Stability and price scaling limit of a Hawkes-process based order book model

Aymen Jedidi, Frédéric Abergel

STABILITY AND PRICE SCALING LIMIT OF A HAWKES PROCESS-BASED ORDER BOOK MODEL

AYMEN JEDIDI AND FRÉDÉRIC ABERGEL

Contents

1. Introduction

Since their introduction in [12], Hawkes processes have been applied to a wide range of research areas, from seismology in the original work by Hawkes, to credit

Date: 8 May 2013.
A. Jedidi, F. Abergel, Chaire de Finance Quantitative, Laboratoire de Mathématiques Appliquées aux Systèmes, École Centrale Paris, 92290 Châtenay-Malabry, France.

risk [11], financial contagion [2] and more recently market microstructure modelling [3, 4, 5, 19, 20, 21].

In market microstructure, and particularly order book modelling, the relevance of these processes comes at least from two empirical properties of the order flow of market and limit orders at the microscopic level:

(1) Time clustering: order arrivals are highly clustered in time.
(2) Mutual dependence: order flow exhibit non-negligible cross-dependences. For instance, as documented in [19], market orders excite limit orders and *vice versa*.

At the microscopic level, point process-based microstructure models capture by construction the intrinsic discreteness of prices and volumes. A question of interest in this context is the microscopic to macroscopic transition in the price dynamics. This strand of research has attracted a lot of interest of late [1, 3, 4, 5, 8, 9, 14, 21].

In this note, we cast a Hawkes process-based order book model into a markovian setting, and using techniques from the theory of Markov chains and stochastic stability [16], show that the order book is stable and leads to a diffusive price limit at large time scales.

Outline. Section 2 is a distillation of some mathematical results about Hawkes processes and Markov chains stochastic stability. Section 3 contains three auxiliary stability results which, apart from their own interest, are useful to prove the stability of the order book; and section 4, the main contribution of this note, is an application to a particular order book model.

Notations. The following notations appear frequently throughout this note, and we recall them here for reference:

- (X_n): discrete-time process,
- $(X(t))$: continuous-time process,
- $|x| = \sum_{i=1}^{p} |x_i|$,
- $[\![1, p]\!] = \{1, 2, \ldots p\}$.

2. Preliminary Remarks

We collect in this section several definitions and results that are useful for the rest of this note. The presentation is rather informal.

2.1. Point Processes.

Definition 2.1 (Point process)**.** *A* point process *is an increasing sequence* $(T_n)_{n\in\mathbb{N}}$ *of positive random variables defined on a measurable space* $(\Omega, \mathcal{F}, \mathbb{P})$.

We will restrict our attention to processes that are *nonexplosive*, that is, for which $\lim_{n\to\infty} T_n = \infty$. To each realization (T_n) corresponds a counting function $(N(t))_{t\in\mathbb{R}^+}$ defined by

$$N(t) = n \text{ if } t \in [T_n, T_{n+1}[, n \geq 0. \tag{1}$$

$(N(t))$ is a right continuous step function with jumps of size 1 and carries the same information as the sequence (T_n), so that $(N(t))$ is also called a point process.

Definition 2.2 (Multivariate point process)**.** *A* multivariate point process *(or marked point process) is a point process (T_n) for which a random variable X_n is associated to each T_n. The variables X_n take their values in a measurable space $(E, \mathcal{E})$.*

We will restrict our attention to the case where $E = \{1, \ldots, M\}$, $m \in \mathbb{N}^*$. For each $m \in \{1, \ldots, M\}$, we can define the counting processes

$$N^m(t) = \sum_{n\geq 1} \mathbb{I}(T_n \leq t)\mathbb{I}(X_n = i). \tag{2}$$

We also call the process

$$N(t) = (N^1(t), \ldots, N^M(t))$$

a multivariate point process.

Definition 2.3 (Intensity of a point process)**.** *A point process $(N(t))_{t\in\mathbb{R}^+}$ can be completely characterized by its (conditional) intensity function, $\lambda(t)$, defined as*

$$\lambda(t) = \lim_{u\to 0} \frac{\mathbb{P}\left[N(t+u) - N(u) = 1|\mathcal{F}_t\right]}{u}, \tag{3}$$

where $\mathcal{F}_t$ is the history of the process up to time t, that is, the specification of all points in $[0, t]$. Intuitively

$$\mathbb{P}\left[N(t+u) - N(u) = 1|\mathcal{F}_t\right] = \lambda(t)\, u + o(u), \tag{4}$$

$$\mathbb{P}\left[N(t+u) - N(u) = 0|\mathcal{F}_t\right] = 1 - \lambda(t)u + o(u), \tag{5}$$

$$\mathbb{P}\left[N(t+u) - N(u) > 1|\mathcal{F}_t\right] = o(u). \tag{6}$$

This is naturally extended to the multivariate case by setting for each m $\in \{1, \ldots, M\}$

$$\lambda^m(t) = \lim_{u\to 0} \frac{\mathbb{P}\left[N^m(t+u) - N^m(u) = 1|\mathcal{F}_t\right]}{u}. \tag{7}$$

2.2. **Hawkes Processes.**

2.2.1. *Hawkes Process.*

Definition 2.4. *A Hawkes process $(N(t))_{t\in\mathbb{R}_+}$ is a point process whose intensity is specified by*

$$\lambda(t) = \mu + \alpha \int_0^t e^{-\beta(t-s)} dN(s) = \mu + \alpha \sum_{0\le s_i \le t} e^{-\beta(t-s_i)}, \tag{8}$$

for a triplet (μ, α, β) of positive real numbers[1].

The process thus defined is *self-excited*: it has a base intensity μ, plus exponentially decaying shocks due to previous jumps. The parameter α characterizes the scale of the excitation and β, its decay in time.

Proposition 2.1. *The process $X(t) = (N(t), \lambda(t))$ is Markov.*

Proof. From a straightforward calculation, we have for any $t_2 > t_1$

$$\begin{aligned}
\lambda(t_2) &= \mu + \alpha \int_0^{t_2} e^{-\beta(t_2-s)} dN(s) && (10)\\
&= \mu + \alpha \int_0^{t_1} e^{-\beta(t_2-s)} dN(s) + \alpha \int_{t_1}^{t_2} e^{-\beta(t_2-s)} dN(s) && (11)\\
&= \mu + e^{-\beta(t_2-t_1)}(\lambda(t_1) - \mu) + \int_{t_1}^{t_2} e^{-\beta(t_2-s)} dN(s), && (12)
\end{aligned}$$

so that, in order to compute $\lambda(t_2)$, we only need to know $\lambda(t_1)$ and $\{N(t) : t_1 \le t \le t_2\}$ - the information contained in $\{N(t), \lambda(t) : 0 \le t < t_1\}$ is irrelevant. Hence

$$\mathbb{P}\left[(N(t_2), \lambda(t_2)) \in A | \{N(t), \lambda(t) : t \in [0, t_1]\}\right] = \mathbb{P}\left[(N(t_2), \lambda(t_2)) \in A | N(t_1), \lambda(t_1)\right], \tag{13}$$

for any measurable set $A \subset \mathbb{N} \times \mathbb{R}^+$, and X is Markov. □

2.2.2. *Multivariate Hawkes Process.*

Definition 2.5. *We say that $N = (N^1, \ldots, N^M)$ is a multivariate Hawkes process when*

$$\lambda^m(t) = \mu_m + \sum_{j=1}^{M} \alpha_{mj} \int_0^t e^{-\beta_{mj}(t-s)} dN^j(s). \tag{14}$$

Proposition 2.2. *Let $Y^{ij}(t) = \alpha_{ij} \int_0^t e^{-\beta_{ij}(t-s)} dN^j(s)$, $1 \le i, j \le M$, and $Y(t) = \{Y^{ij}(t)\}_{1\le i,j\le M}$. The process $X(t) = (N(t), Y(t))$ is Markov.*

[1] A more general definition would have

$$\lambda(t) = \mu + \int_0^t \varphi(t-s) ds, \tag{9}$$

with an unspecified kernel $\varphi > 0$. But we only consider exponentially decaying kernels in this note.

Proof. Let $t_2 > t_1$. Since

$$Y^{mj}(t_2) = e^{-\beta_{mj}(t_2-t_1)}Y^{mj}(t_1) + \int_{t_1}^{t_2} e^{-\beta_{mj}(t_2-s)}dN^j(s), \tag{15}$$

and

$$\lambda^m(t_2) = \mu_m + \sum_{j=1}^{M} Y^{mj}(t_2), \tag{16}$$

the law of $(N(t_2), Y(t_2))$ conditional on $\{(N(t), Y(t)) : 0 \leq t \leq t_1\}$ is the same as when conditioning on $(N(t_1), Y(t_1))$ only—the information contained in $\{(N(t), Y(t)) : 0 \leq t < t_1\}$ is irrelevant, and X is Markov. □

2.2.3. *Stationarity.*

Definition 2.6. *A point process is stationary when for every $r \in \mathbb{N}^*$ and all bounded Borel subsets $A_1, \ldots, A_r$ of the real line, the joint distribution of*

$$\{N(A_1 + t), \ldots, N(A_r + t)\}$$

does not depend on t.

In [13], Hawkes and Oakes show that:

Proposition 2.3. *If*

$$\frac{\alpha}{\beta} < 1 \tag{17}$$

then there exists a (unique) stationary point process $(N(t))$, whose intensity is specified as in definition 2.4.

Brémaud and Massoulié generalize this to the multivariate case in [7]:

Proposition 2.4. *Let the matrix A be defined by*

$$A_{ij} = \frac{\alpha_{ij}}{\beta_{ij}}, \qquad 1 \leq i, j \leq M. \tag{18}$$

If

$$\rho(A) < 1$$

then there exists a (unique) multivariate point process $N(t) = (N^1(t), \ldots, N^m(t))$ whose intensity is specified as in definition 2.5.
$\rho(A)$ is the spectral radius of the matrix A, that is, its largest eigenvalue.

2.3. **The embedded discrete-time Hawkes process.** Throughout this note, we will mostly work with processes sampled in discrete time. We show in this section how to construct a discrete-time version $(X_n)_{n\in\mathbb{N}}$ out of a multivariate Hawkes process $X(t) = (N(t), Y(t))_{t\in\mathbb{R}^+}$, where Y is defined by

$$Y^{ij}(t) = \alpha_{ij} \int_0^t e^{-\beta_{ij}(t-s)} dN^j(s), \qquad 1 \leq i, j \leq M, \tag{19}$$

as in proposition 2.2.

First denote by $(T_n)_{n\geq 1}$ the jump times of the process (and set $T_0 = 0$), and

$$X_n = X(T_n) = (N(T_n), Y(T_n)), \tag{20}$$

and define $E_n = E(T_n) \in \{1, \dots, M\}$ as the mark of the process. The value of E_n indicates which component of $N(t)$ has jumped at time T_n. We also define the waiting times $(\tau_n)_{n\geq 1}$ between two successive jumps as

$$\tau_n = T_{n+1} - T_n. \tag{21}$$

Given that $(N_n, Y_n) = (\xi, y)$, (N_{n+1}, Y_{n+1}) is generated as follows: Set

$$\tau_{n+1} = \min(\tau^1_{n+1}, \dots, \tau^M_{n+1}) \tag{22}$$

where, conditional on $(N_n, Y_n) = (\xi, y)$, the distribution of $\tau^1_{n+1}, \dots, \tau^M_{n+1}$ is that of independent positive random variables whose marginal distributions are determined by hazard rates

$$h^m(t) := \mu_m + \sum_{j=1}^{M} y^{mj} e^{-\beta_{ij} t}, \qquad t \geq 0, 1 \leq m \leq M. \tag{23}$$

Then set

$$E_{n+1} = argmin_{1\leq m\leq M} \tau^m_{n+1}, \tag{24}$$

$$N_{n+1} = (\xi^1, \dots, \xi^{E_n} + 1, \dots, \xi^M), \tag{25}$$

and

$$Y^{mj}_{n+1} = y^{mj}_n e^{-\beta_{mj}\tau_{n+1}} + \alpha^{mj} \mathbb{I}(E_{n+1} = j). \tag{26}$$

2.4. **Drift of a discrete-time Markov process.**

Definition 2.7. *The drift operator $\mathcal{D}$ is defined to act on any nonnegative measurable function V by*

$$\mathcal{D}V(x) = \mathbb{E}[V(X_{n+1}) - V(X_n) | X_n = x]. \tag{27}$$

We will also use the notation

$$\mathcal{P}V(x) = \mathbb{E}[V(X_{n+1}) | X_n = x], \tag{28}$$

hence

$$\mathcal{D}V(x) = \mathcal{P}V(x) - V(x). \tag{29}$$

As will become clear in the next section, the importance of this operator stems from the existence of criteria based on the drift to establish properties of the process. It can be interpreted as the analogue for a process of the derivative for a function[2].

2.5. **A digression on stochastic stability.** Let $(X_n)_{n\in\mathbb{N}}$ be a Markov process on a sate space $\mathcal{S}$ and $(Q^n)_{n\in\mathbb{N}^*}$ its transition probability function, that is

$$Q^n(x, A) = \mathbb{P}\left[X_n \in A | X_0 = x\right], \tag{30}$$

for $x \in \mathcal{S}$ and A a measurable subset of $\mathcal{S}$.

2.5.1. *Ergodicity of a Markov process.* Ergodicity is a strong form of "stability": To quote [16], it means that *"there is an invariant regime described by a measure π such that if the process starts in this regime (that is, if X_0 has distribution π) then it remains in the regime. And moreover if the process starts in some other regime, then it converges in a strong probabilistic sense with π as a limiting distribution."*

Formally, a (aperiodic, irreducible) Markov process is *ergodic* if an invariant[3] probability measure π exists and

$$\lim_{n\to\infty} \|Q^n(x, .) - \pi(.)\| = 0, \forall x \in \mathcal{S}, \tag{32}$$

where $\|.\|$ designates for a signed measure ν the total variation norm[4] defined as

$$\|\nu\| := \sup_{f:|f|<1} |\nu(f)| = \sup_{A\in\mathcal{B}(\mathcal{S})} \nu(A) - \inf_{A\in\mathcal{B}(\mathcal{S})} \nu(A). \tag{34}$$

In (34), $\mathcal{B}(\mathcal{S})$ is the Borel σ-field generated by $\mathcal{S}$, and for a measurable function f on $\mathcal{S}$, $\nu(f) := \int_{\mathcal{S}} f d\nu$.

2.5.2. *V-uniform ergodicity.* We say that a Markov process is *V-uniformly ergodic* if there exists a coercive[5] function $V > 1$, an invariant distribution π, and constants

[2]Cf. Dynkyn's formula or its discrete-time formulation.

[3]That is, satisfying the invariance equations

$$\pi(A) = \int_{\mathcal{S}} \pi(dx) Q(x, A), \qquad A \in \mathcal{B}(\mathcal{S}). \tag{31}$$

[4]If the state space $\mathcal{S}$ is countable (this is *not* the case for $(X(t), Y(t))$ of proposition 2.2.), the convergence in total variation norm implies the more familiar pointwise convergence

$$\lim_{n\to\infty} |Q^n(x, y) - \pi(y)| = 0, \forall x, y \in \mathcal{S}. \tag{33}$$

[5]That is, a function such that $V(x) \to \infty$ as $|x| \to \infty$. The condition $V > 1$ is of course arbitrary and 1 can be replaced by any positive constant.

$0 < r < 1$, and $R < \infty$ such that

$$\|Q^n(x,.) - \pi(.)\| \le Rr^n V(x), x \in \mathcal{S}. \tag{35}$$

This is a strong form of ergodicity (note the geometric rate of convergence), and it can be characterized in terms of the drift operator $\mathcal{D}$. Indeed, it is shown in [16, 17] that it is equivalent to the existence of a coercive function V (the "Lyapunov test function") such that

$$\mathcal{D}V(x) \le -K_1 V(x) + K_2 \mathbb{I}_C(x) \quad \text{(Geometric drift condition.)} \tag{36}$$

for some positive constants K_1 and K_2, and $C \subset \mathcal{S}$ a compact set. (Theorem 16.0.1 in [16].) Condition (36) is equivalent to

$$\mathcal{P}V(x) \le \theta V(x) + K_3 \mathbb{I}_C(x) \tag{37}$$

for some $0 < \theta < 1$. Intuitively, it says that the larger $V(X_n)$ the stronger X is pulled back towards the center of the state space $\mathcal{S}$.

Interestingly, it is possible to develop central limit theorems for functionals of V-uniformly ergodic Markov processes. This will be used to show that the price process in a stable Hawkes process-based order book model is asymptotically diffusive. Before that, we need the following auxiliary results.

3. Auxiliary Results

3.1. **$V-$uniform ergodicity of the intensity of a Hawkes process.** Let $(N(t), \lambda(t))_{t\in\mathbb{R}^+}$ be a Hawkes process with parameters (μ, α, β), and $(N_n, \lambda_n)_{n\in\mathbb{N}}$ its embedded discrete-time process as constructed in section 2.3.

Proposition 3.1. *If $\alpha < \beta$, then the process $(\lambda_n)_{n\in\mathbb{N}}$ is $V-$uniformly ergodic, with*

$$V(\lambda) = e^{\gamma\lambda}, \tag{38}$$

and γ a suitably chosen positive number.

Proof. If $\tau_n = T_{n+1} - T_n$ be the waiting time between two successive jumps of $(X(t))$. There holds, for $t' \in [T_n, T_{n+1}[$,

$$\lambda(t') = \lambda_n + (\lambda_n - \mu)e^{-\beta(t'-T_n)}. \tag{39}$$

The hazard rate associated to τ_n, conditional on $\lambda_n = \lambda \in \mathbb{R}^+$, is

$$h(t) := \mu + (\lambda - \mu)e^{-\beta t}, \tag{40}$$

and the p.d.f. of τ_n is

$$f(t) = h(t)e^{-\int_0^t h(s)ds} = \left(\mu + (\lambda - \mu)e^{-\beta t}\right) e^{-\mu t - \frac{\lambda-\mu}{\beta}(1-e^{-\beta t})}. \tag{41}$$

Let

$$V(\lambda) := e^{\gamma\lambda} \tag{42}$$

be a Lyapunov test function with $\gamma > 0$ an arbitrary parameter. Then,

$$\begin{aligned} \mathbb{E}\left[V\left(\lambda_{n+1}\right)|\lambda_n = \lambda\right] &= \int_0^\infty V(\lambda(t^+))f(t)dt \qquad (43) \\ &= \int_0^\infty V\left(\mu + (\lambda-\mu)e^{-\beta t} + \alpha\right)\left(\mu + (\lambda-\mu)e^{-\beta t}\right)e^{-\mu t - \frac{\lambda-\mu}{\beta}(1-e^{-\beta t})}dt \\ &= \int_0^\infty e^{\gamma\left(\mu+(\lambda-\mu)e^{-\beta t}+\alpha\right)}\left(\mu + (\lambda-\mu)e^{-\beta t}\right)e^{-\mu t - \frac{\lambda-\mu}{\beta}(1-e^{-\beta t})}dt. \end{aligned}$$

Hence,

$$\begin{aligned} \frac{\mathcal{P}V(\lambda)}{V(\lambda)} &= e^{-\gamma\lambda}\,\mathbb{E}\left[V\left(\lambda_{n+1}\right)|\lambda_n = \lambda\right] \qquad (44) \\ &= \int_0^\infty e^{-\gamma(\lambda-\mu)(1-e^{-\beta t})+\gamma\alpha}\left(\mu + (\lambda-\mu)e^{-\beta t}\right)e^{-\mu t - \frac{\lambda-\mu}{\beta}(1-e^{-\beta t})}dt \\ &= e^{\gamma\alpha}\mu\int_0^\infty e^{-\gamma(1+\frac{1}{\beta})(\lambda-\mu)(1-e^{-\beta t})-\mu t}dt \\ &+ e^{\gamma\alpha}(\lambda-\mu)\int_0^\infty e^{-\gamma(1+\frac{1}{\beta})(\lambda-\mu)(1-e^{-\beta t})-(\beta+\mu)t}dt. \qquad (45) \end{aligned}$$

Using lemma A.1, we get

$$\begin{aligned} \frac{\mathcal{P}V(\lambda)}{V(\lambda)} &= e^{\gamma\alpha}\mu\,\mathcal{I}\left((\gamma + \frac{1}{\beta})(\lambda-\mu), \beta, \mu\right) \\ &+ e^{\gamma\alpha}(\lambda-\mu)\,\mathcal{I}\left((\gamma + \frac{1}{\beta})(\lambda-\mu), \beta, \beta+\mu\right), \qquad (46) \end{aligned}$$

where

$$\mathcal{I}(a,b,c) := \int_0^\infty e^{-a(1-e^{-bt})-ct}dt. \tag{47}$$

Then

$$\begin{aligned} \lim_{\lambda\to\infty}\frac{\mathcal{P}V(\lambda)}{V(\lambda)} &= 0 + \frac{e^{\gamma\alpha}}{\beta(\gamma + \frac{1}{\beta})} \\ &= \frac{e^{\gamma\alpha}}{1+\gamma\beta} \qquad (48) \end{aligned}$$

and

$$
\begin{aligned}
\lim_{\lambda\to\infty} \frac{\mathcal{D}V(\lambda)}{V(\lambda)} &= \lim_{\lambda\to\infty} \frac{\mathcal{P}V(\lambda)}{V(\lambda)} - 1 \\
&= \frac{e^{\gamma\alpha}}{1+\gamma\beta} - 1 \\
&= \frac{e^{\gamma\alpha} - 1 - \gamma\beta}{1+\gamma\beta}. \qquad (49)
\end{aligned}
$$

A Taylor expansion in γ around 0 yields

$$\lim_{\lambda\to\infty} \frac{\mathcal{D}V(\lambda)}{V(\lambda)} = \gamma(\alpha-\beta) + o(\gamma), \qquad (50)$$

which has the sign of $\alpha - \beta$. Finally, if $\alpha < \beta$, one can choose $\gamma > 0$, $\lambda_0 \in \mathbb{R}_+^*$ and $\kappa > 0$ such that $\forall \lambda > \lambda_0$

$$\mathcal{D}V(\lambda) \leq -\kappa V(\lambda), \qquad (51)$$

and the V-uniform ergodicity of (λ_n) follows. □

3.2. $V-$**uniform ergodicity of the intensity of a multivariate Hawkes process.** Consider now a multivariate setting. Let $X(t) = (N(t), Y(t))$ be a M-variate Hawkes process with parameters

$$\mu = (\mu_1, \dots, \mu_M)^t, \qquad (52)$$

$$\alpha = (\alpha_{ij})_{1\leq i,j\leq M}, \qquad (53)$$

and

$$\beta = (\beta_{ij})_{1\leq i,j\leq M}. \qquad (54)$$

Define also

$$\alpha_{max} = \max\{\alpha_{ij}\}_{1\leq i,j\leq M} \in \mathbb{R}_+, \qquad (55)$$

and

$$\beta_{min} = \min\{\beta_{ij}\}_{1\leq i,j\leq M} \in \mathbb{R}_+^*, \qquad \beta_{max} = \max\{\beta_{ij}\}_{1\leq i,j\leq M} \in \overline{\mathbb{R}}_+^*. \qquad (56)$$

We recall that $Y(t) = (Y^{ij})_{1\leq i,j\leq M}$ is defined by

$$Y^{ij}(t) = \alpha_{ij} \int_0^t e^{-\beta_{ij}s} dN(s). \qquad (57)$$

As in the monovariate case, let (N_n) and (Y_n) be the discrete time processes

$$
\begin{aligned}
N_n &= N(T_n^+), \\
\text{and } Y_n &= Y(T_n^+), \qquad (58)
\end{aligned}
$$

sampled at the jump times (T_n) of (X). We have the following stability result for (Y_n).

Proposition 3.2. *If*

$$\frac{M\alpha_{max}}{\beta_{min}} e^{1-\frac{M\alpha_{max}}{\beta_{max}}} < 1, \tag{59}$$

then the intensity of a multivariate Hawkes process is V-uniformly ergodic, with

$$V(y) = e^{\gamma \sum_{1\leq k,l\leq M} y^{kl}}, \tag{60}$$

and γ a suitably chosen positive number.

Proof. As in the monovariate case, let

$$V(y) = e^{\gamma \sum_{1\leq k,l\leq M} y^{kl}}. \tag{61}$$

Define the hazard rates

$$h_i(t) = \mu_i + \sum_{j=1}^{M} e^{-\beta_{ij}t} y^{ij}, \qquad 1 \leq i \leq M, \tag{62}$$

and

$$h(t) = \sum_{i=1}^{M} h_i(t). \tag{63}$$

We first note that, conditional on $\tau_{n+1} = t$, the probability that the next jump is on $N^i,\ i \in \{1, \dots, M\}$, is

$$\mathbb{P}\left[E_{n+1} = i | Y_n = y, \tau_{n+1} = t\right] = \frac{h^i(t)}{h(t)}. \tag{64}$$

We have then

$$\begin{aligned} \mathbb{E}\left[V(Y_{n+1})|Y_n = y\right] &= \int_0^\infty \sum_{i=1}^{M} e^{\gamma \sum_{1\leq k,l\leq M}\left(e^{-\beta_{k,l}t} y_{kl} + \mathbb{I}(l=i)\alpha_{k,l}\right)} \frac{h_i(t)}{h(t)} \times h(t) e^{-\int_0^t h(s)ds} dt \\ &= \int_0^\infty \sum_{i=1}^{M} e^{\gamma \sum_{k=1}^{M} \alpha_{ki} + \gamma \sum_{1\leq k,l\leq M} e^{-\beta_{k,l}t} y_{kl}} \left(\mu_i + \sum_{j=1}^{M} y_{ij} e^{-\beta_{ij}t}\right) e^{-\sum_{k=1}^{M} \mu_k t - \sum_{1\leq k,l\leq M} (1-e^{-\beta_{kl}t}) \frac{y_{kl}}{\beta_{kl}}} dt. \end{aligned} \tag{65}$$

Dividing by $V(y)$ and rearranging the terms we get

$$\begin{aligned}\frac{\mathcal{P}V(y)}{V(y)} &= \int_0^\infty \sum_{i=1}^M e^{\gamma\sum_{k=1}^M \alpha_{ki}} e^{\gamma\sum_{1\le k,l\le M}(\gamma+\frac{1}{\beta_{kl}})(1-e^{1-\beta_{k,l}t})y_{kl}}\left(\mu_i+\sum_{j=1}^M y_{ij}e^{-\beta_{ij}t}\right)dt \\ &= \sum_{i=1}^M e^{\gamma\sum_{k=1}^M \alpha_{ki}}\mu_i \mathcal{I}_{M^2}\left(\left((\gamma+\frac{1}{\beta_{kl}})y_{kl}\right)_{1\le k,l\le M};(\beta_{kl})_{1\le k,l\le M};\sum_{k=1}^M \mu_k\right) \\ &+ \sum_{i=1}^M e^{\gamma\sum_{k=1}^M \alpha_{ki}}\sum_{j=1}^M y_{ij}\mathcal{I}_{M^2}\left(\left((\gamma+\frac{1}{\beta_{kl}})y_{kl}\right)_{1\le k,l\le n};\left(\beta^{kl}\right)_{1\le k,l\le M};\sum_{k=1}^n \mu_k+\beta_{ij}\right),\end{aligned} \tag{66}$$

where

$$\mathcal{I}_p(a_1,\ldots,a_p;b_1,\ldots,b_p;c) := \int_0^\infty e^{-a_1(1-e^{-b_1t})-\cdots-a_p(1-e^{-b_pt})-ct}dt \tag{67}$$

is defined in lemma A.2. The first term in the r.h.s of (66) vanishes when $|y|\to\infty$ by lemma A.2. Again using lemma A.2, as $|y|\to\infty$, $\forall 1\le i,j\le M$,

$$\begin{aligned}\mathcal{I}_{M^2}\left(\left((\gamma+\frac{1}{\beta_{kl}})y_{kl}\right)_{1\le k,l\le M};(\beta_{kl})_{1\le k,l\le M};\sum_{k=1}^M \mu_k+\beta_{ij}\right) &\le \frac{1}{\beta_{min}\sum_{1\le k,l\le M}(\gamma+\frac{1}{\beta_{kl}})y_{kl}} \\ &\le \frac{1}{\beta_{min}(\gamma+\frac{1}{\beta_{max}})\sum_{1\le k,l\le M} y_{kl}}\end{aligned} \tag{68}$$

Hence, the second term in the r.h.s of (66) is bounded by

$$\frac{\sum_{i=1}^M e^{\gamma\sum_{k=1}^M \alpha^{ki}}|y|}{\beta_{min}(\gamma+\frac{1}{\beta_{max}})|y|} \le \frac{e^{M\alpha_{max}\gamma}}{\beta_{min}(\gamma+\frac{1}{\beta_{max}})}.$$

And for large $|y|$ we have

$$\frac{\mathcal{P}V(y)}{V(y)} \le \frac{e^{M\alpha_{max}\gamma}}{\beta_{min}(\gamma+\frac{1}{\beta_{max}})}. \tag{69}$$

In order to conclude the proof, it is enough to show that there exists a suitably chosen $\gamma>0$ such that

$$h(\gamma) = \frac{e^{M\alpha_{max}\gamma}}{\beta_{min}(\gamma+\frac{1}{\beta_{max}})} < 1. \tag{70}$$

Minimizing h with respect to γ, the minimum is reached at

$$\gamma^* = \frac{1}{M\alpha_{max}} - \frac{1}{\beta_{max}} > 0. \tag{71}$$

and is equal to

$$h(\gamma^*) = \frac{M\alpha_{max}}{\beta^{min}} e^{1-\frac{M\alpha_{max}}{\beta_{max}}}. \tag{72}$$

Note that for γ^* to be positive (and V to be coercive) we need

$$\alpha_{max} \leq \frac{\beta_{max}}{M}, \tag{73}$$

which we assume. Finally if

$$\frac{M\alpha_{max}}{\beta^{min}} e^{1-\frac{M\alpha_{max}}{\beta_{max}}} < 1, \tag{74}$$

then $(Y_n)_{n\in\mathbb{N}}$ is $V-$uniformly ergodic. □

Remark 3.1. *Note that for* $M = 1$ *the condition is*

$$\frac{\alpha}{\beta} e^{1-\frac{\alpha}{\beta}} < 1, \tag{75}$$

which is satisfied i.i.f.

$$\frac{\alpha}{\beta} < 1. \tag{76}$$

($x \mapsto x(1 - e^x)$ *is strictly increasing from 0 to 1 on* $[0, 1]$*). We get the result in the monovariate case.*

Remark 3.2. *A sufficient condition is*

$$\alpha_{max} < \frac{\beta_{min}}{M}. \tag{77}$$

Remark 3.3. *The stability condition* (74) *is not sharp: it is too stringent on the parameters* $\alpha_{i,j}$ *and* $\beta_{i,j}$*, and we suspect the stationarity condition of proposition 2.4 to be sufficient for V-uniform ergodicty.*

3.3. **V-uniform ergodicity of a "birth-death" Hawkes process.** Let $(N_1(t), N_2(t))$ be a bivariate Hawkes process with intensities:

$$\lambda_1(t) = \mu_1 + \alpha_{11}\int_0^t e^{-\beta_{11}s}dN_1(s) + \alpha_{12}\int_0^t e^{-\beta_{12}s}dN_2(s), \tag{78}$$

$$\lambda_2(t) = \mu_2 + \alpha_{21}\int_0^t e^{-\beta_{21}s}dN_1(s) + \alpha_{22}\int_0^t e^{-\beta_{22}s}dN_2(s), \tag{79}$$

and define the queue $(X(t))$ by

- $X(t) \rightarrow X(t) + 1$ when $N_1(t)$ jumps. This happens with (infinitesimal) probability $\lambda_1(t)dt$.
- $X(t) \rightarrow X(t) - 1$ when $N_2(t)$ jumps and $X(t) \neq 0$. This happens with probability $\lambda_2(t)dt$.

- $X(t) \to X(t) - 1$ with probability $\lambda_3 X(t)dt$ for a constant $\lambda_3 > 0$. This corresponds to a proportional death rate, or in the context of order book modelling, to a proportional cancellation rate.

We also denote by $N_3(t)$ a counting process with intensity $\lambda_3 X(t)$ that jumps by 1 when $X(t)$ jumps by -1 due to a "cancellation".

The queue $X(t)$, albeit peculiar, is the building block of the order book model we present in the next section: N_1 represents the flow of limit orders, N_2 that of market orders and N_3 cancellations.

The following result is the key to the proof of the stability of the order book.

Proposition 3.3. *Provided β_{min} is large (specified precisely below), (X_n, Y_n) is V-uniformly ergodic, where*

$$V(x,y) = e^{\omega x + \gamma \sum_{1\le k,l\le 2} y_{kl}}, \tag{80}$$

and $\omega > 0$ and $\gamma > 0$.

Proof. As usual we write

$$\begin{aligned}
\frac{\mathcal{P}V(x,y)}{V(x,y)} &= e^{\omega+\gamma(\alpha_{11}+\alpha_{12})}\int_0^\infty e^{-\sum_{1\le k,l\le 2}(\gamma+\frac{1}{\beta_{kl}})(1-e^{-\beta klt})y_{kl}-(\mu_1+\mu_2)t-\lambda_3 xt}(\mu_1 + y_{11}e^{-\beta_{11}t} + y_{12}e^{-\beta_{12}t})dt \\
&+ e^{-\omega+\gamma(\alpha_{21}+\alpha_{22})}\int_0^\infty e^{-\sum_{1\le k,l\le 2}(\gamma+\frac{1}{\beta_{kl}})(1-e^{-\beta klt})y_{kl}-(\mu_1+\mu_2)t-\lambda_3 xt}(\mu_2 + y_{21}e^{-\beta_{21}t} + y_{22}e^{-\beta_{22}t})dt \\
&+ e^{-\omega}\int_0^\infty e^{-\sum_{1\le k,l\le 2}(\gamma+\frac{1}{\beta_{kl}})(1-e^{-\beta klt})y_{kl}-(\mu_1+\mu_2)t-\lambda_3 xt}\ \lambda_3 x\, dt \\
&= (e^{\omega+\gamma(\alpha_{11}+\alpha_{12})}\mu_1 + e^{\omega+\gamma(\alpha_{21}+\alpha_{22})}\mu_2 + e^{-\omega}\lambda_3 x)\int_0^\infty e^{-\sum_{1\le k,l\le 2}(\gamma+\frac{1}{\beta_{kl}})(1-e^{-\beta klt})y_{kl}-(\mu_1+\mu_2)t-\lambda_3 xt}dt \\
&+ e^{\omega+\gamma(\alpha_{11}+\alpha_{21})}y_{11}\int_0^\infty e^{-\sum_{1\le k,l\le 2}(\gamma+\frac{1}{\beta_{kl}})(1-e^{-\beta klt})y_{kl}-(\mu_1+\mu_2)t-\lambda_3 xt-\beta_{11}t}dt \\
&+ e^{\omega+\gamma(\alpha_{11}+\alpha_{21})}y_{12}\int_0^\infty e^{-\sum_{1\le k,l\le 2}(\gamma+\frac{1}{\beta_{kl}})(1-e^{-\beta klt})y_{kl}-(\mu_1+\mu_2)t-\lambda_3 xt-\beta_{12}t}dt \\
&+ e^{-\omega+\gamma(\alpha_{12}+\alpha_{22})}y_{21}\int_0^\infty e^{-\sum_{1\le k,l\le 2}(\gamma+\frac{1}{\beta_{kl}})(1-e^{-\beta klt})y_{kl}-(\mu_1+\mu_2)t-\lambda_3 xt-\beta_{21}t}dt \\
&+ e^{-\omega+\gamma(\alpha_{12}+\alpha_{22})}y_{22}\int_0^\infty e^{-\sum_{1\le k,l\le 2}(\gamma+\frac{1}{\beta_{kl}})(1-e^{-\beta klt})y_{kl}-(\mu_1+\mu_2)t-\lambda_3 xt-\beta_{22}t}dt.
\end{aligned} \tag{81}$$

As $|x| + |y| \to \infty$,

$$\frac{\mathcal{P}V(x,y)}{V(x,y)} \le e^{-\omega} + \frac{2}{\beta_{min}(\gamma + \frac{1}{\beta_{max}})}(e^{\omega+\gamma(\alpha_{11}+\alpha_{21})} + e^{-\omega+\gamma(\alpha_{12}+\alpha_{22})}). \tag{82}$$

This quantity can be made smaller than 1 if β_{min} is large enough, hence the stated result. □

Remark 3.4. *Intuitively, a large β corresponds to a short memory for the process* (X_n, Y_n).

4. Application to order book modelling

4.1. **Model setup.** We present a stylized order book model whose dynamics is governed by Hawkes processes. A similar Poissonian order book model has been already discussed at length in [1], so the description provided here is brief.

Assume that each side of the order book is fully described by a finite number of limits K, ranging from 1 to K ticks away from the best available opposite quote. We use the notation

$$X(t) = (a(t); b(t)) = (a_1(t), \dots, a_K(t); b_1(t), \dots, b_K(t)), \tag{83}$$

where $a = (a_1, \dots, a_K)$ designates the ask side of the order book and a_i the number of shares available i ticks away from the best opposite quote, and $b = (b_1, \dots, b_K)$ designates the bid side of the book.

Three types of events can happen:

- arrival of a new limit order;
- arrival of a new market order;
- cancellation of an already existing limit order.

Arrival of limit and market orders are described by four self- and mutually exciting Hawkes processes:

- $L^{\pm}(t)$: arrival of a limit order, with intensity $\lambda^{L^{\pm}}(t)$;
- $M^{\pm}(t)$: arrival of new market order, with intensity $\lambda^{M^{\pm}}(t)$.

Cancellations are modelled by a doubly stochastic Poisson process whose intensity is proportional to the number of shares on each side of the order book, that is

$$\lambda^{C^{\pm}} |x^{\pm}|. \tag{84}$$

We denote by q the size of any new incoming order, and the superscript "+" (respectively "−") refers to the ask (respectively bid) side of the book. Buy limit orders $L^{-}(t)$ arrive below the ask price $P^A(t)$, and sell limit orders $L^{+}(t)$ arrive above the bid price $P^B(t)$.

Once a limit order arrives, its position is chosen randomly between 1 and K. Similarly, once a cancellation order arrives, the order to be cancelled is chosen randomly among the outstanding orders.

Furthermore, we impose constant boundary conditions outside the moving frame of size $2K$: every time the moving frame leaves a price level, the number of shares

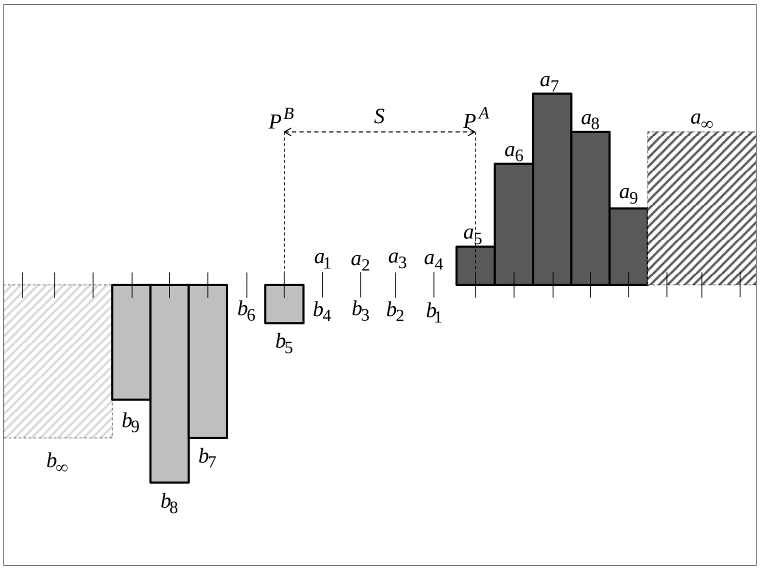

Figure 1. Order book dynamics: in this example, $K = 9$, $q = 1$, $a_\infty = 4$, $b_\infty = -4$. The shape of the order book is such that $a(t) = (0,0,0,0,1,3,5,4,2)$ and $b(t) = (0,0,0,0,-1,0,-4,-5,-3)$. The spread $S(t) = 5$ ticks. Assume that at time $t' > t$ a sell market order $dM^-(t')$ arrives, then $a(t') = (0,0,0,0,0,0,1,3,5)$, $b(t') = (0,0,0,0,0,0,-4,-5,-3)$ and $S(t') = 7$. Assume instead that at $t' > t$ a buy limit order $dL_1^-(t')$ arrives one tick away from the best opposite quote, then $a(t') = (1,3,5,4,2,4,4,4,4)$, $b(t') = (-1,0,0,0,-1,0,-4,-5,-3)$ and $S(t') = 1$.

at that level is set to a_∞ (or b_∞ depending on the side of the book). The quantities a_∞ and b_∞ represent two "reservoirs of liquidity".

Our choice of a finite moving frame and constant boundary conditions has three motivations: firstly, it assures that the order book does not become empty and that P^A, P^B are always well defined. Secondly, it keeps the spread $S = P^A - P^B$ and the increments of P^A, P^B and $P = (P^A + P^B)/2$ bounded - this will be important when addressing the scaling limit of the price. Thirdly, it makes the model markovian, as we do not need to keep track of the price levels that have been visited by the moving frame at some prior time.

Figure 1 is a schematic representation of the order book.

4.2. **Stability of the order book.** We first specify the notations for the 4-variate Hawkes process. We set

$$\lambda_i(t) = \mu_i + \sum_{j=1}^{4} \alpha_{ij} e^{-\beta_{ij} s} dN_j(s), \qquad i \in [\![1, 4]\!], \tag{85}$$

and by convention the index 1 corresponds to L^+, 2 to M^+, 3 to L^- and 4 to M^-.

Proposition 4.1. *Provided β_{min} is large (specified precisely below), the order book (X_n, Y_n) is V-uniformly ergodic, where*

$$V(x, y) = e^{\omega \sum_{i=1}^{K} x_i^{\pm} + \gamma \sum_{1 \leq k,l \leq 4} y_{kl}}, \tag{86}$$

and $\omega > 0$ and $\gamma > 0$.

Proof. We follow the same pattern as the proof of proposition 3.3, and only modify it to account for the fact that the order book is formed from multiple queues, and the role of the boundary conditions a_∞ and b_∞:

$$\begin{aligned}
\frac{\mathcal{P}V(x,y)}{V(x,y)} \quad &\leq \quad e^{\omega q + \omega a_\infty + \gamma \sum_{k=1}^{4} \alpha_{k1}} \int_0^\infty e^{-\sum_{1\leq k,l\leq 4}(\gamma + \frac{1}{\beta_{kl}})(1-e^{-\beta k l t}) y_{kl} - \sum_{k=1}^{4} \mu_k t - \lambda^{C^+} \sum_{k=1}^{K} x_k^+ t} \left(\mu_1 + \sum_{j=1}^{4} y_{1j} e^{-\beta_{1j} t} \right) dt \\
&+ \quad e^{-\omega q + \gamma \sum_{k=1}^{4} \alpha_{k2}} \int_0^\infty e^{-\sum_{1\leq k,l\leq 4}(\gamma + \frac{1}{\beta_{kl}})(1-e^{-\beta k l t}) y_{kl} - \sum_{k=1}^{4} \mu_k t - \lambda^{C^+} \sum_{k=1}^{K} x_k^+ t} \left(\mu_2 + \sum_{j=1}^{4} y_{2j} e^{-\beta_{2j} t} \right) dt \\
&+ \quad e^{-\omega q} \int_0^\infty e^{-\sum_{1\leq k,l\leq 4}(\gamma + \frac{1}{\beta_{kl}})(1-e^{-\beta k l t}) y_{kl} - \sum_{k=1}^{4} \mu_k t - \lambda^{C^+} \sum_{k=1}^{K} x_k^+ t} \left(\lambda^{C^+} \sum_{k=1}^{K} x_k^+ \right) dt \\
&+ \quad \text{similar terms for the bid side of the book} \\
&= \quad \left(e^{\omega q + \omega a_\infty + \gamma \sum_{k=1}^{4} \alpha_{k1}} \mu_1 + e^{-\omega q + \omega a_\infty + \gamma \sum_{k=1}^{4} \alpha_{k2}} \mu_2 + e^{-\omega q} \lambda^{C^+} \sum_{k=1}^{K} x_k^+ \right) \\
&\times \quad \int_0^\infty e^{-\sum_{1\leq k,l\leq 4}(\gamma + \frac{1}{\beta_{kl}})(1-e^{-\beta k l t}) y_{kl} - \sum_{k=1}^{4} \mu_k t - \lambda^{C^+} \sum_{k=1}^{K} x_k^+ t} dt \\
&+ \quad e^{\omega q + \omega a_\infty + \gamma \sum_{k=1}^{4} \alpha_{k1}} \sum_{j=1}^{4} y_{1j} \int_0^\infty e^{-\sum_{1\leq k,l\leq 4}(\gamma + \frac{1}{\beta_{kl}})(1-e^{-\beta k l t}) y_{kl} - \sum_{k=1}^{4} \mu_k t - \lambda^{C^+} \sum_{k=1}^{K} x_k^+ t - \beta_{1j}} dt \\
&+ \quad e^{-\omega q + \gamma \sum_{k=1}^{4} \alpha_{k2}} \sum_{j=1}^{4} y_{2j} \int_0^\infty e^{-\sum_{1\leq k,l\leq 4}(\gamma + \frac{1}{\beta_{kl}})(1-e^{-\beta k l t}) y_{kl} - \sum_{k=1}^{4} \mu_k t - \lambda^{C^+} \sum_{k=1}^{K} x_k^+ t - \beta_{2j}} dt \\
&+ \quad \text{similar terms for the bid side of the book.}
\end{aligned} \tag{87}$$

Again, as $|x| + |y| \to \infty$,

$$\frac{\mathcal{P}V(x,y)}{V(x,y)} \leq e^{-wq} + \frac{4}{\beta_{min}(\gamma + \frac{1}{\beta_{max}})}\left(e^{\omega q+\omega a_\infty+\gamma\sum_{k=1}^4 \alpha_{k1}} + e^{-\omega q+\gamma\sum_{k=1}^4 \alpha_{k2}} + e^{\omega q+\omega b_\infty+\gamma\sum_{k=1} 4\alpha_{k3}} + e^{-\omega q+\gamma\sum_{k=1}^4 \alpha_{k4}}\right). \tag{88}$$

This quantity can be made smaller than 1 if β_{min} is large enough, and this concludes the proof of the proposition. □

4.3. **Large scale limit of the price process.** Given the state (X_{n-1}, Y_{n-1}) of the order book at time $n-1$ and the event E_n, the price increment at time n can be determined. We define the sequence of random variables

$$\eta_n = \Psi(X_{n-1}, Y_{n-1}, E_n), = \Phi(Z_n, Z_{n-1}), \tag{89}$$

as the price increment at time n, where

$$Z_n = (X_n, Y_n). \tag{90}$$

Ψ is a deterministic function giving the elementary "price-impact" of event E_n on the order book at state X_{n-1}. Let μ be the stationary distribution of (Z_n), and M its transition probability function. We are interested in the random sums

$$P_n := \sum_{k=1}^n \overline{\eta}_k = \sum_{k=1}^n \overline{\Phi}(Z_k, Z_{k-1}), \tag{91}$$

where

$$\overline{\eta}_k := \eta_k - \mathbb{E}_\mu[\eta_k] = \overline{\Phi}_k = \Phi_k - \mathbb{E}_\mu[\Phi_k], \tag{92}$$

and the asymptotic behavior of the rescaled-centered price process

$$\widetilde{P}^{(n)}(t) := \frac{P_{\lfloor nt \rfloor}}{\sqrt{n}}, \tag{93}$$

as n goes to infinity.

Proposition 4.2. *In event time, the large-scale limit of the price process is a Brownian motion. Formally, the series*

$$\sigma^2 = \mathbb{E}_\mu[\overline{\eta}_0^2] + 2\sum_{n=1}^\infty \mathbb{E}_\mu[\overline{\eta}_0 \overline{\eta}_n] \tag{94}$$

converges absolutely, and

$$\widetilde{P}^{(n)}(t) \overset{n\to\infty}{\longrightarrow} \sigma B(t), \tag{95}$$

where $(B(t))$ is a standard Brownian motion.

Proof. This is an application of the functional central limit theorem for (stationary and ergodic) sequences of weakly dependent random variables with finite variance, and is identical to the proof of theorem 6.1 in [1]. Firstly, we note that the variance of the price increments η_n is finite since it is bounded by $K + 1$. Secondly, the V-uniform ergodicity of (Z_n) is equivalent to

$$||M^n(z, .) - \mu(.)|| \leq R\rho^n V(z), n \in \mathbb{N}, \tag{96}$$

for some $R < \infty$ and $\rho < 1$. This implies thanks to theorem 16.1.5 in [16][6] that for any g, h such that $Max(g^2, h^2) \leq V$, $k, n \in \mathbb{N}$, and any initial condition z

$$|\mathbb{E}_z[g(Z_k)h(Z_{n+k})] - \mathbb{E}_z[g(Z_k)]\mathbb{E}_z[h(Z_k)]| \leq R\rho^n[1 + \rho^k V(z)], \tag{97}$$

where $\mathbb{E}_z[.]$ means $\mathbb{E}[.|Z_0 = z]$. This in turn implies

$$|\mathbb{E}_z[\overline{h}(Z_k)\overline{g}(Z_{k+n})]| \leq R_1\rho^n[1 + \rho^k V(z)] \tag{98}$$

for some $R_1 < \infty$, where $\overline{h} = h - \mathbb{E}_\mu[h], \overline{g} = g - \mathbb{E}_\mu[g]$. By taking the expectation over μ on both sides of (98) and noting that $\mathbb{E}_\mu[V(Z_0)]$ is finite by theorem 14.3.7 in [16], we get

$$|\mathbb{E}_\mu[\overline{g}(Z_k)\overline{h}(Z_{k+n})]| \leq R_2\rho^n = \rho(n), k, n \in \mathbb{N}. \tag{99}$$

Hence the stationary version of (Z_n) satisfies a geometric mixing condition, and in particular

$$\sum_n \rho(n) < \infty. \tag{100}$$

Theorems 19.2 and 19.3 in [6] on functions of mixing processes allow us to conclude that

$$\sigma^2 := \mathbb{E}_\mu[\overline{\eta}_0^2] + 2\sum_{n=1}^{\infty} \mathbb{E}_\mu[\overline{\eta}_0\overline{\eta}_n] \tag{101}$$

is well-defined - the series in (101) converges absolutely - and coincides with the asymptotic variance

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}_\mu\left[\sum_{k=1}^{n} (\overline{\eta}_k)^2\right] = \sigma^2. \tag{102}$$

Moreover

$$\widetilde{P}^{(n)}(t) \overset{n\to\infty}{\longrightarrow} \sigma B(t), \tag{103}$$

where $(B(t))$ is a standard Brownian motion. The convergence in (103) happens in $D[0, \infty)$, the space of $\mathbb{R}$-valued càdlàg functions, equipped with the Skorohod topology. □

[6]We refer to §16.1.2 *"V-geometric mixing and V-uniform ergodicity"* in [16] for more details.

Appendix A. Technical Lemmas

Lemma A.1. *Let $a, b, c > 0$ be three positive real numbers. Then*

$$\begin{aligned} \mathcal{I}(a,b,c) &= \int_0^\infty e^{-a(1-e^{-bt})-ct} dt \\ &= \frac{(-a)^{-c/b}}{b}\ e^{-a}\ \Big(\Gamma(\frac{c}{b}) - \Gamma(\frac{c}{b}, -a)\Big), \end{aligned} \tag{104}$$

(105)

where the Gamma function *is defined for all complex numbers p such that $\mathcal{R}[p] > 0$ as*

$$\Gamma(p) = \int_0^\infty t^{p-1} e^{-t} dt, \tag{106}$$

and the incomplete Gamma function *is defined for all $p \in \mathbb{C}$, $\mathcal{R}[p] > 0$ and all $z \in \mathbb{C}$ as*

$$\Gamma(p,z) = \int_z^\infty t^{p-1} e^{-t} dt. \tag{107}$$

In particular, for all $b > 0, c > 0$

$$\lim_{a\to\infty} \mathcal{I}(a,b,c) = 0, \tag{108}$$

and

$$\lim_{a\to\infty} a\ \mathcal{I}(a,b,c) = \frac{1}{b}. \tag{109}$$

Proof. This representation and the limits can be obtained with a symbolic computation system such as Mathematica. □

Lemma A.2. *More generally, if*

$$\mathcal{I}_p(a_1,\dots,a_p;b_1,\dots,b_p;c) = \int_0^\infty e^{-a_1(1-e^{-b_1 t})-\dots-a_p(1-e^{-b_p t})-ct} dt, \tag{110}$$

with $a_i > 0, b_i > 0\ \forall i \in [\![1,p]\!]$, and $c > 0$. Let

$$b_{min} = \min\{b_i\}_{1\le i\le p}. \tag{111}$$

Then

$$\begin{aligned} \mathcal{I}_p(a_1,\dots,a_p;b_1,\dots,b_p;c) &\le \int_0^\infty e^{-\sum_{i=1}^p a_i\ (1-e^{-b_{min}t})-ct} \\ &= \mathcal{I}(|a|, b_{min}, c), \end{aligned} \tag{112}$$

whith

$$|a| = \sum_{i=1}^p a_i. \tag{113}$$

Hence

$$\mathcal{I}_p(a;b;c) \leq \frac{1}{b_{min}|a|}, \quad as\ |a| \to \infty. \tag{114}$$

References

[1] F. Abergel and A. Jedidi, A mathematical approach to order book modeling, *International Journal of Theoretical and Applied Finance* (2013).

[2] Y. Aït-Sahalia, J. Cacho-Diaz and R.J.A. Laeven, Modelling financial contagion using mutually exciting Hawkes processes, Preprint (2013).

[3] E. Bacry, S. Delattre, M. Hoffmann and J.F. Muzy, Modelling microstructure noise with mutually exciting point processes, *Quantitative Finance* (2013).

[4] E. Bacry, S. Delattre, M. Hoffmann and J.F. Muzy, Scaling limits for Hawkes processes and application to financial statistics, *Stochastic Processes and Applications* (2013).

[5] E. Bacry and J.F. Muzy, Hawkes model for price and trades high-frequency dynamics, Preprint (2013).

[6] P. Billingsley, *Convergence of Probability Measures*, Wiley, 2nd ed. (1999).

[7] P. Brémaud, *Point Processes and Queues: Martingale Dynamics*, Springer-Verlag (1981).

[8] R. Cont and A. de Larrard, Price dynamics in a Markovian limit order book market, *SIAM Journal for Financial Mathematics* (2013).

[9] R. Cont and A. de Larrard, Order book dynamics in liquid markets: limit theorems and diffusion approximations, Preprint (2012).

[10] D.R. Cox and V. Isham, *Point Processes*, Chapman & Hall/CRC (1980).

[11] E. Errais, K. Giesecke, and L. Goldberg, Affine point processes and portfolio credit risk, *SIAM Journal on Financial Mathematics* (2010).

[12] A. Hawkes, Spectra of some self-exciting and mutually exciting point processes, *Biometrika* (1971).

[13] A. Hawkes and D. Oakes, A cluster process representation of a self-exciting process, Journal of Applied Probability, (1974).

[14] U. Horst and M. Paulsen, A law of large numbers for limit order books, Preprint (2013).

[15] J. Jacod, Multivariate point processes: predictable projections, Radon-Nikodym derivatives, representation of martingales, *Z. Wahrsch. Verw. Gebiete* (1975).

[16] S. Meyn and R.L. Tweedie, *Markov Chains and Stochastic Stability*, Cambridge University Press, 2nd ed. (2009).

[17] S. Meyn and R.L. Tweedie, Stability of Markovian processes I: criteria for discrete-time chains, *Advances in Applied Probability* (1992).

[18] D.J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes*, Springer, 2nd ed. (2003).

[19] I. Muni Toke, Market making behavior in an order book model and its impact on the spread, *Econophysics of Order-driven Markets*, Springer (2011).

[20] I. Muni Toke and F. Pomponio, Modelling trades-through in a limit order book using Hawkes processes. *Economics: The Open-Access Open-Assessment E-Journal* (2012).

[21] B. Zheng, F. Roueff and F. Abergel, Ergodicity and scaling limit of a constrained multivariate Hawkes process, Preprint (2013).

E-mail address: aymen.jedidi@ecp.fr, frederic.abergel@ecp.fr.