol{V}}_t &= \frac{\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)}\int_0^t (t-s)^{\alpha-1}\left[\begin{pmatrix}\mu_2\\ \mu_2\end{pmatrix} - \begin{pmatrix}\tilde{V}_s^2\\ \tilde{V}_s^2\end{pmatrix}\right] ds \\ &\quad + \frac{\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)}\int_0^t (t-s)^{\alpha-1}\sqrt{\tilde{V}_s^2 + H_{21}\tilde{V}_s^1}\begin{pmatrix}dB_s^3 + dB_s^4\\ dB_s^3 + dB_s^4\end{pmatrix}.\end{aligned}$$

Note that the above implies that $V^{1+} = V^{1-}$ and $V^{2+} = V^{2-}$. This property is due to the symmetric structure of the baselines and kernels. Therefore, the joint dynamics can be fully captured by considering the joint dynamics of (V^{1+}, V^{2+}). Thus, writing $V^1 := V^{1+} = V^{1-}$ and $V^2 := V^{2+} = V^{2-}$, we have

$$\begin{aligned}\Gamma(\alpha)\frac{\Gamma(1-\alpha)}{\alpha}V_t^1 &= \int_0^t (t-s)^{\alpha-1}\big(\mu_1 - \tilde{V}_s^1\big)ds + \int_0^t \sqrt{V_t^1}\big(dB_s^1 + dB_s^2\big), \\ \Gamma(\alpha)\frac{\Gamma(1-\alpha)}{\alpha}V_t^2 &= \int_0^t (t-s)^{\alpha-1}\big(\mu_2 - \tilde{V}_s^2\big)ds + \int_0^t \sqrt{V_t^2}\big(dB_s^3 + dB_s^4\big).\end{aligned}$$

We can write the above without $\tilde{\boldsymbol{V}}$ as

$$\begin{aligned}\Gamma(\alpha)\frac{\Gamma(1-\alpha)}{\alpha}\begin{pmatrix}V_t^1\\ V_t^2\end{pmatrix} &= \int_0^t (t-s)^{\alpha-1}\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} - \boldsymbol{K}^{-1}\begin{pmatrix}V_s^1\\ V_s^2\end{pmatrix}\right) ds \\ &\quad + \int_0^t (t-s)^{\alpha-1}\begin{pmatrix}\sqrt{V_s^1}\big(dB_s^1 + dB_s^2\big)\\ \sqrt{V_s^2}\big(dB_s^3 + dB_s^4\big)\end{pmatrix}.\end{aligned}$$

Limiting price process. Turning now to the price process, we compute $\boldsymbol{\Delta}$ (see Equation (10)) using the definition. We have

$$\begin{aligned}{}^{\top}\|\boldsymbol{\psi}\|_1\boldsymbol{O}_3 &= \sum_{k\geq 1} {}^{\top}\|\boldsymbol{\phi}\|_1{}^{k}\boldsymbol{O}_3 \\ &= \boldsymbol{O}\sum_{k\geq 1}\left[\left(\int_0^\infty \boldsymbol{C}(s)ds\right)_{11}^{k}\boldsymbol{e}_3 + \left(\int_0^\infty \boldsymbol{C}(s)ds\right)_{12}^{k}\boldsymbol{e}_4\right] \\ &= \sum_{k\geq 1}\left[\left(\int_0^\infty \boldsymbol{C}(s)ds\right)_{11}^{k}\boldsymbol{O}_3 + \left(\int_0^\infty \boldsymbol{C}(s)ds\right)_{12}^{k}\boldsymbol{O}_4\right] \\ &= \left[(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds)^{-1} - \boldsymbol{I}\right]_{11}\boldsymbol{O}_3 + \left[(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds)^{-1} - \boldsymbol{I}\right]_{12}\boldsymbol{O}_4,\end{aligned}$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

which, by definition of $\boldsymbol{\Delta}$, yields

$$\Delta_{11} = \left[\left(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds\right)^{-1} - \boldsymbol{I}\right]_{11} = \frac{2\gamma_2}{4\gamma_1\gamma_2 - \left(H_{12}^c - H_{12}^a\right)\left(H_{21}^c - H_{21}^a\right)} - 1,$$

$$\Delta_{12} = \left[\left(\boldsymbol{I} - \int_0^\infty \boldsymbol{C}(s)ds\right)^{-1} - \boldsymbol{I}\right]_{12} = \frac{H_{21}^c - H_{21}^a}{4\gamma_1\gamma_2 - \left(H_{12}^c - H_{12}^a\right)\left(H_{21}^c - H_{21}^a\right)}.$$

Therefore,

$$\boldsymbol{\Delta} = \frac{1}{4\gamma_1\gamma_2 - \left(H_{12}^c - H_{12}^a\right)\left(H_{21}^c - H_{21}^a\right)} \begin{pmatrix} 2\gamma_2 & H_{21}^c - H_{21}^a \\ H_{12}^c - H_{12}^a & 2\gamma_1 \end{pmatrix} - \boldsymbol{I}.$$

Finally, by application of Theorem 1, any limit point $\boldsymbol{P}$ of the sequence of microscopic price processes satisfies the following equation:

$$\boldsymbol{P}_t = \frac{1}{4\gamma_1\gamma_2 - \left(H_{12}^c - H_{12}^a\right)\left(H_{21}^c - H_{21}^a\right)} \begin{pmatrix} 2\gamma_2 & H_{21}^c - H_{21}^a \\ H_{12}^c - H_{12}^a & 2\gamma_1 \end{pmatrix}$$
$$\begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix} \int_0^t \begin{pmatrix} \sqrt{V_s^1}dB_s^1 \\ \sqrt{V_s^1}dB_s^2 \\ \sqrt{V_s^2}dB_s^3 \\ \sqrt{V_s^2}dB_s^4 \end{pmatrix}$$
$$= \frac{1}{4\gamma_1\gamma_2 - \left(H_{12}^c - H_{12}^a\right)\left(H_{21}^c - H_{21}^a\right)} \begin{pmatrix} 2\gamma_2 & H_{21}^c - H_{21}^a \\ H_{12}^c - H_{12}^a & 2\gamma_1 \end{pmatrix} \int_0^t \begin{pmatrix} \sqrt{V_s^1}\left(dB_s^1 - dB_s^2\right) \\ \sqrt{V_s^2}\left(dB_s^3 - dB_s^4\right) \end{pmatrix}.$$

Introducing the independent bi-dimensional Brownian motions

$$\boldsymbol{Z} := \frac{1}{\sqrt{2}}\begin{pmatrix} B^1 + B^2 \\ B^3 + B^4 \end{pmatrix}, \qquad \boldsymbol{W} := \frac{1}{\sqrt{2}}\begin{pmatrix} B^1 - B^2 \\ B^3 - B^4 \end{pmatrix},$$

this concludes the proof of Corollary 1.

A.4. Proof of Corollary 2

We define the interaction kernel between Asset i and Asset j. For $1 \le i, j \le m$, define

$$\boldsymbol{\phi}_{ij}^T(t) := \begin{cases} \alpha(1 - T^{-\alpha})\mathbb{1}_{t\ge1}t^{-(\alpha+1)} \begin{pmatrix} (1-\gamma) & \gamma \\ \gamma & (1-\gamma) \end{pmatrix} & \text{if } i = j, \\ \alpha T^{-\alpha}\mathbb{1}_{t\ge1}t^{-(\alpha+1)} \begin{pmatrix} H^c & H^a \\ H^a & H^c \end{pmatrix} & \text{if Asset } i \text{ and Asset } j \text{ belong to the same sector,} \\ \alpha T^{-\alpha}\mathbb{1}_{t\ge1}t^{-(\alpha+1)} \begin{pmatrix} H^c + H_r^c & H^a + H_r^a \\ H^a + H_r^a & H^c + H_r^c \end{pmatrix} & \text{otherwise.} \end{cases}$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Finally, the complete Hawkes baseline and kernel structure is

$$\boldsymbol{\mu}^T = T^{\alpha-1}\begin{pmatrix}\mu^1\\ \mu^1\\ \vdots\\ \mu^m\\ \mu^m\end{pmatrix}, \qquad \boldsymbol{\phi}^T = \begin{pmatrix}\boldsymbol{\phi_{11}}^T & \boldsymbol{\phi_{12}}^T & \dots & \boldsymbol{\phi_{1m}}^T\\ \boldsymbol{\phi_{21}}^T & \boldsymbol{\phi_{22}}^T & \dots & \boldsymbol{\phi_{2m}}^T\\ \vdots & \dots & \ddots & \vdots\\ \boldsymbol{\phi_{m1}}^T & \dots & \dots & \boldsymbol{\phi_{mm}}^T\end{pmatrix}.$$

As in the previous example, the proof is split into two steps. First, we show that the kernel satisfies the assumptions required to apply Theorem 1. Then, we compute the equations satisfied by the limiting variance and price processes.

Checking for the assumptions of Theorem 3. We can examine the structure of the kernel as in the two-asset example. Define the following basis:

$$\boldsymbol{O}_i := \begin{cases}\boldsymbol{e}_{2i} + \boldsymbol{e}_{2i+1} \text{if } 1 \le i \le m,\\ \boldsymbol{e}_{2i} - \boldsymbol{e}_{2i} \quad \text{if } m+1 \le i \le 2m.\end{cases}$$

Using the notation of Section 2, straightforward computations allow us to write

$$\boldsymbol{\phi}^T = \boldsymbol{O}\begin{pmatrix}\boldsymbol{A}^T & \boldsymbol{0}\\ \boldsymbol{B}^T & \boldsymbol{C}^T\end{pmatrix}\boldsymbol{O}^{-1} = \boldsymbol{O}\begin{pmatrix}\boldsymbol{A}^T & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{C}^T\end{pmatrix}\boldsymbol{O}^{-1},$$

where we can compute $\boldsymbol{A}^T$ and $\boldsymbol{C}^T$. Checking the assumptions is done as in the two-asset case, though the conditions have changed here because of the new structure of the kernel. For example, since

$$\begin{aligned}\lim_{T\to\infty}\int_0^\infty \boldsymbol{\phi}^T(s)ds\boldsymbol{O}_{m+i} &= (1-2\gamma)\boldsymbol{O}_{n+i} + (H^c - H^a)\sum_{1\le j\ne i\le m}\boldsymbol{O}_{m+j}\\ &\quad + \sum_{1\le j\ne i\le m}\sum_{1\le r\le R}\big(H_r^c - H_r^a\big)\boldsymbol{O}_{m+j},\end{aligned}$$

we have, writing $\boldsymbol{J} := \boldsymbol{e}_1{}^\top\boldsymbol{e}_1 + \dots + \boldsymbol{e}_m{}^\top\boldsymbol{e}_m$ and for any $1 \le r \le R$, $\boldsymbol{J}_r := \boldsymbol{e}_{i_r}{}^\top\boldsymbol{e}_{i_r} + \dots + \boldsymbol{e}_{i_r+m_r}{}^\top\boldsymbol{e}_{i_r+m_r}$,

$$\int_0^\infty \boldsymbol{C}(s)ds = (1-2\gamma)\boldsymbol{I} + \big(H^c - H^a\big)\boldsymbol{J} + \sum_{1\le r\le R}\big(H_r^c - H_r^a\big)\boldsymbol{J_r}.$$

Therefore, as the eigenvalues of $\int_0^\infty \boldsymbol{C}(s)ds$ can be made explicit, if

$$\left|\lambda^- + \sum_{1\le r\le R}\lambda_r^-\right| < 2\gamma,$$

then $\rho(\int_0^\infty \boldsymbol{C}^T(s)ds) < 1$ and $\rho(\int_0^\infty \boldsymbol{C}(s)ds) < 1$. Similarly, we can easily check that a necessary condition for $\rho(\int_0^\infty \boldsymbol{A}^T) < 1$ for T large enough is

$$|\, H^c + H^a + \sum_{1\le r\le R}\frac{m_r - 1}{m-1}\big(H_r^c + H_r^a\big)\,| < \frac{1}{m-1}.$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Since we are interested in the limit where the number of assets grows to infinity, we also impose

$$\left|\lambda^- + \sum_{1\le r\le R} \eta_r\lambda_r^-\right| < 2\gamma,$$

$$\left|\lambda^+ + \sum_{1\le r\le R} \eta_r\lambda_r^+\right| < 1.$$

Combined, we have verified all the assumptions on the structure of the kernel that are needed to apply Theorem 1. We thus move to assumptions on $\boldsymbol{K}$ and $\boldsymbol{\Lambda} = \boldsymbol{K}\boldsymbol{M}^{-1}$. As in the two-asset example, we have here $\boldsymbol{M} = \alpha\boldsymbol{I}$. Since $\boldsymbol{K} = \boldsymbol{I} - (H^c + H^a)\boldsymbol{J} - \sum_{1\le r\le R}(H_r^c + H_r^a)\boldsymbol{J}_r$, the eigenvalues of $\boldsymbol{K}$ (and therefore those of $\boldsymbol{\Lambda}$) are all strictly positive. Thus we have checked all conditions necessary to apply Theorem 1. We can now state the equations satisfied by the limiting variance and price processes.

Limiting variance process. As in the previous example, we have $V^{i+} = V^{i-}$. Thus, we write the underlying variance of asset i as V^i and, using a (slight) abuse of notation, define $\boldsymbol{V} := (V^1, V^2, \cdots, V^m)$. Then $\boldsymbol{V}$ satisfies

$$\boldsymbol{V}_t = \frac{\alpha}{\Gamma(\alpha)\Gamma(1-\alpha)}\int_0^t (t-s)^{\alpha-1}(\boldsymbol{\theta} - \boldsymbol{K}^{-1}\boldsymbol{V}_s)ds + \frac{\alpha\sqrt{2}}{\Gamma(\alpha)\Gamma(1-\alpha)}\int_0^t (t-s)^{\alpha-1}\mathrm{diag}\left(\sqrt{\boldsymbol{V}_s}\right)d\boldsymbol{B}_s,$$

where $\boldsymbol{B}$ is a Brownian motion. We can rewrite $\boldsymbol{K}^{-1}$ as

$$\boldsymbol{K}^{-1} = \left(\boldsymbol{I} - (H^c + H^a)\boldsymbol{J} - \sum_{1\le r\le R}(H_r^c + H_r^a)\boldsymbol{J}_r\right)^{-1} = \left(\boldsymbol{I} - (H^c + H^a)(m-1)\boldsymbol{w}^\top\boldsymbol{w} - \sum_{1\le r\le R}(H_r^c + H_r^a)(m_r-1)\boldsymbol{w}_r{}^\top\boldsymbol{w}_r - \epsilon\right)^{-1},$$

with the small term $\boldsymbol{\epsilon}$ given by

$$\boldsymbol{\epsilon} := (H^c + H^a)(\boldsymbol{J} - (m-1)\boldsymbol{w}^\top\boldsymbol{w}) + \sum_{1\le r\le R}\left(H_r^c + H_r^a\right)\left(\boldsymbol{J}_r - (m_r-1)\boldsymbol{w}_r{}^\top\boldsymbol{w}_r\right).$$

It is easy to check that

$$\rho(\boldsymbol{\epsilon}) \underset{m\to\infty}{=} o\left(\frac{1}{m}\right),$$

which concludes our study of the variance process. We now turn to the equation satisfied by the limiting price process.

Limiting price process. Using the same approach as in the two-asset case, computing $\boldsymbol{\Delta}$ boils down to computing $(I - \int_0^\infty \boldsymbol{C}(s)ds)^{-1}$. Using the expression for $\int_0^\infty \boldsymbol{C}(s)ds$ derived previously, we have

$$(\boldsymbol{I} - \boldsymbol{C})^{-1} = \frac{1}{2\gamma}\left(\boldsymbol{I} - \frac{H^c - H^a}{2\gamma}\boldsymbol{J} - \sum_{1\le r\le R}\frac{H_r^c - H_r^a}{2\gamma}\boldsymbol{J}_r\right)^{-1}.$$

https://doi.org/10.1017/apr.2020.60 Published online by Cambridge University Press

Therefore, repeating the same approach we used for $\boldsymbol{K}^{-1}$ yields

$$(\boldsymbol{I}-\boldsymbol{C})^{-1}=\Big(2\gamma\boldsymbol{I}-\lambda^{-}\boldsymbol{w}^{\top}\boldsymbol{w}-\sum_{1\le r\le R}\eta_r\lambda_r^{-}\boldsymbol{w}_r{}^{\top}\boldsymbol{w}_r-\boldsymbol{\epsilon}\Big)^{-1},$$

with

$$\rho(\boldsymbol{\epsilon})=o\Big(\frac{1}{m}\Big).$$

Thus, we have the following expression for $\boldsymbol{\Delta}$:

$$\boldsymbol{\Delta}=\Big(2\gamma\boldsymbol{I}-\lambda^{-}\boldsymbol{w}^{\top}\boldsymbol{w}-\sum_{1\le r\le R}\eta_r\lambda_r^{-}\boldsymbol{w}_r{}^{\top}\boldsymbol{w}_r-\boldsymbol{\epsilon}\Big)^{-1}-\boldsymbol{I}.$$

Plugging this into Theorem 1, we have the equation satisfied by any limit point $\boldsymbol{P}$ of the sequence $\boldsymbol{P}^T$, which concludes the proof of Corollary 2.

Acknowledgements

M. Rosenbaum and M. Tomas gratefully acknowledge the financial support of the ERC Grant 679836 Staqamof and the Chair Analytics and Models for Regulation. M. Tomas gratefully acknowledges the support of the CFM Chair of Econophysics and Complex Systems. The authors thank Michael Benzaquen and Iacopo Mastromatteo for their helpful comments and are grateful to Eduardo Abi-Jaber, Jean-Philippe Bouchaud, Antoine Fosset, and Paul Jusselin for very fruitful discussions and suggestions.

References

[1] Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J.-F. (2013). Modelling microstructure noise with mutually exciting point processes. *Quant. Finance* **13**, 65–77.

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

[3] Bayer, C., Friz, P. and Gatheral, J. (2016). Pricing under rough volatility. *Quant. Finance* **16**, 887–904.

[4] Benzaquen, M., Mastromatteo, I., Eisler, Z. and Bouchaud, J.-P. (2017). Dissecting cross-impact on stock markets: an empirical analysis. *J. Statist. Mech. Theory Experiment* **2017**, 23406.

[5] Cuchiero, C. and Teichmann, J. (2019). Markovian lifts of positive semidefinite affine Volterra-type processes. *Decisions Econom. Finance* **42**, 407–448.

[6] Da Fonseca, J. and Zhang, W. (2019). Volatility of volatility is (also) rough. *J. Futures Markets* **39**, 600–611.

[7] Dandapani, A., Jusselin, P. and Rosenbaum, M. (2019). From quadratic Hawkes processes to super-Heston rough volatility models with Zumbach effect. Preprint. Available at https://arxiv.org/abs/1907.06151.

[8] El Euch, O., Fukasawa, M. and Rosenbaum, M. (2018). The microstructural foundations of leverage effect and rough volatility. *Finance Stoch.* **22**, 241–280.

[9] El Euch, O., Gatheral, J. and Rosenbaum, M. (2018). Roughening Heston. Available at https://ssrn.com/abstract=3116887.

[10] El Euch, O., Gatheral, J., Radoičić, R. and Rosenbaum, M. (2020). The Zumbach effect under rough Heston. *Quant. Finance* **20**, 235–241.

[11] Gatheral, J., Jaisson, T. and Rosenbaum, M. (2018). Volatility is rough. *Quant. Finance* **18**, 933–949.

[12] Hardiman, S. J., Bercot, N. and Bouchaud, J.-P. (2013). Critical reflexivity in financial markets: a Hawkes process analysis. *Europ. Phys. J. B* **86**, 442.

[13] Hawkes, A. G. (1971). Point spectra of some mutually exciting point processes. *J. R. Statist. Soc. B [Statist. Methodology]* **33**, 438–443.

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

[15] Hawkes, A. G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *J. Appl. Prob.* **11**, 493–503.

[16] HIGHAM, N. J. (2008). *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA.
[17] HORVATH, B., MUGURUZA, A. AND TOMAS, M. (2019). Deep learning volatility. Available at https://ssrn.com/abstract=3322085.
[18] JABER, E. A., CUCHIERO, C., LARSSON, M. AND PULIDO, S. (2019). *A weak solution theory for stochastic Volterra equations of convolution type*. Preprint. Available at https://arxiv.org/abs/1909.01166.
[19] JACOD, J. AND SHIRYAEV, A. (2013). *Limit Theorems for Stochastic Processes*. Springer, Berlin, Heidelberg.
[20] JAISSON, T. AND ROSENBAUM, M. (2015). Limit theorems for nearly unstable Hawkes processes. *Ann. Appl. Prob.* **25**, 600–631.
[21] JAISSON, T. AND ROSENBAUM, M. (2016). Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *Ann. Appl. Prob.* **26**, 2860–2882.
[22] JUSSELIN, P. AND ROSENBAUM, M. (2018). No-arbitrage implies power-law market impact and rough volatility. Preprint. Available at https://arxiv.org/abs/1805.07134.
[23] LALOUX, L., CIZEAU, P., BOUCHAUD, J.-P. AND POTTERS, M. (1999). Noise dressing of financial correlation matrices. *Phys. Rev. Lett.* **83**, 1467.
[24] LIVIERI, G., MOUTI, S., PALLAVICINI, A. AND ROSENBAUM, M. (2018). Rough volatility: evidence from option prices. *IISE Trans.* **50**, 767–776.
[25] REIGNERON, P.-A., ALLEZ, R. AND BOUCHAUD, J.-P. (2011). Principal regression analysis and the index leverage effect. *Physica A* **390**, 3026–3035.
[26] REVUZ, D. AND YOR, M. (2013). *Continuous martingales and Brownian motion*. Springer, Berlin, Heidelberg.
[27] VERAAR, M. (2012). The stochastic Fubini theorem revisited. *Stochastics* **84**, 543–551.