

Some limit theorems for Hawkes processes and application to financial statistics

E. Bacry*, S. Delattre†, M. Hoffmann‡ and J.F. Muzy§

Abstract

In the context of statistics for random processes, we prove a law of large numbers and a functional central limit theorem for multivariate Hawkes processes observed over a time interval $[0, T]$ when $T \rightarrow \infty$. We further exhibit the asymptotic behaviour of the covariation of the increments of the components of a multivariate Hawkes process, when the observations are imposed by a discrete scheme with mesh Δ over $[0, T]$ up to some further time shift τ . The behaviour of this functional depends on the relative size of Δ and τ with respect to T and enables to give a full account of the second-order structure. As an application, we develop our results in the context of financial statistics. We introduced in [7] a microscopic stochastic model for the variations of a multivariate financial asset, based on Hawkes processes and that is confined to live on a tick grid. We derive and characterise the exact macroscopic diffusion limit of this model and show in particular its ability to reproduce important empirical stylised fact such as the Epps effect and the lead-lag effect. Moreover, our approach enables to track these effects across scales in rigorous mathematical terms.

Keywords: Point processes. Hawkes processes. Limit theorems. Discretisation of stochastic processes. Statistics of random processes

Mathematical Subject Classification: 60F05, 60G55, 62M10.

*CMAP CNRS-UMR 7641 and École Polytechnique, 91128 Palaiseau, France

†Université Paris Diderot and LPMA CNRS-UMR 7599, Boîte courrier 7012 75251 Paris Cedex 16, France

‡Université Paris-Dauphine and CEREMADE CNRS-UMR 7534, Place du maréchal de Lattre de Tassigny, 75775 Paris, France

§SPE CNRS-UMR 6134 and Université de Corte, 20250 Corte, France

1 Introduction

1.1 Motivation and setting

Point processes have long served as a representative model for event-time based stochastic phenomena that evolve in continuous time. A comprehensive mathematical development of the theory of point processes can be found in the celebrated textbook of Daley and Vere-Jones [12], see also the references therein. In this context, mutually exciting processes form a specific but quite important class of point processes that are mathematically tractable and widely used in practice. They were first described by Hawkes in 1971 [17, 16]. Informally, to a multivariate d -dimensional counting process $N = (N_1, \dots, N_d)$ with values in \mathbb{N}^d is associated an intensity function $(\lambda_1, \dots, \lambda_d)$ defined by

$$P(N_i \text{ has a jump in } [t, t + dt] \mid \mathcal{F}_t) = \lambda_{i,t} dt, \quad i = 1, \dots, d$$

where P stands for probability and \mathcal{F}_t is the sigma-field generated by N up to present time t . A multivariate Hawkes process has intensity

$$\lambda_{i,t} = \mu_i + \int_{(0,t)} \sum_{j=1}^d \varphi_{ij}(t-s) dN_{j,s}, \quad i = 1, \dots, d \quad (1)$$

and is specified by $\mu_i \in \mathbb{R}_+ = [0, \infty)$ and for $i = 1, \dots, d$, the φ_{ij} are functions from \mathbb{R}_+ to \mathbb{R}_+ . More in Section 2 below for rigorous definitions.

The properties of Hawkes processes are fairly well known: from a probabilistic point a view, the aforementioned book of Daley and Vere-Jones [12] gives a concise synthesis of earlier results published by Hawkes [17, 16, 18] that focus on spectral analysis following Bartlett [9] and cluster representation, see Hawkes and Oakes [19]. From a statistical perspective, Ogata studied in [25] the maximum likelihood estimator. Recently, the nonparametric estimation of the intensity functions has been investigated by Reynaud-Bouret and Schbath [26] and Al Dayri *et al.* [13].

However, in all these papers and the references therein, the focus is on the “microscopic properties” of Hawkes processes, *i.e.* their infinitesimal evolution, possibly under a stationary regime or for a large time horizon $[0, T]$ in order to guarantee a large number of jumps for statistical inference purposes. In the present work, we are rather interested in the “macroscopic properties” of Hawkes processes, in the sense of obtaining a limit behaviour for the multivariate process $(N_{Tv})_{v \in [0,1]}$ as $T \rightarrow \infty$, for a suitable normalisation. Our interest originates in financial data modelling: in [7], we introduced a stochastic model for the price $S = (S_1, \dots, S_n)$ of a multivariate asset, based on a

$d = 2n$ dimensional Hawkes process of the form (1), with representation

$$S_1 = N_1 - N_2, S_2 = N_3 - N_4, \dots, S_n = N_{d-1} - N_d. \quad (2)$$

In this context, the fact that the S_i take values on \mathbb{Z} accounts for the discreteness of the price formation under a limit order book. If we take $\mu_{2i-1} = \mu_{2i}$ for every $1 \leq i \leq d$, the process S is centred and the mutually exciting properties of the intensity processes λ_i under (1) allow to reproduce empirically microstructure noise and the Epps effect, as demonstrated in [7]. Microstructure noise – a major stylised fact in high frequency financial data (see *e.g.* [6, 4, 5, 27, 28, 8]) – is characterised by the property that in microscopic scales¹, an upward change of price for the component S_i is more likely to be followed by a downward jump of S_i and *vice versa*. Similarly, a jump of a component N_{2i-1} or N_{2i} of S_i at time t will increase the conditional probability, given the history up to time t , that a jump occurs for the component N_{2j-1} or N_{2j} of S_j through the exciting effect of the kernels $\varphi_{2i-1,2j-1}$, $\varphi_{2i-1,2j}$, $\varphi_{2i,2j-1}$ or $\varphi_{2i,2j}$, thus creating a certain dependence structure and the Epps effect, another well-documented stylised fact implying that the empirical correlation between two assets vanishes on microscopic scales [15]. However, both microstructure noise and the Epps effect vanish at coarser scales where a standard diffusion regime dominates. Thus, a certain macroscopic stability property is desirable, namely the property that S behaves like a continuous diffusion on coarse scales. Also, one would like to be able to track the aforementioned microscopic effects produced by the kernels φ_{ij} in the diffusion limit of S . This is the main topic of the paper.

The importance of mutually exciting processes has expanded in financial econometrics over the last years, see [3, 10, 14, 20] among others. From a wider perspective however, a mathematical analysis of the limiting behaviour of Hawkes processes has a relevance beyond finance: these include seismology of course, for which Hawkes processes were originally introduced (see [24] for instance), and in a more prospective way, one could mention recent related areas such as traffic networks [11] or genomic analysis [26] where mutually exciting processes take a growing importance.

1.2 Main results and organisation of the paper

We present our results in a general setting in Sections 2, 3 and 4 with relatively minimal assumptions. It is only in the subsequent Section 5 that we consistently apply our results to the multivariate price model S described by (1)-(2) for financial data modelling.

¹when the data are sampled every few seconds or less.

In Section 2, we recall the construction of Hawkes processes in a rigorous and general setting. We state in particular the fact that as soon as the kernels φ_{ij} that define the intensity process λ are locally integrable, the existence of N with uniqueness in law is guaranteed on a rich enough space, as follows from the theory of predictable projection of integer-valued random measures of Jacod [22]. We state in Section 3 a law of large numbers (Theorem 1) and a functional central limit theorem (Theorem 2). The law of large numbers takes the form

$$\sup_{v \in [0,1]} \|T^{-1}N_{Tv} - v(\mathbf{Id} - \mathbf{K})^{-1}\mu\| \rightarrow 0 \text{ as } T \rightarrow \infty \quad (3)$$

almost surely and in $L^2(P)$. We set $\|\cdot\|$ for the Euclidean norm on \mathbb{R}^d . The limit is described by $\mu = (\mu_1, \dots, \mu_d)$ identified with a column vector and the $d \times d$ matrix $\mathbf{K} = \int_0^\infty \varphi(t) dt$, where $\varphi = (\varphi_{ij})$ is a matrix of functions from \mathbb{R}_+ to \mathbb{R}_+ . The convergence (3) holds under the assumption that the spectral radius of \mathbf{K} is strictly less than 1, which guarantees in particular all the desirable integrability properties for N . As for a central limit theorem, under the additional condition that $t^{1/2}\varphi(t)$ is integrable componentwise, a consequence of Theorem 2 is the convergence of the processes

$$\sqrt{T} \left(T^{-1}N_{Tv} - v(\mathbf{Id} - \mathbf{K})^{-1}\mu \right), \quad v \in [0, 1]$$

to $(\mathbf{Id} - \mathbf{K})^{-1}\Sigma^{1/2}W_v, v \in [0, 1]$ in law for the Skorokod topology. The limit is described by a standard d -dimensional Brownian motion $(W_v)_{v \in [0,1]}$ and the $d \times d$ diagonal matrix Σ defined by $\Sigma_{ii} = ((\mathbf{Id} - \mathbf{K})^{-1}\mu)_i$. The proof of Theorems 1 and 2 relies on the intensity theory of point processes: using that $N_t - \int_0^t \lambda_s ds$ is a d -dimensional martingale, if we set $X_t = N_t - E(N_t)$ with $E(\cdot)$ denoting expectation, we then have the fundamental representation $X_t = M_t + \int_0^t \varphi(t-s)X_s ds$ of X as a sum of a martingale and a convolution product of φ with X itself. Separating the two components, we can then take advantage of the powerful theory of limit theorems of semimartingales, as exposed in the comprehensive book of Jacod and Shiryaev [23].

In Section 4, we consider yet another angle of study that is useful for applications. Given two square integrable point processes N and N' with values in \mathbb{N}^d , setting $X_t = N_t - E(N_t)$ and $X'_t = N'_t - E(N'_t)$, one can define the empirical cross-correlation of N and N' at scale Δ over $[0, T]$ as

$$\mathbf{V}_{\Delta, T}(N, N') = \frac{1}{T} \sum_{i=1}^{\lfloor T/\Delta \rfloor} (X_{i\Delta} - X_{(i-1)\Delta})(X'_{i\Delta} - X'_{(i-1)\Delta})^\top,$$

where X_t and X'_t are identified as column vectors and $(\cdot)^\top$ denotes transposition in \mathbb{R}^d . In rigorous terms, $\mathbf{V}_{\Delta, T}(N, N')$ shall rather be called the empirical

correlation matrix of the increments of X and X' sampled at scale Δ over $[0, T]$. This object is of major importance in practice: viewed as a function of Δ , it reveals the correlation structure across scales between N and N' . It is crucial for the understanding of the transition of $(N_{Tv})_{v \in [0,1]}$ from a microscopic regime of a point process to the macroscopic behaviour of a Brownian diffusion, up to appropriate normalisation, as $T \rightarrow \infty$. More precisely, given a time shift $\tau = \tau_T \in \mathbb{R}$, we focus in the paper on the τ -shifted empirical correlation matrix of N at scale Δ , namely² $\mathbf{V}_{\Delta, T}(N, N_{\tau+})$. We prove in Theorem 3 that $\mathbf{V}_{\Delta, T}(N, N_{\tau+})$ is close in $L^2(P)$ to a deterministic explicit counterpart

$$\mathbf{v}_{\Delta, \tau} = \int_{\mathbb{R}_+^2} \left(1 - \frac{|t-s-\tau|}{\Delta}\right)^+ (\mathbf{Id} \delta_0(ds) + \boldsymbol{\psi}(s)ds) \boldsymbol{\Sigma} (\mathbf{Id} \delta_0(dt) + \boldsymbol{\psi}(t)^\top dt)$$

under appropriate conditions on τ and Δ relative to T as $T \rightarrow \infty$. The approximation $\mathbf{v}_{\Delta, \tau}$ is described by $\boldsymbol{\Sigma}$ that already appears in Theorems 1 and 2 and $\boldsymbol{\psi} = \sum_{n \geq 1} \boldsymbol{\varphi}_n$, where $\boldsymbol{\varphi}_n$ is the matrix of n -fold convolution product of $\boldsymbol{\varphi}$. This apparently cumbersome formula paves the way to explicit computations that yield crucial information about the dependence structure of functionals of N , as later developed in the paper.

Section 5 is devoted to some applications of Theorems 1, 2 and 3 to the study of the multivariate price model S described by (1)-(2) for financial data modelling. We elaborate on particular examples the insight given by $\mathbf{v}_{\Delta, \tau}$. Concerning macroscopic limits of a univariate price $S = N_1 - N_2$ obtained for $d = 2$, we derive in Proposition 1 the limiting variance of the rescaled process $T^{-1/2}(N_{1, Tv} - N_{2, Tv})$ in the special case where $\varphi_{1,2} = \varphi_{2,1} = \varphi$ and $\varphi_{1,1} = \varphi_{2,2} = 0$. This oversimplification enables to focus on the phenomenon of microstructure noise and reveals the macroscopic trace of φ in the variance of the limiting Brownian motion. We discuss the theoretical implications of the obtained formulas in statistical finance. In the same way, we explicitly characterise in Proposition 2 the macroscopic correlation structure obtained with a bivariate price process $S = (S_1, S_2) = (N_1 - N_2, N_3 - N_4)$ when only cross-excitations $\varphi_{1,3} = \varphi_{3,1}$ and $\varphi_{2,4} = \varphi_{4,2}$ are considered, the other components φ_{ij} being set to 0 for simplicity. Again, we obtain an illuminating formula that characterises the influence of the microscopic dynamics on the macroscopic correlation of the Brownian limit. This enables in particular in Proposition 3 to demonstrate how Hawkes process can account for the Epps effect [15] and also the lead-lag effect between two financial assets that has been given some attention in the literature recently [21, 2, 1].

²This is actually a more general object since we can formally recover $\mathbf{V}_{\Delta, T}(N, N')$ from $\mathbf{V}_{\Delta, T}(N'', N''_{\tau+})$, where $N'' = (N, N')$ is a $2d$ -dimensional point process by letting $\tau \rightarrow 0$ in $\mathbf{V}_{\Delta, T}(N'', N''_{\tau+})$.

Section 6 develops some preliminary tools for the proofs. Section 7, 8 and 9 are devoted to the proof of Theorems 1, 2 and 3 respectively. Some technical results and computations of the application Section 5 are delayed until an appendix.

2 Multivariate Hawkes processes

Consider a measurable space (Ω, \mathcal{F}) on which is defined a non-decreasing sequence of random variables $(T_n)_{n \geq 1}$ taking their values in $(0, \infty]$, and such that $T_n < T_{n+1}$ on the event $\{T_n < \infty\}$, for all $n \geq 1$. Let $(Z_n)_{n \geq 1}$ be a sequence of discrete random variables taking their values in $\{1, \dots, d\}$ for some positive integer d . Define, for $t \geq 0$

$$N_{i,t} = \sum_{n \geq 1} \mathbf{1}_{\{T_n \leq t\} \cap \{Z_n = i\}}.$$

Remark that $N_{i,0} = 0$ by construction. We endow Ω with the filtration $(\mathcal{F}_t)_{t \geq 0}$ where \mathcal{F}_t is the σ -algebra generated by the random variables $N_{i,s}$, $s \leq t$, $1 \leq i \leq d$. According to Jacod [22], for any progressively measurable non-negative processes $(\lambda_{1,t})_{t \geq 0}, \dots, (\lambda_{d,t})_{t \geq 0}$ satisfying $\int_0^{T_n} \lambda_{i,s} ds < \infty$ almost surely, there exists at most one probability measure P on $(\Omega, \mathcal{F}_\infty)$ such that the compensator (or predictable projection) of the integer-valued random measure

$$N(dt, dx) = \sum_{n \geq 1} \mathbf{1}_{\{T_n < \infty\}} \delta_{(T_n, Z_n)}(dt, dx)$$

on $(0, \infty) \times \{1, \dots, d\}$ is $\nu(dt, dx) = \sum_{i=1}^d \lambda_{i,t} dt \otimes \delta_i(dx)$, where δ is the Dirac mass. In other words, for all $n \geq 1$, for all $i \in \{1, \dots, d\}$, the process

$$N_{i,t \wedge T_n} - \int_0^{t \wedge T_n} \lambda_{i,s} ds$$

is a (\mathcal{F}_t) -martingale. This implies that the law of the d -dimensional process (N_1, \dots, N_d) is characterised by $(\lambda_1, \dots, \lambda_d)$. Moreover if Ω is rich enough we have the existence of such a probability measure P .

Definition 1. We say that $N = (N_1, \dots, N_d)$ is a multivariate Hawkes process when

$$\lambda_{i,t} = \mu_i + \int_{(0,t)} \sum_{j=1}^d \varphi_{ij}(t-s) dN_{j,s} \quad (4)$$

where $\mu_i \in \mathbb{R}_+$ and $\varphi_{i,j}$ is a function from \mathbb{R}_+ to \mathbb{R}_+ .

We have a non-explosion criterion, as a consequence for instance of equality (10) in Lemma 2, Section 6 below.

Lemma 1 (Non-explosion criterion). *Set $T_\infty = \lim_n T_n$. Assume that the following holds:*

$$\int_0^t \varphi_{ij}(s) ds < \infty \quad \text{for all } i, j, t. \quad (5)$$

Then $T_\infty = \infty$ almost surely.

3 Law of large numbers and functional central limit theorem

On a rich enough probability space (Ω, \mathcal{F}, P) , we consider a Hawkes process $N = (N_t)_{t \geq 0}$ according to Definition 1, satisfying (5) and specified by the vector $\mu = (\mu_1, \dots, \mu_d)$ and the $d \times d$ -matrix valued function $\varphi = (\varphi_{i,j})_{1 \leq i, j \leq d}$. Note that in this setting, we do not assume a stationary regime for N . Consider the assumption

$$\text{For all } i, j \text{ we have } \int_0^\infty \varphi_{ij}(t) dt < \infty \text{ and the spectral radius } \rho(\mathbf{K}) \text{ of the matrix } \mathbf{K} = \int_0^\infty \varphi(t) dt \text{ satisfies } \rho(\mathbf{K}) < 1. \quad (\mathbf{A1})$$

First we have a law of large numbers in the following sense:

Theorem 1. *Assume that (A1) holds. Then $N_t \in L^2(P)$ for all $t \geq 0$ and we have $\sup_{v \in [0,1]} \|T^{-1} N_{Tv} - v(\mathbf{Id} - \mathbf{K})^{-1} \mu\| \rightarrow 0$ as $T \rightarrow \infty$ almost-surely and in $L^2(P)$.*

Next we have an associated functional central-limit theorem. Introduce the functions φ_n defined on \mathbb{R}_+ and with values in the set of $d \times d$ -matrices with entries in $[0, \infty]$ by

$$\varphi_1 = \varphi, \quad \varphi_{n+1}(t) = \int_0^t \varphi(t-s) \varphi_n(s) ds, \quad n \geq 1. \quad (6)$$

Under (A1) we have $\int_0^\infty \varphi_n(t) dt = \mathbf{K}^n$ hence the series $\sum_{n \geq 1} \varphi_n$ converges in $L^1(dt)$. We set

$$\psi = \sum_{n \geq 1} \varphi_n. \quad (7)$$

Theorem 2. *Assume that (A1) holds. We have*

$$E(N_t) = t\mu + \left(\int_0^t \psi(t-s) s ds \right) \mu$$

where $\psi \in L^1(dt)$ is given by (7). Moreover, the processes

$$\frac{1}{\sqrt{T}} \left(N_{Tv} - E(N_{Tv}) \right), v \in [0, 1]$$

converge in law for the Skorokod topology to $(\mathbf{Id} - \mathbf{K})^{-1} \boldsymbol{\Sigma}^{1/2} W_v, v \in [0, 1]$ as $T \rightarrow \infty$, where $(W_v)_{v \in [0, 1]}$ is a standard d -dimensional Brownian motion and $\boldsymbol{\Sigma}$ is the diagonal matrix such that $\Sigma_{ii} = ((\mathbf{Id} - \mathbf{K})^{-1} \mu)_i$.

Consider now the following restriction on φ :

$$\int_0^\infty \varphi(t) t^{1/2} dt < \infty \quad \text{componentwise.} \quad (\mathbf{A2})$$

Using Theorem 1 and Assumption (A2), we may replace $T^{-1}E(N_{Tv})$ by its limit in Theorem 2 and obtain the following corollary.

Corollary 1. *Assume that (A1) and (A2) hold. Then the processes*

$$\sqrt{T} \left(\frac{1}{T} N_{Tv} - v(\mathbf{Id} - \mathbf{K})^{-1} \mu \right), \quad v \in [0, 1]$$

converge in law for the Skorokod topology to $(\mathbf{Id} - \mathbf{K})^{-1} \boldsymbol{\Sigma}^{1/2} W_v, v \in [0, 1]$ as $T \rightarrow \infty$.

4 Empirical covariation across time scales

For two square integrable counting processes N and N' with values in \mathbb{N}^d , set

$$X_t = N_t - E(N_t), \quad X'_t = N'_t - E(N'_t).$$

The empirical covariation of N and N' is the process $\mathbf{V}_{\Delta, T}(N, N')$, $T > 0$, $\Delta > 0$, taking values in the set of $d \times d$ matrices and defined as

$$\mathbf{V}_{\Delta, T}(N, N') = \frac{1}{T} \sum_{i=1}^{\lfloor T/\Delta \rfloor} (X_{i\Delta} - X_{(i-1)\Delta}) (X'_{i\Delta} - X'_{(i-1)\Delta})^\top$$

where $(X_{i\Delta} - X_{(i-1)\Delta})$ is identified as a column vector and $(X_{i\Delta} - X_{(i-1)\Delta})^\top$ denotes its transpose, and we set $X_t = 0$ for $t \leq 0$. More precisely, given a time shift $\tau \in \mathbb{R}$, we are interested in the behaviour of $\mathbf{V}_{\Delta, T}(N, N_{\tau+})$. In essence, $\mathbf{V}_{\Delta, T}(N, N_{\tau+})$ can be viewed as a multivariate cross-correlogram across scales Δ of N : it can be consistently measured from empirical data and its limiting behaviour as $T \rightarrow \infty$ plays a key tool in understanding the second-order structure of linear functions of N across scales Δ .

Theorem 3. *In the same setting as in Theorem 2, let $(\Delta_T)_{T>0}$ and $(\tau_T)_{T>0}$ be two families of real numbers such that $\Delta_T > 0$. If $\Delta_T/T \rightarrow 0$ and $\tau_T/T \rightarrow 0$ as $T \rightarrow \infty$, we have*

$$\mathbf{V}_{\Delta_T, T}(X, X_{\tau_T+}) - \mathbf{v}_{\Delta_T, \tau_T} \rightarrow 0 \text{ as } T \rightarrow \infty \text{ in } L^2(P)$$

where

$$\begin{aligned} \mathbf{v}_{\Delta, \tau} &= \left(1 - \frac{|\tau|}{\Delta}\right)^+ \boldsymbol{\Sigma} + \int_{\mathbb{R}_+^2} ds dt \left(1 - \frac{|t-s-\tau|}{\Delta}\right)^+ \boldsymbol{\psi}(s) \boldsymbol{\Sigma} \boldsymbol{\psi}(t)^\top + \\ &\quad + \int_0^\infty ds \left(1 - \frac{|s+\tau|}{\Delta}\right)^+ \boldsymbol{\psi}(s) \boldsymbol{\Sigma} + \int_0^\infty ds \left(1 - \frac{|s-\tau|}{\Delta}\right)^+ \boldsymbol{\Sigma} \boldsymbol{\psi}(s)^\top, \end{aligned} \quad (8)$$

or equivalently,

$$\mathbf{v}_{\Delta, \tau} = \int_{\mathbb{R}_+^2} \left(1 - \frac{|t-s-\tau|}{\Delta}\right)^+ (\mathbf{Id} \delta_0(ds) + \boldsymbol{\psi}(s) ds) \boldsymbol{\Sigma} (\mathbf{Id} \delta_0(dt) + \boldsymbol{\psi}(t)^\top dt),$$

where $\boldsymbol{\Sigma}$ is the diagonal matrix such that $\boldsymbol{\Sigma}_{ii} = ((\mathbf{Id} - \mathbf{K})^{-1} \boldsymbol{\mu})_i$ and the function $\boldsymbol{\psi}$ is given by (7).

Remark 1. One can check that $\mathbf{v}_{\Delta, \tau} = E((N_\Delta - N_0)(N_{\Delta+\tau} - N_\tau)^\top)$ where N is a counting process of the (unique in law) stationary multivariate Hawkes process on \mathbb{R} associated to $\boldsymbol{\mu}$ and $\boldsymbol{\varphi}$. Thus, another way to obtain $\mathbf{v}_{\Delta, \tau}$ is to compute $E((N_\Delta - N_0)(N_{\Delta+\tau} - N_\tau)^\top)$ in the stationary regime, as in [17, 16, 18] by means of the Bartlett spectrum of N , see [9]. However, the stationary restriction is superfluous and moreover, only very specific parametric form of φ_{ij} like exponential functions enable to carry such computations.

Remark 2. Obviously, we have $\mathbf{v}_{\Delta, \tau} = \mathbf{v}_{\Delta, -\tau}^\top$.

Remark 3. For fixed $\tau \in \mathbb{R}$, we retrieve the macroscopic variance of Theorem 2 by letting $\Delta \rightarrow \infty$. More precisely, we have

$$\mathbf{v}_{\Delta, \tau} \rightarrow \boldsymbol{\Gamma} \boldsymbol{\Sigma} \boldsymbol{\Gamma}^\top \text{ as } \Delta \rightarrow \infty$$

where $\boldsymbol{\Gamma} = (\mathbf{Id} - \mathbf{K})^{-1}$, and the effect of τ vanishes as $\Delta \rightarrow \infty$. For all $\tau \neq 0$ we have $\mathbf{v}_{\Delta, \tau} \rightarrow \mathbf{0}$ as $\Delta \rightarrow 0$. This convergence simply expresses the fact that the two processes N and $N_{\tau+}$ cannot jump at the same time, producing a flat autocorrelogram for sufficiently small sampling mesh Δ .

In the same way as Corollary 1 is obtained from Theorem 2, we have the following refinement of Theorem 3.

Corollary 2. *Finally, in the same setting as in Theorem 3, assume moreover that (A2) holds. Define*

$$\tilde{X}_t = \begin{cases} N_t - t(\mathbf{Id} - \mathbf{K})^{-1}\mu & \text{if } t \geq 0, \\ 0 & \text{if } t < 0. \end{cases}$$

We have $\mathbf{V}_{\Delta T, T}(\tilde{X}, \tilde{X}_{\tau_T+}) - \mathbf{v}_{\Delta T, \tau_T} \rightarrow 0$ as $T \rightarrow \infty$ in $L^2(P)$, where $\mathbf{v}_{\Delta, \tau}$ is given by (8).

5 Application to financial statistics

5.1 The macroscopic trace of microstructure noise

Following [7], we introduce a univariate price process $S = (S_t)_{t \geq 0}$ by setting

$$S = N_1 - N_2,$$

where (N_1, N_2) is a Hawkes process in the case $d = 2$, with

$$\begin{cases} \lambda_{1,t} &= \nu + \int_{(0,t)} \varphi(t-s) dN_{2,s}, \\ \lambda_{2,t} &= \nu + \int_{(0,t)} \varphi(t-s) dN_{1,s} \end{cases}$$

for some $\nu \in \mathbb{R}_+$ and $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$. With our notation, this corresponds to having $\mu = (\nu, \nu)$ and

$$\varphi = \begin{pmatrix} 0 & \varphi \\ \varphi & 0 \end{pmatrix}$$

If $\varphi = 0$, we find back a compound Poisson process with intensity ν and symmetric Bernoulli jumps. This corresponds to the simplest model for a random walk in continuous time, constrained to live on a lattice, the tick-grid in financial statistics accounting for the discreteness of price at fine scales. Microstructure noise corresponds to the property that an upward jump of S will be more likely followed by a downward jump and vice versa. This phenomenon lays its roots in microeconomic analysis of price manipulation of agents [6, 4, 5, 27, 28, 8]. In our simple phenomenological setting, it will be reproduced by the introduction of the kernel φ , as empirically demonstrated in [7]. The question we can now address is the macroscopic stability of the model. Do we retrieve a standard diffusion in the limit $T \rightarrow \infty$ for an appropriate scaling of S and how does the effect of φ influence the macroscopic volatility? By Theorems 1 and 2, we readily obtain an explicit answer:

Proposition 1 (Macroscopic trace of microstructure noise). *Assume that $\|\varphi\|_{L^1} = \int_0^\infty \varphi(t)dt < 1$. Then*

$$(T^{-1/2}S_{Tv}, v \in [0, 1]) \rightarrow (\sigma W_v, v \in [0, 1]) \text{ as } T \rightarrow \infty,$$

in law for the Skorokod topology, where $(W_v)_{v \in [0,1]}$ is a standard Brownian motion and

$$\sigma^2 = \frac{2\nu}{(1 - \|\varphi\|_{L^1})(1 + \|\varphi\|_{L^1})^2}.$$

Remark 4. Of course, a Brownian limit is too simple if financial assets are considered over large amount of times such as a month or a year, where path dependent or even stochastic volatility models are more appropriate. The term macroscopic in our context shall rather be considered in the scales of a few hours, see [7] for a discussion on empirical data.

Remark 5. Note that if we take $\varphi = 0$, we retrieve the standard convergence of a compound Poisson process with symmetric jumps to a Brownian motion.

Remark 6. By assumption, $0 \leq \|\varphi\|_{L^1} < 1$ and a closer inspection of the function

$$x \rightsquigarrow \sigma(\nu, x)^2 = \frac{2\nu}{(1-x)(1+x)^2} \text{ for } x \in [0, 1)$$

reveals an interesting feature: for small microstructure effect (namely if $x = \|\varphi\|_{L^1}$ less than 1/3) the effect of microstructure tends to stabilise the macroscopic variance in the sense that $\sigma(\nu, x)^2 \leq \sigma(\nu, 0)^2$, whereas beyond a critical value $x \approx 0.61$, we have $\sigma(\nu, x)^2 \geq \sigma(\nu, 0)^2$ and even $\sigma(\nu, x)^2 \rightarrow \infty$ as $x \rightarrow 1$.

5.2 Macroscopic correlations for bivariate assets

We now turn to a bivariate price model $S = (S_1, S_2)$ obtained from a Hawkes process in dimension $d = 4$, of the form

$$(S_1, S_2) = (N_1 - N_2, N_3 - N_4)$$

with $\mu = (\mu_1, \mu_2, \mu_3, \mu_4)$ such that $\mu_1 = \mu_2$ and $\mu_3 = \mu_4$ together with

$$\varphi = \begin{pmatrix} 0 & 0 & h & 0 \\ 0 & 0 & 0 & h \\ g & 0 & 0 & 0 \\ 0 & g & 0 & 0. \end{pmatrix}$$

The upward jumps of S_1 excite the intensity of the upward jumps of S_2 and, in a symmetric way, the downward jumps of S_1 excite the downward jumps

of S_2 via the kernel $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$. Likewise, the upward jumps of S_2 excite the upward jumps of S_1 and the downward jumps of S_2 excite the downward jumps of S_1 via the kernel $h : \mathbb{R}_+ \rightarrow \mathbb{R}_+$. For simplicity we ignore other cross terms that could produce microstructure noise within the inner jumps of S_1 and S_2 . This relatively simple dependence structure at a microscopic level enables to obtain a non-trivial form of the macroscopic correlation of the diffusion limits of S_1 and S_2 .

Proposition 2. *Assume that $\|h\|_{L^1}\|g\|_{L^1} < 1$. Then, the 2-dimensional processes $T^{-1/2}(S_{1,Tv}, S_{2,Tv})_{v \in [0,1]}$ converge in law as $T \rightarrow \infty$ for the Skorokod topology to*

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \frac{\sqrt{2}}{(1 - \|h\|_{L^1}\|g\|_{L^1})^{3/2}} \begin{pmatrix} \nu_1^{1/2} W_1 + \nu_2^{1/2} \|h\|_{L^1} W_2 \\ \nu_1^{1/2} \|g\|_{L^1} W_1 + \nu_2^{1/2} W_2 \end{pmatrix}$$

with

$$\nu_1 = \mu_1 + \|h\|_{L^1}\mu_3, \quad \nu_2 = \mu_3 + \|g\|_{L^1}\mu_1. \quad (9)$$

and where $(W_1, W_2) = (W_{1,t}, W_{2,t})_{t \in [0,1]}$ is a standard Brownian motion.

The proof is a consequence of Theorems 1 and 2 and is given in appendix.

Remark 7. The macroscopic correlation between S_1 and S_2 is thus equal to the cosine of the angle of the two vectors

$$(\nu_1^{1/2}, \nu_2^{1/2} \|h\|_{L^1}) \quad \text{and} \quad (\nu_1^{1/2} \|g\|_{L^1}, \nu_2^{1/2}).$$

Obviously, it is always nonnegative and strictly less than 1 since the determinant $\nu_1^{1/2}\nu_2^{1/2}(1 - \|h\|_{L^1}\|g\|_{L^1})$ of the two above vectors is positive unless the μ_i are all 0.

5.3 Lead-lag and Epps effect through the cross correlations across-scales

We keep up with the model and the notation of Section 5.2 but we now study the quantities

$$V_{\Delta,T}(S_1, S_{1,\tau+}), \quad V_{\Delta,T}(S_2, S_{2,\tau+}), \quad V_{\Delta,T}(S_1, S_{2,\tau+}).$$

In particular, the quantity $V_{\Delta,T}(S_1, S_{2,\tau+})$ is a powerful tool for the statistical study of lead-lag effects, *i.e.* the fact that the jumps of S_1 can anticipate on those of S_2 and vice-versa, see for instance [1, 21]. The Epps effect – *i.e.* the stylised fact statement that the correlation between the increments of two

assets vanishes at fine scales – can be tracked down likewise. Theorem 3 enables to characterise in principle the limiting behaviour of these functionals. This is described in details in Proposition 3 below.

We consider g and h as functions defined on \mathbb{R} by setting $g(t) = h(t) = 0$ is $t < 0$. Assume that $\|h\|_{L^1}\|g\|_{L^1} < 1$. Then the series $F := \sum_{n \geq 1} (h \star g)^{\star n}$ converges in $L^1(\mathbb{R}, dt)$. If f is a function on \mathbb{R} we define \check{f} by $\check{f}(t) = f(-t)$. We have

Proposition 3. *Assume that $\|h\|_{L^1}\|g\|_{L^1} < 1$. Let $(\Delta_T)_{T>0}$ and $(\tau_T)_{T>0}$ be two families of real numbers such that $\Delta_T > 0$. If $\Delta_T/T \rightarrow 0$ and $\tau_T/T \rightarrow 0$ as $T \rightarrow \infty$ we have*

$$V_{\Delta_T, T}(S_1, S_{1, \tau_T+}) - C_{11}(\Delta_T, \tau_T) \rightarrow 0 \text{ as } T \rightarrow \infty \text{ in } L^2(P),$$

where

$$C_{11}(\Delta, \tau) = \frac{2}{1 - \|h\|_{L^1}\|g\|_{L^1}} \gamma_\Delta \star (\delta_0 + F + \check{F} + F \star \check{F}) \star (\nu_1 \delta_0 + \nu_2, h \star \check{h})(\tau),$$

with $\gamma_\Delta(x) = (1 - |x|/\Delta)^+$. We also have

$$V_{\Delta_T, T}(S_1, S_{2, \tau_T+}) - C_{12}(\Delta_T, \tau_T) \rightarrow 0 \text{ as } T \rightarrow \infty \text{ in } L^2(P)$$

and

$$V_{\Delta_T, T}(S_2, S_{1, \tau_T+}) - C_{12}(\Delta_T, -\tau_T) \rightarrow 0 \text{ as } T \rightarrow \infty \text{ in } L^2(P),$$

with

$$C_{12}(\Delta, \tau) = \frac{2}{1 - \|h\|_{L^1}\|g\|_{L^1}} \gamma_\Delta \star (\delta_0 + F + \check{F} + F \star \check{F}) \star (\nu_2 \check{h} + \nu_1 g)(\tau).$$

Remark 8 (The Epps effect). For $f \in L^1$ we have $\gamma_\Delta \star f \rightarrow 0$ pointwise as $\Delta \rightarrow 0$. Therefore we obtain

$$C_{12}(\Delta, \tau) \rightarrow 0 \text{ as } \Delta \rightarrow 0 \text{ for every } \tau \in \mathbb{R}$$

and this characterises the Epps effect. The same argument, together with $\gamma_\Delta(0) = 1$, yields the convergence

$$C_{11}(\Delta, \tau) \rightarrow \frac{2\nu_1}{1 - \|h\|_{L^1}\|g\|_{L^1}} \mathbf{1}_{\{\tau=0\}} \text{ as } \Delta \rightarrow 0.$$

Remark 9 (The Lead-Lag effect). Following [7] and as a consensus in the literature, we say that a lead-lag effect is present between S_1 and S_2 if there exists $\Delta > 0$ and $\tau \neq 0$ such that

$$C_{12}(\Delta, \tau) \neq C_{12}(\Delta, -\tau).$$

Therefore, an absence of any lead-lag effect is obtained if and only if the function $(\delta_0 + F + \check{F} + F \star \check{F}) \star (\nu_2 \check{h} + \nu_1 g)$ is even. This is the case if $\nu_1 g = \nu_2 h$. Now, let $\varepsilon > 0$. If $g = h \star \delta_\varepsilon$ and $\mu_1 = \mu_3$, then

$$C_{12}(\Delta, -\tau) = C_{12}(\Delta, \tau + \varepsilon) \text{ for every } \Delta > 0, \tau \in \mathbb{R}.$$

This particular choice for h and g models in particular the property that S_1 acts on S_2 in the same manner as S_2 acts on S_1 with an extra temporal shift of ε . Since we always have

$$\lim_{\tau \rightarrow \pm\infty} C_{12}(\Delta, \tau) = 0,$$

there exists τ_0 such that $C_{12}(\Delta, \tau_0 + \varepsilon) \neq C_{12}(\Delta, \tau_0)$, or in other words, we have a lead-lag effect.

Remark 10 (Macroscopic correlations). Since $\gamma_\Delta \rightarrow 1$ as $\Delta \rightarrow \infty$, we obtain the convergence

$$\begin{aligned} C_{11}(\Delta, \tau) &\rightarrow \frac{2}{1 - \|h\|_{L^1} \|g\|_{L^1}} \int (\delta_0 + F + \check{F} + F \star \check{F}) \star (\nu_1 \delta_0 + \nu_2 h \star \check{h}) \\ &= \frac{2}{1 - \|h\|_{L^1} \|g\|_{L^1}} \left(1 + 2 \int F + (\int F)^2\right) (\nu_1 + \nu_2 (\int h)^2) = \text{Var}(X_1). \end{aligned}$$

Likewise, we have $C_{12}(\Delta, \tau) \rightarrow \text{Cov}(X_1, X_2)$ as $\Delta \rightarrow \infty$.

6 Preparation for the proofs

In the sequel, we work in the setting of Sections 2 and 3 under Assumption (A1).

Lemma 2. *For all finite stopping time S one has:*

$$E(N_S) = \mu E(S) + E\left(\int_0^S \varphi(S-t) N_t dt\right) \quad (10)$$

$$E(N_S) \leq (\mathbf{Id} - \mathbf{K})^{-1} \mu E(S) \quad \text{componentwise.} \quad (11)$$

Proof. Recall that $(T_p)_{p \geq 1}$ denote the successive jump times of N and set $S_p = S \wedge T_p$. Since the stochastic intensities λ_i are given by (4) one has

$$E(N_{S_p}) = \mu E(S_p) + E\left(\int_0^{S_p} dt \int_{(0,t)} \varphi(t-s) dN_s\right).$$

Moreover, by Fubini theorem $\int_0^{S_p} dt \int_{(0,t)} \varphi(t-s) dN_s$ is equal to

$$\int_{[0, S_p]} \left(\int_s^{S_p} \varphi(t-s) dt\right) dN_s = \int_{[0, S_p]} \left(\int_0^{S_p-s} \varphi(t) dt\right) dN_s.$$

Now, integrating by part with $\Phi(t) = \int_0^t \varphi(s) ds$, we can write

$$0 = \Phi(0)N_{S_p} - \Phi(S_p)N_0 = \int_{(0, S_p]} \Phi(S_p - t) dN_t - \int_0^{S_p} \varphi(S_p - t) N_t dt.$$

Remark that both sides of the above equality are finite since $\sum_{i=1}^d N_{i, S_p} \leq p$. We obtain

$$E(N_{S_p}) = \mu E(S_p) + E\left(\int_0^{S_p} \varphi(S_p - t) N_t dt\right)$$

and derive (10) using that $N_{S_p} \uparrow N_S$ as $p \rightarrow \infty$ and

$$\int_0^{S_p} \varphi(S_p - t) N_t dt = \int_0^{S_p} \varphi(t) N_{S_p-t} dt \uparrow \int_0^S \varphi(t) N_{S-t} dt = \int_0^S \varphi(S-t) N_t dt.$$

We next prove (11). We have (componentwise)

$$\begin{aligned} E(N_{S_p}) &= E\left(S_p \mu + \int_0^{S_p} \varphi(S_p - t) N_t dt\right) \\ &\leq E(S_p) \mu + E\left(\int_0^\infty \varphi(S_p - t) N_t dt\right) = E(S_p) \mu + \mathbf{K} E(N_{S_p}). \end{aligned}$$

By induction, $E(N_{S_p}) \leq (\mathbf{Id} + \mathbf{K} + \dots + \mathbf{K}^{n-1}) E(S_p) \mu + \mathbf{K}^n E(N_{S_p})$ componentwise for all integer n . On the one hand, $\sum_{i=1}^d N_{i, S_p} \leq p$. On the other hand, since $\rho(\mathbf{K}) < 1$ we have $\mathbf{K}^n \rightarrow 0$ as $n \rightarrow \infty$ and $\sum_{n=0}^\infty \mathbf{K}^n = (\mathbf{Id} - \mathbf{K})^{-1}$, therefore $E(N_{S_p}) \leq (\mathbf{Id} - \mathbf{K})^{-1} E(S) \mu$. This readily yields (11) since $E(N_S) = \lim_p E(N_{S_p})$. \square

Let φ_n , $n \geq 1$, and $\psi = \sum_{n \geq 1} \varphi_n$ be defined as in Theorem 2. By induction it is easily shown that $\int_0^\infty \varphi_n(t) dt = \mathbf{K}^n$ for all n . Therefore $\int_0^\infty \psi(t) dt = \sum_{n \geq 1} \mathbf{K}^n$ is finite componentwise by assumption (A1). We next state a multivariate version of the well known renewal equation (we recall the proof for sake of completeness).

Lemma 3. Let h be a Borel and locally bounded function from \mathbb{R}_+ to \mathbb{R}^d . Then there exists a unique locally bounded function $f : \mathbb{R}_+ \rightarrow \mathbb{R}^d$ solution to

$$f(t) = h(t) + \int_0^t \varphi(t-s)f(s)ds \quad \forall t \geq 0, \quad (12)$$

given by

$$f_h(t) = h(t) + \int_0^t \psi(t-s)h(s)ds.$$

Proof. Since $\psi \in L^1(dt)$ and h is locally bounded, the function f_h is locally bounded. Moreover f_h satisfies (12). It follows that

$$\begin{aligned} \int_0^t \varphi(t-s)f_h(s)ds &= \int_0^t \varphi(t-s)h(s)ds + \int_0^t ds\varphi(t-s) \int_0^s \psi(s-r)h(r)dr \\ &= \int_0^t \varphi(t-s)h(s)ds + \int_0^t dr \int_r^t ds\varphi(t-s)\psi(s-r)h(r) \\ &= \int_0^t \psi(t-r)h(r)dr \end{aligned}$$

since $\int_0^t \varphi(t-s)\psi(s)ds = \psi(t) - \varphi(t)$. As for the uniqueness, if f satisfies (12) then

$$f_h(t) - f(t) = \int_0^t \varphi(t-s)(f_h(s) - f(s))ds$$

thus if $g_i(t) = |f_{h,i}(t) - f_i(t)|$, $1 \leq i \leq d$, one has $g(t) \leq \int_0^t \varphi(t-s)g(s)ds$ componentwise, which yields $\int_0^\infty g(t)dt \leq \mathbf{K} \int_0^\infty g(t)dt$ componentwise. Since $\rho(\mathbf{K}) < 1$ it follows that $f = f_h$ almost everywhere. Therefore

$$\int_0^t \varphi(t-s)f(s)ds = \int_0^t \varphi(t-s)f_h(s)ds \text{ for all } t$$

and thus $f = f_h$ since both function satisfy (12). \square

Define the d -dimensional martingale $(M_t)_{t \geq 0}$ by $M_t = N_t - \int_0^t \lambda_s ds$ with $\lambda = (\lambda_1, \dots, \lambda_d)$.

Lemma 4. For all $t \geq 0$:

$$E(N_t) = t\mu + \left(\int_0^t \psi(t-s)s ds \right) \mu, \quad (13)$$

$$N_t - E(N_t) = M_t + \int_0^t \psi(t-s)M_s ds. \quad (14)$$

Proof. By (10) of Lemma 2 and Fubini theorem, we get

$$E(N_t) = t\mu + \int_0^t \varphi(t-s)E(N_s) ds.$$

Besides, $t \rightsquigarrow E(N_t)$ is locally bounded in view of (11). Applying Lemma 3 we obtain (13). The second formula follows from Lemma 3 and the fact that, if $X_t = N_t - E(N_t)$, representation (13) entails $X_t = M_t + \int_0^t \varphi(t-s)X_s ds$. \square

7 Proof of Theorem 1

Lemma 5. *Let $p \in [0, 1)$ and assume that $\int_0^\infty t^p \varphi(t) dt < \infty$ componentwise. Then*

1. *If $p < 1$, we have $T^p \left(T^{-1} E(N_{Tv}) - v(\mathbf{Id} - \mathbf{K})^{-1} \mu \right) \rightarrow 0$ as $T \rightarrow \infty$ uniformly in $v \in [0, 1]$.*
2. *If $p = 1$, we have (as $T \rightarrow \infty$)*

$$T \left(\frac{1}{T} E(N_T) - (\mathbf{Id} - \mathbf{K})^{-1} \mu \right) \rightarrow -(\mathbf{Id} - \mathbf{K})^{-1} \left(\int_0^\infty t \varphi(t) dt \right) (\mathbf{Id} - \mathbf{K})^{-1} \mu.$$

Proof. Let $p \in [0, 1]$ and assume that $\int_0^\infty t^p \varphi(t) dt < \infty$ componentwise. We first prove that $\int_0^\infty t^p \psi(t) dt < \infty$ componentwise. For $n \geq 1$, setting $\mathbf{A}_n = \int_0^\infty t^p \varphi_n(t) dt$, we can write

$$\begin{aligned} \mathbf{A}_{n+1} &= \int_0^\infty t^p \left(\int_0^t \varphi(t-s) \varphi_n(s) ds \right) dt = \int_0^\infty \left(\int_0^\infty (t+s)^p \varphi(t) dt \right) \varphi_n(s) ds \\ &\leq \int_0^\infty t^p \varphi(t) dt \mathbf{K}^n + \mathbf{K} \int_0^\infty s^p \varphi_n(s) ds \quad \text{with equality if } p = 1, \\ &= \mathbf{A}_1 \mathbf{K}^n + \mathbf{K} \mathbf{A}_n. \end{aligned}$$

Therefore for all integer N , we have

$$\begin{aligned} \sum_{n=1}^N \mathbf{A}_n &\leq \mathbf{A}_1 + \mathbf{A}_1 \sum_{n=1}^{N-1} \mathbf{K}^n + \mathbf{K} \sum_{n=1}^{N-1} \mathbf{A}_n, \\ (\mathbf{Id} - \mathbf{K}) \sum_{n=1}^{N-1} \mathbf{A}_n + \mathbf{A}_N &\leq \mathbf{A}_1 + \mathbf{A}_1 \sum_{n=1}^{N-1} \mathbf{K}^n, \\ \sum_{n=1}^{N-1} \mathbf{A}_n &\leq (\mathbf{Id} - \mathbf{K})^{-1} (\mathbf{A}_1 + \mathbf{A}_1 \sum_{n=1}^{N-1} \mathbf{K}^n). \end{aligned}$$

Letting $N \rightarrow \infty$ we derive

$$\int_0^\infty t^p \boldsymbol{\psi}(t) dt = \sum_{n \geq 1} \mathbf{A}_n \leq (\mathbf{Id} - \mathbf{K})^{-1} \mathbf{A}_1 (\mathbf{Id} - \mathbf{K})^{-1}$$

with equality if $p = 1$. From (13) it follows that for all $v \in [0, 1]$:

$$v(\mathbf{Id} - \mathbf{K})^{-1} \boldsymbol{\mu} - \frac{1}{T} E(N_{Tv}) = \left(v \int_{Tv}^\infty \boldsymbol{\psi}(t) dt + \frac{1}{T} \int_0^{Tv} t \boldsymbol{\psi}(t) dt \right) \boldsymbol{\mu}. \quad (15)$$

Since $t^p \boldsymbol{\psi}(t)$ is integrable, we have $T^p \int_{Tv}^\infty \boldsymbol{\psi}(t) dt \leq v^{1-p} \int_{Tv}^\infty t^p \boldsymbol{\psi}(t) dt \rightarrow 0$ as $T \rightarrow \infty$ and this convergence is uniform in $v \in [0, 1]$ in the case $p < 1$. It remains to prove that if $p < 1$, we have

$$\frac{1}{T^{1-p}} \int_0^T t \boldsymbol{\psi}(t) dt \rightarrow 0 \text{ as } T \rightarrow \infty.$$

With $\mathbf{G}(t) = \int_0^t s^p \boldsymbol{\psi}(s) ds$, integrating by part, we obtain

$$T^{1-p} \mathbf{G}(T) = \int_0^T t \boldsymbol{\psi}(t) dt + (1-p) \int_0^T t^{-p} \mathbf{G}(t) dt$$

and

$$\frac{1}{T^{1-p}} \int_0^T \boldsymbol{\psi}(t) t dt = \mathbf{G}(T) - \frac{1-p}{T^{1-p}} \int_0^T t^{-p} \mathbf{G}(t) dt.$$

Since $\mathbf{G}(t)$ is convergent as $t \rightarrow \infty$ we finally derive that the right-hand-side in the above equality converges to 0 as $T \rightarrow \infty$. \square

Denote by $\|\cdot\|$ the Euclidean norm either on \mathbb{R}^d or on the set of $d \times d$ matrices.

Lemma 6. *There exists a constant $C_{\mu, \varphi}$ such that for all $t, \Delta \geq 0$:*

$$E\left(\sup_{t \leq s \leq t+\Delta} \|M_s - M_t\|^2\right) \leq C_{\mu, \varphi} \Delta.$$

Proof. Doob's inequality yields

$$E\left(\sup_{t \leq s \leq t+\Delta} \|M_s - M_t\|^2\right) \leq 4 \sum_{i=1}^d E((M_{i,t+\Delta} - M_{i,t})^2).$$

For each $i \in \{1, \dots, d\}$, the quadratic variation of the martingale $(M_{i,t})_{t \geq 0}$ is

$$[M_i, M_i]_t = \sum_{s \leq t} (M_{i,s} - M_{i,s-})^2 = N_{i,t}.$$

Thus we have

$$E((M_{i,t+\Delta} - M_{i,t})^2) = E(N_{i,t+\Delta} - N_{i,t}).$$

Besides, in view of Lemma 4 and the fact that $\int_0^\infty \psi(t)dt = (\mathbf{Id} - \mathbf{K})^{-1} - \mathbf{Id}$, we obtain

$$E(N_{t+\Delta} - N_t) \leq \Delta(\mathbf{Id} - \mathbf{K})^{-1}\mu \quad \text{componentwise.}$$

□

Completion of proof of Theorem 1. Lemma 5 with $p = 0$ implies that it is enough to prove the following convergence

$$T^{-1} \sup_{v \in [0,1]} \|N_{Tv} - E(N_{Tv})\| \rightarrow 0 \quad \text{as } T \rightarrow \infty \quad (16)$$

almost surely and in $L^2(P)$. Thanks to (14) of Lemma 4, we have

$$\sup_{v \in [0,1]} \|N_{Tv} - E(N_{Tv})\| \leq \left(1 + \int_0^T \|\psi(t)\| dt\right) \sup_{t \leq T} \|M_t\| \leq C_\varphi \sup_{t \leq T} \|M_t\|$$

since ψ is integrable. Moreover $E(\sup_{t \leq T} \|M_t\|^2) \leq C_{\mu,\varphi}T$ by Lemma 6, therefore convergence (16) holds in $L^2(P)$. In order to prove the almost-sure convergence, it is enough to show that

$$T^{-1} \sup_{v \in [0,1]} \|M_{Tv}\| \rightarrow 0 \quad \text{as } T \rightarrow \infty \quad \text{almost-surely.}$$

Let $Z_t = (Z_{1,t}, \dots, Z_{d,t}) = \int_{(0,t]} \frac{1}{s+1} dM_s$. The quadratic variation of the martingale Z_i satisfies

$$[Z_i, Z_i]_t = \sum_{0 < s \leq t} (Z_{i,s} - Z_{i,s-})^2 = \int_{(0,t]} \frac{1}{(s+1)^2} dN_{i,s}$$

and moreover, using integration by part and (13), we have

$$E\left(\int_{(0,\infty)} \frac{1}{(s+1)^2} dN_s\right) = 2E\left(\int_{(0,\infty)} \frac{N_s}{(1+s)^3} ds\right) < \infty.$$

Therefore $\lim_{t \rightarrow \infty} Z_t$ exists and is finite almost surely. It follows that

$$\frac{1}{t+1} M_t = Z_t - \frac{1}{t+1} \int_0^t Z_s ds \rightarrow 0 \quad \text{as } T \rightarrow \infty \quad \text{almost surely.}$$

We deduce that almost surely, for all family $v_T \in [0, 1]$, $T > 0$, such that $Tv_T \rightarrow \infty$ the convergence $M_{Tv_T}/T \rightarrow 0$ holds. Moreover, we have $M_{Tv_T}/T \rightarrow 0$ if $\sup_T Tv_T < \infty$. In other words $T^{-1}M_{Tv} \rightarrow 0$ uniformly in $v \in [0, 1]$, almost-surely. The proof of Theorem 1 is complete. □

8 Proof of Theorem 2

Let $W = (W_1, \dots, W_d)$ be a standard d -dimensional Brownian motion. For $i = 1, \dots, d$, put $\sigma_i = (\Sigma_{ii})^{1/2}$.

Lemma 7. *The martingales $M^{(T)} := (T^{-1/2}M_{Tv})_{v \in [0,1]}$ converge in law for the Skorokod topology to $(\sigma_1 W_1, \dots, \sigma_d W_d)$.*

Proof. According to Theorem VIII-3.11 of [23], since the martingales $M^{(T)}$ have uniformly bounded jumps, a necessary and sufficient condition to obtain the lemma is: for all $v \in [0, 1]$, for all $1 \leq i < j \leq d$

$$[M_i^{(T)}, M_i^{(T)}]_v \rightarrow \sigma_i^2 v, \quad [M_i^{(T)}, M_j^{(T)}]_v \rightarrow 0, \quad \text{as } T \rightarrow \infty \text{ in probability.}$$

We have $[M_i^{(T)}, M_i^{(T)}]_v = \frac{1}{T} N_{i,Tv} \rightarrow \sigma_i^2 v$ in $L^2(P)$ by Theorem 1 and also $[M_i^{(T)}, M_j^{(T)}]_v = 0$ since the processes N_i for $1 \leq i \leq d$, have no common jump by construction. \square

Completion of proof of Theorem 2. Set $X_v^{(T)} = T^{-1/2}(N_{Tv} - E(N_{Tv}))$. In view of Lemma 7, it is enough to prove that

$$\sup_{v \in [0,1]} \|X_v^{(T)} - (\mathbf{Id} - \mathbf{K})^{-1} M_v^{(T)}\| \rightarrow 0 \text{ as } T \rightarrow \infty \text{ in probability.}$$

By Lemma 4, we have $X_v^{(T)} = M_v^{(T)} + \int_0^v T \psi(Tu) M_{v-u}^{(T)} du$, hence we need to prove that

$$\sup_{v \in [0,1]} \left\| \int_0^v T \psi(Tu) M_{v-u}^{(T)} du - \left(\int_0^\infty \psi(t) dt \right) M_v^{(T)} \right\| \rightarrow 0 \text{ as } T \rightarrow \infty \quad (17)$$

in probability. We plan to use the fact that ψ is integrable and the C -tightness of the family $(M^{(T)})_{T>0}$. The tightness is a consequence of Lemma 7 and reads:

$$\forall \varepsilon > 0 \quad \limsup_T P \left(\sup_{|u-u'| \leq \eta} \|M_u^{(T)} - M_{u'}^{(T)}\| > \varepsilon \right) \rightarrow 0 \text{ as } \eta \rightarrow 0. \quad (18)$$

using also that $M_0^{(T)} = 0$. For $\eta > 0$ and $v \in [0, 1]$ we have

$$\begin{aligned} \left\| \int_{v \wedge \eta}^v T \psi(Tu) M_{v-u}^{(T)} du \right\| &\leq \sup_{0 \leq u \leq 1} \|M_u^{(T)}\| \int_\eta^1 T \|\psi(Tu)\| du \\ &\leq \sup_{0 \leq u \leq 1} \|M_u^{(T)}\| \int_{T\eta}^\infty \|\psi(t)\| dt \rightarrow 0 \end{aligned}$$

as $T \rightarrow \infty$ in probability, since $\sup_{0 \leq u \leq 1} \|M_u^{(T)}\|$ is bounded in probability and $\int_{T\eta}^{\infty} \|\psi(t)\| dt \rightarrow 0$ as $T \rightarrow \infty$. Moreover

$$\left\| \int_0^{v \wedge \eta} T\psi(Tu)(M_v^{(T)} - M_{v-u}^{(T)}) du \right\| \leq \sup_{|u-u'| \leq \eta} \|M_u^{(T)} - M_{u'}^{(T)}\| \int_0^{\infty} \|\psi(t)\| dt,$$

therefore, in order to prove (17) it suffices to show that for all $\eta > 0$, the convergence

$$\sup_{v \in [0,1]} \left\| \left(\int_0^{\infty} \psi(t) dt - \int_0^{v \wedge \eta} T\psi(Tu) du \right) M_v^{(T)} \right\| \rightarrow 0 \text{ as } T \rightarrow \infty$$

holds in probability. It readily follows from (18) and the fact that $\left\| \left(\int_0^{\infty} \psi(t) dt - \int_0^{v \wedge \eta} T\psi(Tu) du \right) M_v^{(T)} \right\|$ is less than $\int_{T\delta}^{\infty} \|\psi(t)\| dt \sup_u \|M_u^{(T)}\|$ if $v > \delta$, and less than $\int_0^{\infty} \|\psi(t)\| dt \sup_{u \leq \delta} \|M_u^{(T)}\|$ if $v \leq \delta$, with $0 < \delta < \eta$. \square

Proof of Corollary 1. By (A2), Lemma 5 with $p = 1/2$ yields

$$T^{1/2}(T^{-1}E(N_{Tv}) - v(\mathbf{Id} - \mathbf{K})^{-1}\mu) \rightarrow 0 \text{ as } T \rightarrow \infty$$

uniformly in $v \in [0, 1]$. Moreover, by (A1), Theorem 2 yields

$$T^{1/2}(T^{-1}N_{Tv} - T^{-1}E(N_{Tv})) \rightarrow (\mathbf{Id} - \mathbf{K})^{-1}\Sigma^{1/2}W$$

in distribution as $T \rightarrow \infty$ and the result follows. \square

9 Proof of Theorem 3

Set

$$Y_t = \int_0^t \psi(t-s)M_s ds \tag{19}$$

in order that $X = M + Y$, see Lemma 4. For all $0 \leq \varepsilon \leq \eta \leq 1$, for all integer $1 \leq k_0 \leq T/\Delta$, define

$$\mathcal{D}_{k_0, \varepsilon, \eta}(X)_{T, \Delta} = \frac{1}{T} \sum_{k=k_0}^{\lfloor T/\Delta \rfloor} (X_{(k-1)\Delta + \eta\Delta} - X_{(k-1)\Delta + \varepsilon\Delta}).$$

Lemma 8. *There exists a function $T \rightsquigarrow \xi_{\mu, \varphi}(T)$ such that $\xi_{\mu, \varphi}(T) \rightarrow 0$ as $T \rightarrow \infty$ and such that for all $0 \leq \varepsilon \leq \eta \leq 1$ and all integer $1 \leq k_0 \leq T/\Delta$, we have*

$$E\left(\left\| \left(1 - \frac{k_0\Delta}{T}\right)(\eta - \varepsilon)(\mathbf{Id} - \mathbf{K})^{-1}\mu - \mathcal{D}_{k_0, \varepsilon, \eta}(N)_{T, \Delta} \right\|^2\right) \leq \xi_{\mu, \varphi}(T).$$

Proof. First we prove that

$$\mathbb{E}(\|\mathcal{D}_{k_0, \varepsilon, \eta}(X)_{T, \Delta}\|^2) \leq C_{\mu, \varphi} T^{-1} \quad (20)$$

for some constant $C_{\mu, \varphi}$ that depends on μ and φ only. Using

$$\begin{aligned} X_{(k-1)\Delta + \eta\Delta} - X_{(k-1)\Delta + \varepsilon\Delta} &= M_{(k-1)\Delta + \eta\Delta} - M_{(k-1)\Delta + \varepsilon\Delta} \\ &\quad + \int_0^\infty ds \psi(s) (M_{(k-1)\Delta + \eta\Delta - s} - M_{(k-1)\Delta + \varepsilon\Delta - s}) \end{aligned}$$

and the fact that ψ is integrable, it suffices to prove that

$$E(\|\mathcal{D}_{k_0, \varepsilon, \eta}(M_{\cdot - s})_{T, \Delta}\|^2) \leq C_{\mu, \varphi} T^{-1}.$$

Set $S_n = \sum_{k=1}^n (M_{(k-1)\Delta + \eta\Delta - s} - M_{(k-1)\Delta + \varepsilon\Delta - s})$. Clearly $(S_n)_{n \geq 1}$ is a discrete martingale thus

$$E(\|S_{\lfloor T\Delta \rfloor} - S_{k_0-1}\|^2) = \sum_{k=k_0}^{\lfloor T\Delta \rfloor} E(\|M_{(k-1)\Delta + \eta\Delta - s} - M_{(k-1)\Delta + \varepsilon\Delta - s}\|^2).$$

By Lemma 6 we have $E(\|S_{\lfloor T\Delta \rfloor}\|^2) \leq C_{\mu, \varphi}(\eta - \varepsilon)T$ and (20) follows. It remains to prove that

$$\left\| \left(1 - \frac{k_0\Delta}{T}\right)(\eta - \varepsilon)(\mathbf{Id} - \mathbf{K})^{-1}\mu - E(\mathcal{D}_{k_0, \varepsilon, \eta}(N)_{T, \Delta}) \right\| \leq \tilde{\xi}_{\mu, \varphi}(T)$$

where $\tilde{\xi}_{\mu, \varphi}(T) \rightarrow 0$ as $T \rightarrow \infty$. By Lemma 4, we have

$$\begin{aligned} &E(N_{(k-1)\Delta + \eta\Delta} - N_{(k-1)\Delta + \varepsilon\Delta}) \\ &= (\eta - \varepsilon)\Delta \left(\mu + \int_0^{(k-1)\Delta + \varepsilon\Delta} dr \psi(r) \mu \right) \\ &\quad + \int_{(k-1)\Delta + \varepsilon\Delta}^{(k-1)\Delta + \eta\Delta} dr \psi(r) ((k-1)\Delta + \eta\Delta - r) \mu \\ &= (\eta - \varepsilon)\Delta (\mathbf{Id} - \mathbf{K})^{-1}\mu - (\eta - \varepsilon)\Delta \int_{(k-1)\Delta + \varepsilon\Delta}^\infty dr \psi(r) \\ &\quad + \int_{(k-1)\Delta + \varepsilon\Delta}^{(k-1)\Delta + \eta\Delta} dr \psi(r) ((k-1)\Delta + \eta\Delta - r) \mu. \end{aligned}$$

Finally

$$\int_{(k-1)\Delta + \varepsilon\Delta}^{(k-1)\Delta + \eta\Delta} dr \psi(r) ((k-1)\Delta + \eta\Delta - r) \leq (\eta - \varepsilon)\Delta \int_{(k-1)\Delta + \varepsilon\Delta}^\infty dr \psi(r).$$

We conclude noting that

$$\begin{aligned} \frac{\Delta}{T} \sum_{k=k_0}^{\lfloor T/\Delta \rfloor} \int_{(k-1)\Delta+\varepsilon\Delta}^{\infty} dr \boldsymbol{\psi}(r) &= \frac{\Delta}{T} \int_{(k_0-1)\Delta+\varepsilon\Delta}^{\infty} dr \boldsymbol{\psi}(r) \sum_{k=k_0}^{\lfloor T/\Delta \rfloor} \mathbf{1}_{\{(k-1)\Delta+\varepsilon\Delta < r\}} \\ &\leq \frac{1}{T} \int_0^{\infty} dr \boldsymbol{\psi}(r)(r \wedge T), \end{aligned}$$

This last quantity converges to 0 as $T \rightarrow \infty$ by an argument similar to the end of proof of Lemma, 5, using that $\boldsymbol{\psi}$ is integrable. \square

Lemma 9. *There exists a constant $C_{\mu,\varphi}$ such that for all $t, h \geq 0$, we have*

$$E\left(\sup_{t \leq s \leq t+h} \|M_s - M_t\|^4\right) \leq C_{\mu,\varphi} (h + h^2)$$

Proof. By BDG-inequality, $E(\sup_{t \leq s \leq t+h} \|M_s - M_t\|^4 | \mathcal{F}_t)$ is less than

$$C \sum_{i=1}^d E\left(\left([M_i, M_i]_{t+h} - [M_i, M_i]_t\right)^2 | \mathcal{F}_t\right) = C \sum_{i=1}^d E\left(\left(N_{i,t+h} - N_{i,t}\right)^2 | \mathcal{F}_t\right),$$

hence $E(\sup_{t \leq s \leq t+h} \|M_s - M_t\|^4) \leq C(E(\|X_{t+h} - X_t\|^2) + \|E(N_{t+h} - N_t)\|^2)$. By Lemma 4, we have $E(N_{t+h} - N_t) \leq h(\mathbf{Id} - \mathbf{K})^{-1}\mu$ componentwise and

$$\left(E(\|X_{t+h} - X_t\|^2)\right)^{1/2} \leq \int_0^{\infty} \|\boldsymbol{\psi}(s)\| \left(E(\|M_{t+h-s} - M_t\|^2)\right)^{1/2} ds.$$

The conclusion follows using Lemma 6 and the fact that $\boldsymbol{\psi}$ is integrable. \square

With the notation introduced in Section 3, the quantity $\mathbf{V}_{\Delta,T}(M_{\cdot-s}, M_{\cdot-t+\tau})$ is equal to

$$\frac{1}{T} \sum_{k=1}^{\lfloor T/\Delta \rfloor} (M_{k\Delta-s} - M_{(k-1)\Delta-s})(M_{k\Delta-t+\tau} - M_{(k-1)\Delta-t+\tau})^\top.$$

Lemma 10. *For all $s, t \geq 0$ we have*

$$\mathbf{V}_{\Delta,T}(M_{\cdot-s}, M_{\cdot-t+\tau_T}) - \left(1 - \frac{|t-s-\tau_T|}{\Delta_T}\right)^+ \boldsymbol{\Sigma} \rightarrow 0 \text{ as } T \rightarrow \infty \text{ in } L^2(P)$$

and

$$E(\|\mathbf{V}_{\Delta,T}(M_{\cdot-s}, M_{\cdot-t+\tau_T})\|^2) \leq C_{\mu,\varphi}.$$

Proof. We start with proving preliminary estimates. Let b_1, b_2, b'_1, b'_2 be real numbers such that $b_1 \leq b_2, b'_1 \leq b'_2$ and $(b_1, b_2) \cap (b'_1, b'_2) = \emptyset$. Using that M is a martingale, we successively obtain

$$\begin{aligned}
& E\left(\left\|T^{-1} \sum_{k=1}^{\lfloor T/\Delta \rfloor} (M_{(k-1)\Delta+b_2} - M_{(k-1)\Delta+b_1})(M_{(k-1)\Delta+b'_2} - M_{(k-1)\Delta+b'_1})^\top\right\|^2\right) \\
&= \frac{1}{T^2} \sum_{k=1}^{\lfloor T/\Delta \rfloor} E\left(\left\|(M_{(k-1)\Delta+b_2} - M_{(k-1)\Delta+b_1})(M_{(k-1)\Delta+b'_2} - M_{(k-1)\Delta+b'_1})^\top\right\|^2\right) \\
&= \frac{1}{T^2} \sum_{k=1}^{\lfloor T/\Delta \rfloor} E\left(\|M_{(k-1)\Delta+b_2} - M_{(k-1)\Delta+b_1}\|^2 \|(M_{(k-1)\Delta+b'_2} - M_{(k-1)\Delta+b'_1})^\top\|^2\right).
\end{aligned}$$

By Cauchy-Schwarz, this last quantity is less than

$$\frac{1}{T^2} \sum_{k=1}^{\lfloor T/\Delta \rfloor} \left(E(\|M_{(k-1)\Delta+b_2} - M_{(k-1)\Delta+b_1}\|^4) E(\|(M_{(k-1)\Delta+b'_2} - M_{(k-1)\Delta+b'_1})^\top\|^4)\right)^{\frac{1}{2}}$$

which in turn, using Lemma 9, is bounded by

$$C_{\mu, \varphi} \frac{1}{T\Delta} (b_2 - b_1 + (b_2 - b_1)^2)^{1/2} (b'_2 - b'_1 + (b'_2 - b'_1)^2)^{1/2}. \quad (21)$$

Moreover, if $b_2 - b_1 \leq \Delta$, using that $[M, M] = \mathbf{diag}(N)$ and Lemma 9, we obtain that

$$\begin{aligned}
& E\left(\left\|\frac{1}{T} \sum_{k=1}^{\lfloor T/\Delta \rfloor} \mathbf{diag}(N_{(k-1)\Delta+b_2} - N_{(k-1)\Delta+b_1})\right.\right. \\
& \left. - \frac{1}{T} \sum_{k=1}^{\lfloor T/\Delta \rfloor} (M_{(k-1)\Delta+b_2} - M_{(k-1)\Delta+b_1})(M_{(k-1)\Delta+b_2} - M_{(k-1)\Delta+b_1})^\top\right\|^2\right)
\end{aligned}$$

is less than

$$C_{\mu, \varphi} \frac{1}{T\Delta} (b_2 - b_1 + (b_2 - b_1)^2). \quad (22)$$

We are ready to prove Lemma 10. It is a consequence of Lemma 8 and the fact that there exists $a_2 = a_2(s, t, \Delta_T, \tau_T) \leq a_3 = a_3(s, t, \Delta_T, \tau_T)$ such that

$$a_3 - a_2 = (\Delta_T - |t - s - \tau_T|)^+, \quad \frac{a_2}{T} \rightarrow 0 \text{ as } T \rightarrow \infty$$

holds, together with the estimate

$$\begin{aligned}
& E\left(\left\|\frac{1}{T} \sum_{k=1}^{\lfloor T/\Delta_T \rfloor} \mathbf{diag}(N_{(k-1)\Delta_T+a_3} - N_{(k-1)\Delta_T+a_2}) - \mathbf{V}_{\Delta_T, T}(M_{\cdot-s}, M_{\cdot-t+\tau_T})\right\|^2\right) \\
& \leq C_{\mu, \varphi} \frac{1}{T} (1 + \Delta_T).
\end{aligned} \tag{23}$$

Indeed, if $t - \tau_T + \Delta_T \leq s$ or $s \leq t - \tau_T - \Delta_T$ then the upper bound we obtained in (21) entails

$$E\left(\left\|\mathbf{V}_{\Delta_T, T}(M_{\cdot-s}, M_{\cdot-t+\tau_T})\right\|^2\right) \leq C_{\mu, \varphi} \frac{1}{T} (1 + \Delta_T).$$

Let us first consider the case $t - \tau_T \leq s \leq t - \tau_T + \Delta_T$. Set

$$a_1 := s \leq a_2 := -t + \tau_T \leq a_3 := \Delta_T - s \leq a_4 := \Delta_T - t + \tau_T.$$

We use the following decomposition

$$\begin{aligned}
& (M_{k\Delta_T-s} - M_{(k-1)\Delta_T-s})(M_{k\Delta_T-t+\tau_T} - M_{(k-1)\Delta_T-t+\tau_T})^\top \\
& = (M_{(k-1)\Delta_T+a_3} - M_{(k-1)\Delta_T+a_2})(M_{(k-1)\Delta_T+a_3} - M_{(k-1)\Delta_T+a_2})^\top \\
& \quad + (M_{(k-1)\Delta_T+a_3} - M_{(k-1)\Delta_T+a_1})(M_{(k-1)\Delta_T+a_4} - M_{(k-1)\Delta_T+a_3})^\top \\
& \quad + (M_{(k-1)\Delta_T+a_2} - M_{(k-1)\Delta_T+a_1})(M_{(k-1)\Delta_T+a_3} - M_{(k-1)\Delta_T+a_2})^\top.
\end{aligned}$$

On the one hand, (21) readily yields that both

$$E\left(\left\|\frac{1}{T} \sum_{k=1}^{\lfloor T/\Delta_T \rfloor} (M_{(k-1)\Delta_T+a_3} - M_{(k-1)\Delta_T+a_1})(M_{(k-1)\Delta_T+a_4} - M_{(k-1)\Delta_T+a_3})^\top\right\|^2\right)$$

and

$$E\left(\left\|\frac{1}{T} \sum_{k=1}^{\lfloor T/\Delta_T \rfloor} (M_{(k-1)\Delta_T+a_2} - M_{(k-1)\Delta_T+a_1})(M_{(k-1)\Delta_T+a_3} - M_{(k-1)\Delta_T+a_2})^\top\right\|^2\right)$$

are less than $C_{\mu, \varphi}(1 + \Delta_T)/T$. On the other hand, by (22), the same estimate holds for

$$\begin{aligned}
& E\left(\left\|\frac{1}{T} \sum_{k=1}^{\lfloor T/\Delta_T \rfloor} \mathbf{diag}(N_{(k-1)\Delta_T+a_3} - N_{(k-1)\Delta_T+a_2})\right.\right. \\
& \left. - \frac{1}{T} \sum_{k=1}^{\lfloor T/\Delta_T \rfloor} (M_{(k-1)\Delta_T+a_3} - M_{(k-1)\Delta_T+a_2})(M_{(k-1)\Delta_T+a_3} - M_{(k-1)\Delta_T+a_2})^\top\right\|^2\right),
\end{aligned}$$

therefore (23) holds in that case. If now $t - \tau_T - \Delta_T \leq s \leq t - \tau_T$, setting

$$a_1 := -t + \tau_T \leq a_2 := -s \leq a_3 := \Delta_T - t + \tau_T \leq a_4 := \Delta_T - s,$$

one readily checks that (23) holds using the same arguments. \square

Completion of proof of Theorem 3. Since $X = M + Y$ by (19), we have the following decomposition

$$\begin{aligned} & \mathbf{V}_{\Delta,T}(X, X_{\tau+}) \\ &= \mathbf{V}_{\Delta,T}(M, M_{\tau+}) + \mathbf{V}_{\Delta,T}(Y, Y_{\tau+}) + \mathbf{V}_{\Delta,T}(M, Y_{\tau+}) + \mathbf{V}_{\Delta,T}(Y, M_{\tau+}). \end{aligned}$$

Setting $M_t = 0$ for $t \leq 0$, we can write $Y_t = \int_0^\infty \psi(s) M_{t-s} ds$, therefore

$$\begin{aligned} & (Y_{i\Delta} - Y_{(i-1)\Delta})(Y_{i\Delta+\tau} - Y_{(i-1)\Delta+\tau})^\top \\ &= \int_{\mathbb{R}_+^2} ds dt \psi(s) (M_{i\Delta-s} - M_{(i-1)\Delta-s}) (M_{i\Delta-t+\tau} - M_{(i-1)\Delta-t+\tau})^\top \psi(t)^\top, \end{aligned}$$

hence

$$\mathbf{V}_{\Delta,T}(Y, Y_{\tau+}) = \int_{\mathbb{R}_+^2} ds dt \psi(s) \mathbf{V}_{\Delta,T}(M_{-s}, M_{-t+\tau}) \psi(t)^\top.$$

Likewise

$$\begin{aligned} \mathbf{V}_{\Delta,T}(M, Y_{\tau+}) &= \int_0^\infty dt \mathbf{V}_{\Delta,T}(M, M_{-t+\tau}) \psi(t)^\top, \\ \mathbf{V}_{\Delta,T}(Y, M_{\tau+}) &= \int_0^\infty dt \psi(t) \mathbf{V}_{\Delta,T}(M_{-t}, M_{+\tau}). \end{aligned}$$

In view of Lemma 10 and the fact that ψ is integrable, by Lebesgue dominated convergence theorem, we successively obtain

$$\mathbf{V}_{\Delta_T,T}(Y, Y_{\tau_T+}) - \int_{\mathbb{R}_+^2} ds dt \psi(s) (1 - |t - s - \tau_T|/\Delta_T)^+ \Sigma \psi(t)^\top \rightarrow 0,$$

$$\mathbf{V}_{\Delta_T,T}(M, Y_{\tau_T+}) - \int_0^\infty dt (1 - |t - \tau_T|/\Delta_T)^+ \Sigma \psi(t)^\top \rightarrow 0,$$

and

$$\mathbf{V}_{\Delta_T,T}(Y, M_{\tau_T+}) - \int_0^\infty dt \psi(t) (1 - |t + \tau_T|/\Delta_T)^+ \Sigma \rightarrow 0$$

in $L^2(P)$ as $T \rightarrow \infty$. The proof is complete. \square

Proof of Corollary 2. In view of Theorem 3, we need to show

$$\mathbf{V}_{\Delta_T, T}(X, X_{\tau T+}) - \mathbf{V}_{\Delta_T, T}(\tilde{X}, \tilde{X}_{\tau T+}) \rightarrow 0 \text{ as } T \rightarrow \infty \text{ in } L^2(P).$$

By Cauchy-Schwarz inequality, this convergence is a consequence of

$$\sup_{\tau \in \mathbb{R}} \frac{1}{T} \sum_{k=1}^{\lfloor T/\Delta_T \rfloor} \left(\eta_i(\tau + k\Delta_T) - \eta_i(\tau + (k-1)\Delta_T) \right)^2 \rightarrow 0, \quad 1 \leq i \leq d,$$

where $\eta(t) = (t(\mathbf{Id} - \mathbf{K})^{-1}\mu - E(N_t)) \mathbf{1}_{\{t \geq 0\}}$. By (15), for $t \geq 0$, we have the decomposition

$$\eta(t) = \left(t \int_t^\infty \boldsymbol{\psi}(s) ds + \int_0^t \boldsymbol{\psi}(s) s ds \right) \mu$$

therefore η is absolutely continuous and we have $\eta'(t) = \int_t^\infty \boldsymbol{\psi}(s) ds \mu \mathbf{1}_{\{t \geq 0\}}$. We derive

$$\frac{1}{T} \sum_{k=1}^{\lfloor T/\Delta_T \rfloor} \left(\eta_i(\tau + k\Delta_T) - \eta_i(\tau + (k-1)\Delta_T) \right)^2 \leq \frac{\Delta_T}{T} \int_0^\infty \eta'_i(s)^2 ds.$$

It remains to prove $\int_0^\infty \|\eta'(t)\|^2 dt < \infty$. We have

$$\eta'(t) \leq t^{-1/2} \int_t^\infty s^{1/2} \boldsymbol{\psi}(s) ds \mu \text{ and } \|\eta'(t)\| \leq C_{\mu, \varphi} t^{-1/2}$$

since $s \rightsquigarrow s^{1/2} \boldsymbol{\psi}(s)$ is integrable. Finally $\int_0^\infty \|\eta'(t)\|^2 dt$ is less than

$$C_{\mu, \varphi} \int_0^\infty dt t^{-1/2} \int_t^\infty ds \|\boldsymbol{\psi}(s)\| = 2 C_{\mu, \varphi} \int_0^\infty ds \|\boldsymbol{\psi}(s)\| s^{1/2} < \infty$$

and the result follows. \square

Appendix

Proof of Proposition 2

The spectral radius of $\mathbf{K} = \begin{pmatrix} 0 & 0 & fh & 0 \\ 0 & 0 & 0 & fh \\ fg & 0 & 0 & 0 \\ 0 & fg & 0 & 0 \end{pmatrix}$ is equal to $fgfh$, and we have

$$(\mathbf{Id} - \mathbf{K})^{-1} = \frac{1}{1 - fhfg} \begin{pmatrix} 1 & 0 & fh & 0 \\ 0 & 1 & 0 & fh \\ fg & 0 & 1 & 0 \\ 0 & fg & 0 & 1 \end{pmatrix},$$

therefore $(\mathbf{Id} - \mathbf{K})^{-1}\mu = \frac{1}{1-fhg} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_2 \end{pmatrix}$ where ν_1 and ν_2 are given by (9). Set $X = N - E(N)$. By symmetry, we have $N = (N_2, N_1, N_4, N_3)$ in distribution, thus $E(N_{1,t}) = E(N_{2,t})$ and $E(N_{3,t}) = E(N_{4,t})$ for all t . Consequently

$$S_1 = X_1 - X_2 \quad \text{and} \quad S_2 = X_3 - X_4. \quad (24)$$

According to Theorem 2 the processes $T^{-1/2}X_{Tv}$ converge in law to the process $Y_v = (\mathbf{Id} - \mathbf{K})^{-1}\Sigma^{1/2}W_v$ with $\Sigma = \frac{1}{1-fhg} \begin{pmatrix} \nu_1 & 0 & 0 & 0 \\ 0 & \nu_1 & 0 & 0 \\ 0 & 0 & \nu_2 & 0 \\ 0 & 0 & 0 & \nu_2 \end{pmatrix}$. Therefore the processes $T^{-1/2}(S_{1,Tv}, S_{2,Tv})_{v \in [0,1]}$ converge in distribution to $(Y_1 - Y_2, Y_3 - Y_4)$ and Proposition 2 is proved.

Proof of Proposition 3

From (24) it follows that $V_{\Delta_T, T}(S_1, S_{1, \tau_T+}) = (1 \ -1 \ 0 \ 0) V_{\Delta_T, T}(X, X_{\tau_T+}) \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$ and $V_{\Delta_T, T}(S_1, S_{2, \tau_T+}) = (1 \ -1 \ 0 \ 0) V_{\Delta_T, T}(X, X_{\tau_T+}) \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$. Consequently Proposition 3 follows from Theorem 3 with

$$C_{11}(\Delta, \tau) = (1 \ -1 \ 0 \ 0) \mathbf{v}_{\Delta, \delta} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad C_{12}(\Delta, \tau) = (1 \ -1 \ 0 \ 0) \mathbf{v}_{\Delta, \delta} \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}.$$

It remains to compute C_{11} and C_{12} . First we compute $\psi = \sum_{n \geq 1} \varphi_n$. We readily check that for all $s, t \geq 0$: $\varphi(t-s)\varphi(s) = h(t-s)g(s)\mathbf{Id}$. Thus $\varphi_2 = (h \star g)\mathbf{Id}$. We derive

$$\varphi_{2n} = (h \star g)^{\star n} \mathbf{Id}, \quad \varphi_{2n+1} = \begin{pmatrix} 0 & 0 & (h \star g)^{\star n} \star h & 0 \\ 0 & 0 & 0 & (h \star g)^{\star n} \star h \\ (h \star g)^{\star n} \star g & 0 & 0 & 0 \\ 0 & (h \star g)^{\star n} \star g & 0 & 0 \end{pmatrix}.$$

and we obtain

$$\begin{aligned} \psi &= \sum_{n \geq 1} \varphi_{2n} + \sum_{n \geq 0} \varphi_{2n+1} \\ &= \begin{pmatrix} F & 0 & (\delta_0 + F) \star h & 0 \\ 0 & F & 0 & (\delta_0 + F) \star h \\ (\delta_0 + F) \star g & 0 & F & 0 \\ 0 & (\delta_0 + F) \star g & 0 & F \end{pmatrix} \end{aligned}$$

where $F = \sum_{n \geq 1} (h \star g)^{\star n}$. Set $\tilde{F}(ds) = \delta_0(ds) + F(ds)$. Standard computations yield

$$\mathbf{Id}\delta_0(ds) + \boldsymbol{\psi}(s)ds = \begin{pmatrix} \tilde{F}(ds) & 0 & \tilde{F} \star h(s)ds & 0 \\ 0 & \tilde{F}(ds) & 0 & \tilde{F} \star h(s)ds \\ \tilde{F} \star g(s)ds & 0 & \tilde{F}(ds) & 0 \\ 0 & \tilde{F} \star g(s)ds & 0 & \tilde{F}(ds) \end{pmatrix}$$

and

$$\begin{aligned} & (\mathbf{Id}\delta_0(ds) + \boldsymbol{\psi}(s)ds)\boldsymbol{\Sigma}(\mathbf{Id}\delta_0(dt) + \boldsymbol{\psi}(t)^*dt) = \\ & \frac{1}{1 - \int hfg} \begin{pmatrix} a_{11}(ds, dt) & 0 & a_{13}(ds, dt) & 0 \\ 0 & a_{11}(ds, dt) & 0 & a_{13}(ds, dt) \\ a_{31}(ds, dt) & 0 & a_{33}(ds, dt) & 0 \\ 0 & a_{31}(ds, dt) & 0 & a_{33}(ds, dt) \end{pmatrix} \end{aligned}$$

with:

$$\begin{aligned} a_{11}(ds, dt) &= \nu_1 \tilde{F}(ds) \tilde{F}(dt) + \nu_2 \tilde{F} \star h(ds) \tilde{F} \star h(dt), \\ a_{13}(ds, dt) &= \nu_1 \tilde{F}(ds) \tilde{F} \star g(dt) + \nu_2 \tilde{F} \star h(ds) \tilde{F}(dt), \\ a_{31}(ds, dt) &= \nu_2 \tilde{F}(ds) \tilde{F} \star h(dt) + \nu_1 \tilde{F} \star g(ds) \tilde{F}(dt), \\ a_{33}(ds, dt) &= \nu_2 \tilde{F}(ds) \tilde{F}(dt) + \nu_1 \tilde{F} \star g(ds) \tilde{F} \star g(dt). \end{aligned}$$

Therefore

$$C_{11}(\Delta, \tau) = \int_{[0, \infty)^2} \gamma_\Delta(t - s - \tau) 2a_{11}(ds, dt),$$

and

$$C_{12}(\Delta, \tau) = \int_{[0, \infty)^2} \gamma_\Delta(t - s - \tau) 2a_{31}(ds, dt).$$

To complete the proof of Proposition 3, it suffices to use that for two finite measures μ and ν on \mathbb{R} one has (for all $\Delta > 0$, $\tau \in \mathbb{R}$) $\int_{\mathbb{R}_+^2} \gamma_\Delta(t - s - \tau) \mu(ds) \nu(dt) = \gamma_\Delta \star \nu \star \check{\mu}(\tau)$ where $\check{\mu}$ is the image of μ by $x \rightsquigarrow -x$.

Acknowledgement

The research of E. Bacry and J.F. Muzy is supported in part by the *Chair Financial Risks of the Risk Foundation*. The research of M. Hoffmann is supported in part by the *Agence Nationale de la Recherche*, Grant No. ANR-08-BLAN-0220-01.

References

- [1] F. Abergel and N. Huth. High frequency lead/lag relationship-empirical facts. *Arxiv preprint. arXiv:1111.7103v1*, 2011.
- [2] F. Abergel and F. Pomponio. *Trade-throughs: Empirical facts and application to lead-lag measures*. In *Econophysics of order-driven markets*, F. Abergel, B.K. Chakrabarti, A. Chakrabart and M. Mitra, editors, New Economic Windows. Springer Berlin Heidelberg, 2011.
- [3] Y. Ait-Sahalia, J. Cacho-Diaz, and R. Laeven. Modeling financial contagion using mutually exciting jump processes. *Working paper*, 2011.
- [4] Y. Ait-Sahalia, P.A. Mykland, and L. Zhang. How often to sample a continuous-time process in the presence of market microstructure noise. *The Review of Financial Studies*, 18:351–416, 2005.
- [5] Y. Ait-Sahalia, P.A. Mykland, and L. Zhang. Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160:160–175, 2011.
- [6] T. Andersen, T. Bollerslev, F.X. Diebold, and P. Labys. (understanding, optimizing, using and forecasting) realized volatility and correlation. "Great Realizations," *Risk*,, pages 105–108, 2000.
- [7] E. Bacry, S. Delattre, M. Hoffmann, and J.F. Muzy. Modelling microstructure noise with mutually exciting point processes. *Quantitative Finance*, 13:65–77, 2013.
- [8] O. Barndorff-Nielsen, P. Hansen, A. Lunde, and N. Stephard. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536, 2008.
- [9] M.S. Bartlett. The spectral analysis of point processes. *Journal of the Royal Statistical Society. Series B.*, 25:264–296, 1963.
- [10] L. Bauwens and N. Hautsch. *Modelling financial high frequency data using point processes*. In T. Mikosch, J-P. Kreiss, R. A. Davis, and T. G. Andersen, editors, *Handbook of Financial Time Series*. Springer Berlin Heidelberg, 2009.
- [11] P. Brémaud and L. Massoulié". Power spectra of general shot noises and hawkes point processes with a random excitation. *Advances in Applied Probability*, 34:205–222, 2002.

- [12] D.J. Daley and D. Vere-Jones. *An introduction to the theory of point processes. Vol. I.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2003. Elementary theory and methods.
- [13] K. Dayri, E. Bacry, and J.F. Muzy. Non-parametric kernel estimation for symmetric Hawkes processes. application to high frequency financial data. <http://arxiv.org/abs/1112.1838>, 2011.
- [14] P. Embrechts, J. T. Liniger, and L. Lu. Multivariate Hawkes processes: an application to financial data. *Journal of Applied Probability*, 48:367–378, 2011.
- [15] T. W. Epps. Comovements in stock prices in the very short run. *Journal of the American Statistical Association*, 74:291–298, 1979.
- [16] A.G. Hawkes. Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society, Series B*, 33:438–443, 1971.
- [17] A.G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58:83–90, 1971.
- [18] A.G. Hawkes. *Spectra of some mutually exciting point processes with associated variables.* In *Stochastic Point Processes*. ed. P. A. W. Lewis. Wiley, New York. 1972.
- [19] A.G. Hawkes and D. Oakes. A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11:493–503, 1974.
- [20] P. Hewlett. Clustering of order arrivals, price impact and trade path optimisation. *Workshop on Financial Modeling with Jump processes, Ecole Polytechnique*, 2006.
- [21] M. Hoffmann, M. Rosenbaum, and N. Yoshida. Estimation of the lead-lag effect from nonsynchronous data. *Bernoulli*, 19:426–461, 2013.
- [22] J. Jacod. Multivariate point processes: predictable projection, Radon-Nikodým derivatives, representation of martingales. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 31:235–253, 1974/75.
- [23] J. Jacod and A.N. Shiryaev. *Limit theorems for stochastic processes*, volume 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1987.

- [24] J.Zhuang, Y.Ogata, and D.Vere-Jones. Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association*, 97:369–380, 2002.
- [25] Y. Ogata. The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics*, 30:243–261, 1978.
- [26] P. Reynaud-Bouret and S. Schbath. Adaptive estimation for Hawkes processes; application to genome analysis. *Ann. Statist.*, 38:2781–2822, 2010.
- [27] C.Y. Robert and M. Rosenbaum. A new approach for the dynamics of ultra high frequency data: the model with uncertainty zones. *Journal of Financial Econometrics*, 9:344–366, 2011.
- [28] M. Rosenbaum. A new microstructure noise index. *Quantitative Finance*, 6:883–899, 2011.