Double Master's degree in Mathematics – MAPPA

Università degli studi di Padova – Université Paris Dauphine PSL

Financial Hawkes based modeling across time scales and application to the study of market impact

Davide Zanni

N° étudiant: 22300002 Matricola: 2089126

Supervisor:
Prof. Marc Hoffmann

Co-supervisor:
Prof. Giorgia Callegaro

Academic year 2023/2024

3 September 2024

Abstract

We present a law of large numbers and a central limit theorem for a sequence of multivariate Hawkes processes observed over the interval $[0, T]$ as T tends to infinity. Moved by empirical observations, we focus on nearly unstable Hawkes processes, proving that they asymptotically behave like integrated CIR processes. We apply these results to financial statistics. We consider a Hawkes based price model that accurately captures discrete asset price variations at microscopic level, and we obtain that it converges to a Heston model at macroscopic scale. Then we move to rough volatility setting. We look at the asymptotic behavior of Hawkes processes with a power-law tail and prove they asymptotically resemble integrated fractional CIR processes. Finally, we apply the previous results to describe the market impact of a significant volume of buy orders. We get the shape of the market impact function under no-arbitrage conditions, linking the parameter associated to the Hawkes kernel and the Hurst parameter of the rough volatility model obtained in the limit. Most of the results are based on [BDHM13b], [JR15], [JR16], [JR20] and they are revisited to achieve a unitary work.

Contents

Introduction

In recent years, Hawkes processes have gained increasing importance, finding successful applications in various fields due to their tractability. In this work we aim to explain some of their properties and look at their natural applications in financial price modeling. Most of the results are based on [BDHM13b], [JR15], [JR16], [JR20] and they are revisited to achieve a unitary work. In particular, the thesis is organized as follows. The first part is dedicated to presenting the basic definitions, some statistical properties of Hawkes processes and some limit theorems in different settings, while the second part is focused on their application to the market impact.

Chapter 1

In Chapter 1, we introduce the definition of Hawkes processes and some first limit theorems, as presented in [BDHM13a] and [BDHM13b]. To introduce Hawkes processes in an informal way, we can think about a counting process N_t with values in $\mathbb{N}$ and jumps of unit size. We denote by λ_t the intensity process such that, for all $t \geq 0$,

$$\mathbb{P}(N \text{ has a jump in } [t, t+dt) \,|\mathcal{F}_{t^-}) = \lambda_t \, dt,$$

where $\mathcal{F}_{t^-}$ is the σ-algebra generated by $(N_s)_{s<t}$. The intensity process characterizes the law of the Hawkes process and it is defined by

$$\lambda_t = \mu + \int_{(0,t)} \phi(t-s) dN_s, \quad t \geq 0,$$

where $\mu \in \mathbb{R}^+$ and $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ is a non-negative integrable functions. The application of Hawkes processes in financial modeling is based on the fact that they allow to properly describe some empirical phenomenons, such as the microstructure noise, and they can reproduce the instantaneous movements of the price market at the tick-by-tick level being associated to the positive and negative jumps of an asset price. But in this work we are also interested in their features at large scales. In Chapter 1, we state two important results, a law of large numbers and a functional central limit theorem, taken by [BDHM13b]. The main fact arising by these theorems is that these processes behave like a continuous diffusion at large scales.

Our contribution: We begin the chapter with a brief discussion of the theoretical construction of Hawkes processes as outlined in [Jac75], to facilitate the presentation of the definitions and the results of [BDHM13b]. We then provide the proofs for the two main theorems of the chapter, adapting them to the one-dimensional case, as this will be the framework used in the sequel of the thesis.

Chapter 2

In Chapter 2, moved by empirical considerations in financial applications, we want to deal with the so called *nearly unstable Hawkes processes*, following the main results of [JR15]. For any fixed $T > 0$, consider the Hawkes process $(N_t^T)_{t\in[0,T]}$, with $N_0^T = 0$, characterized by the intensity

$$\lambda_t^T = \mu + \int_{(0,t)} \phi^T(t-s)dN_s^T, \quad t \in [0,T],$$

where $\mu > 0$ and the kernel ϕ^T is a non-negative measurable function on $\mathbb{R}^+$. We always assume to work under the stability condition $\|\phi^T\|_1 < 1$, but in this case we ask that it is almost violated. In particular, we assume that

$$\phi^T(t) = a_T\phi(t), \quad t \geq 0,$$

where ϕ is a non negative measurable function such that $\|\phi\|_1 = 1$, while $(a_T)_{T>0}$ is a sequence of real numbers with $a_T \in (0,1)$ and $a_T \to 1$ as $T \to +\infty$. Moreover, we assume that the function ϕ satisfies another regularity condition, that will be relaxed in the following.

Assumption 1. $m := \displaystyle\int_0^{+\infty} s\phi(s)\,ds < +\infty.$

The goal is to study the limit of the sequence of Hawkes processes $(N^T)_{T>0}$ with these properties as $T \to +\infty$. To do so, the idea is to look at a properly renormalized intensity process and show that its limiting process satisfies a Cox-Ingersoll-Ross (CIR) stochastic differential equation of the form:

$$X_t = \frac{\lambda}{m}\int_0^t (\mu - X_s)ds + \frac{\sqrt{\lambda}}{m}\int_0^t \sqrt{X_s}\,dB_s. \tag{1}$$

Hence, the Hawkes process at appropriate scale with L^1 norm of the kernel close to one looks like an integrated CIR process. This result is then applied to a Hawkes based price model.

Hawkes processes are largely used in price modeling thanks to the possibility of describing properly the main features of high frequency market, reproducing the tick-by-tick microscopic movements of the price, but still preserving a diffusive behaviour when looking at macroscopic scale. This passage from a microscopic process to its macroscopic limit is obtained applying the previous results. The idea is to consider a price process $(P_t)_{t\geq 0}$ defined as

$$P_t = N_t^+ - N_t^-, \quad t \geq 0,$$

where (N^+, N^-) is a two dimensional Hawkes process. Intuitively, this process replicates exactly the upward and downward variations of the price process at microstructure level. In the same spirit as before, we show that the price process asymptotically behaves like a Heston model.

Our contribution: We focus on the presentation of the heuristic derivation of equation (1) starting from the Hawkes model. Indeed, we carry out all the computations to move from the Hawkes intensity to the limiting process, adding some details and explanations to better capture the original idea that led to these results. On the other hand, we provide an outline of the rigorous proofs, referring to [JR15] for full details.

Chapter 3

In Chapter 3, we present a generalization of the results of the previous chapter. Indeed, in practice the intensity of the market order flow does not really behave as a CIR process. To develop a more realistic model, we move to the rough volatility setting and generalize the previous results in the case of a weaker assumption, following [JR16].

It has been empirically established that the log-volatility process of an asset essentially behaves as a fractional Brownian motion with Hurst parameter of order 0.1. This implies the need to develop some fractional models to reproduce the most important financial stylized facts: for example, we work with the so called rough Heston model, which is a natural generalization of the classical case. Similarly to the previous part, we start by a microscopic Hawkes based model and then look at its limit. To achieve a fractional setting, we work with *nearly unstable heavy tailed Hawkes processes.* This means that the sequence of processes $(N^T)_{T>0}$ is exactly as before, but now we relax Assumption 1.

Assumption 2. There exist constants $\alpha \in (0,1)$ and $K > 0$ such that

$$\lim_{x \to +\infty} \alpha x^{\alpha} \int_x^{+\infty} \phi(s)ds = K.$$

This assumption (intuitively denoted by the expression "heavy tails") leads to a different analysis than the previous case. Indeed under this hypothesis, the limit we provide is no longer an integrated semimartingale, making the treatment more complicated. The main result is that the limiting distribution of the nearly unstable heavy tailed Hawkes processes is the integral of a fractional version of the CIR process. This dynamics will be fundamental to find a rough version of the Heston model when we will introduce a price process in the next chapter.

Our contribution: As in the previous chapter, we focus on presenting the heuristic derivation of the results from [JR16], highlighting how fractional dynamics emerge in this new context and outlining the fundamental reasoning behind the rigorous theorem.

Chapter 4

Chapter 4 of the thesis is dedicated to the application of the results of previous chapters to the study of market impact. In this case, the main reference we used is [JR20].

When we consider movements of large amount of orders in the market, a particular attention has to be devoted to the market impact, which is the connection between an incoming order and the consequent price change: indeed, on average we can say that a buy order causes an upward movement of the price, while a sell order a downward change. Our first goal is to define a function which describes this market impact, and in general we can work with metaorders. The concept of metaorder refers to a large amount of orders which cannot be executed in a single transaction since the cumulated volume is much larger than the liquidity available in the order book, but it has to be split into several transactions.

In this context, it is possible to define a model for the order flow which depends on the orders executed in the market and add our buy metaorder to study its impact. The choice is still to consider a Hawkes based model, but in this case using two independent Hawkes processes, say $N^{a,T}$ for the buy orders and $N^{b,T}$ for the sell orders of all the other agents, so each order is assumed to have unit size, and add a non-homogeneous Poisson process n^T for the metaorder. The price process can be defined as

$$P_t^T = P_0 + \int_0^t \xi^T(t-s)\, d(N_s^{a,T} - N_s^{b,T} + n_s^T), \quad t \geq 0, \tag{2}$$

with ξ^T a proper kernel depending on the Hawkes kernel ϕ^T. Then the market impact function MI^T of the metaorder is defined as

$$MI^T(t) = \mathbb{E}[P_t^T - P_0].$$

After some computations, the function MI^T can be decomposed into the sum of two addends, called permanent part (PMI) and transient part (TMI), since in the limit the transient part is vanishing. The theorems presented in this chapter show that, after a suitable rescaling, it is possible to pass from the microscopic market impact to the macroscopic limit to prove that the macroscopic transient market impact has a power law of type

$$TMI(t) \sim t^{1-\alpha}, \quad \text{with } \alpha \in (0,1).$$

Finally, the focus comes back to the price dynamics and we show that the macroscopic dynamics is diffusive with rough volatility. Furthermore, in a particular range of the parameters, it is possible to differentiate the integrated variance in order to get the volatility dynamics that gives, together with the price process, the requested rough Heston model in the limit.

Our contribution: To work with the price process (2) introduced in [JR20], we begin by presenting a construction from [Jai15] to underline the relationship between price and transaction volume. We present the fundamental proofs of the results about market impact, including additional details and computations to facilitate a clearer understanding of the proofs and emphasizing the application of theorems from previous chapters.

Chapter 1

Hawkes processes: definition and asymptotic behaviour

The first chapter is dedicated to introducing some definitions and the main tools we will need in the following, in particular Hawkes processes and some of their properties. These counting processes were introduced by A. G. Hawkes in the early seventies, see for example [Haw71a], [Haw71b], [HO74]. In the last years, they have been largely used in financial modeling thanks to the possibility of reproducing properly the main features of the market, as we will explain in the following chapters. We will present the procedure to build Hawkes-based models, which allow us to describe the movements of an asset price from a microscopic point of view, and then move to the macroscopic scale passing to the limit. To do so, following [BDHM13b], the first step is to present a law of large numbers (LLN) and a functional central limit theorem (CLT) for Hawkes processes, the two main results of this chapter. The most important consequence is that the sequence of properly rescaled Hawkes processes behaves like a continuous diffusion at large scales.

1.1 Preliminary definitions

Let $(\Omega, \mathcal{F})$ be a measurable space and $(\mathcal{F}_t)_{t\geq 0}$ an increasing and right-continuous family of σ-algebras included in $\mathcal{F}$. We want to state the definitions of predictable process and predictable measure, which are independent of the presence of a probability measure in this space.

Definition 1.1. Let $(X_t)_{t\geq 0}$ be a real-valued process and let $\mathcal{P}$ be the σ-algebra of $\Omega \times [0, \infty)$ generated by the applications $(\omega, t) \mapsto Y_t(\omega)$ which are $\mathcal{F}_t$-measurable in ω and left-continuous in t. Then we say that the process $(X_t)_{t\geq 0}$ is *predictable* if it is measurable with respect to $\mathcal{P}$.

Consider $(E, \mathcal{E})$ a measurable state space. We denote by η a random measure, which is a map

$$\eta(\omega; dt, dx) : (\Omega, \mathcal{F}) \to (E, \mathcal{E}).$$

Definition 1.2. A random measure η is called *predictable* if, for any measurable random process $X : (\Omega \times [0, \infty) \times E, \mathcal{P} \otimes \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, the process $(\eta X)_t$ defined by

$$(\eta X)_t(\omega) = \int_E \int_0^t X(\omega, s, x)\eta(\omega; ds, dx)$$

is $\mathcal{P}$-measurable.

Now we can give the specific random measure that allows to define multivariate point processes. We introduce a sequence $(T_n, Z_n)_{n\geq 1}$ of random variables with values in $(0,\infty)\times E$ such that $T_n < T_{n+1}$ and Z_n is $\mathcal{F}_{T_n}$-measurable, for all $n \geq 1$. We set $T_0 = 0$ and $T_\infty = \lim_{n\to\infty} T_n$.

Definition 1.3. We denote by μ the integer-valued random measure defined by

$$\mu(\omega; dt, dx) = \sum_{n\geq 1} \mathbb{1}_{\{T_n(\omega)<\infty\}} \delta_{(T_n(\omega), Z_n(\omega))}(dt, dx). \tag{1.1}$$

The passage to a counting process is just given by the definition

$$N_t := \mu((0,t] \times E),$$

which is an ordinary point process on $(0,\infty)$ when E is only a point, and in this case we will speak directly of N instead of the measure μ. We now introduce the concept of predictable projection of a measure.

Definition 1.4. Given a probability measure P on $(\Omega, \mathcal{F})$, we call *predictable projection* the unique predictable random measure ν such that, for all $\mathcal{P}\otimes\mathcal{E}$-measurable processes X, we have

$$\mathbb{E}\left[\int_E X(t,x)\mu(dt,dx)\right] = \mathbb{E}\left[\int_E X(t,x)\nu(dt,dx)\right].$$

In particular, the predictable projection has the following properties:

1. $(\nu((0,t] \times B))_{t\geq 0}$ is predictable, for all $B \in \mathcal{E}$;
2. $(\mu((0,t\wedge T_n] \times B) - \nu((0,t\wedge T_n] \times B))_{t\geq 0}$ is a local martingale, for all $n \in \mathbb{N}$ and $B \in \mathcal{E}$.

Given these general notions, our goal is to start by the definition of a particular random measure ν and find a probability measure which makes ν the correct predictable projection of the random measure μ introduced in (1.1). Thanks to this procedure, it will be possible to define properly the counting processes used in all the next results. Following the results of [Jac75], we can state the next theorem.

Theorem 1.1. (Theorem 3.6 in [Jac75]) *Let P_0 be a probability measure on $(\Omega, \mathcal{F}_0)$ and ν a predictable random measure such that $\nu(\{t\}\times E) \leq 1$ and $\nu([T_\infty,\infty)\times E) = 0$. Then there exists a unique probability measure P on $(\Omega, \mathcal{F}_\infty)$ whose restriction to $\mathcal{F}_0$ is P_0, and for which ν is the predictable projection of μ.*

We are now ready to introduce the specific elements to achieve the definition of Hawkes process. Consider $E = \{1,\dots,d\}$ and the counting process

$$N_t^i = \mu((0,t]\times\{i\}), \quad \forall t \geq 0, \forall i \in \{1,\dots,d\},$$

where μ is the random measure defined in (1.1), or, equivalently,

$$N_t^i = \sum_{n\geq 1} \mathbb{1}_{\{T_n\leq t\}\cap\{Z_n = i\}} \tag{1.2}$$

with $N^i_0 = 0$ by construction. Introduce the filtration $(\mathcal{F}_t)_{t\geq 0}$, where $\mathcal{F}_t = \sigma(\{N^i_s : s \leq t, 1 \leq i \leq d\})$. Moreover, consider $(\lambda^1_t)_{t\geq 0}, \dots, (\lambda^d_t)_{t\geq 0}$ a set of progressively measurable non-negative processes such that $\int_0^{T_n} \lambda^i_s ds < \infty$, $\forall n \geq 1, \forall i \in \{1, \dots, d\}$, and define

$$\nu(dt, dx) = \sum_{i=1}^{d} \lambda^i_t \, dt \otimes \delta_i(dx),$$

where δ is the Dirac mass. Using Theorem 1.1, we can say that there exists at most one probability measure P on $(\Omega, \mathcal{F}_\infty)$ such that ν is the predictable projection of the random measure μ on $(0, \infty) \times \{1, \dots, d\}$. Therefore in the following we will assume to work in the probability space $(\Omega, \mathcal{F}, P)$ endowed with the filtration $(\mathcal{F}_t)_{t\geq 0}$.

Rephrased in term of point processes, we will refer to $\int_0^t \lambda^i_s \, ds$ as the compensator of the process N^i. It follows that

$$N^i_{t\wedge T_n} - \int_0^{t\wedge T_n} \lambda^i_s \, ds$$

is a $(\mathcal{F}_t)$-martingale.

Thanks to this construction, the process $N = (N^1, \dots, N^d)$ is completely characterized by the intensity process $\lambda = (\lambda^1, \dots, \lambda^d)$ and we can use this fact to define the concept of multivariate Hawkes process in the next section.

1.2 Multivariate Hawkes processes

Definition 1.5. Consider a sequence $(T_n, Z_n)_{n\geq 1}$ with the properties of the previous section and a d-dimensional process $N = (N^1, ..., N^d)$ associated to it, as in the definition (1.2). Define the intensity process as

$$\lambda^i_t = \mu^i + \int_{(0,t)} \sum_{j=1}^{d} \phi^{ij}(t-s) dN^j_s, \quad \forall t \geq 0, \forall i \in \{1, \dots, d\}, \tag{1.3}$$

where $\mu^i \geq 0$ and $\phi^{ij} : \mathbb{R}_+ \to \mathbb{R}_+$ are non-negative measurable functions. Then the process N is called *multivariate Hawkes process.*

In particular, the multivariate Hawkes process $(N_t)_{t\geq 0}$ is characterized by the vector $\mu = (\mu^1, \dots, \mu^d)$ called baseline, and by the $d \times d$-matrix valued function $\phi = (\phi^{ij})_{1\leq i,j\leq d}$ called kernel. Looking at the interpretation of the population model (see Appendix A.2), the constant μ is related to the exogenous movements of the model, while the kernel ϕ to the endogenous movements: this will be explained better when we will use Hawkes processes to build a market model for financial applications.

Remark. The integral in the definition of the intensity process (1.3) has to be thought as the discrete sum over the jump times of the counting process before t. For example, in the case $d = 1$, it can be written as

$$\lambda_t = \mu + \sum_{T_n < t} \phi(t - T_n),$$

where $(T_n)_{n\geq 1}$ are the jump times.

We always assume to work with Hawkes processes that satisfy a stability condition reported in the following assumption. We denote by K the L^1-norm of the matrix-valued function ϕ:

$$K = \int_0^\infty \phi(t)dt.$$

Assumption 1. (Stability condition.) The kernel of the Hawkes process $(N_t)_{t\geq 0}$ satisfies the following properties:

- $\displaystyle\int_0^\infty \phi^{ij}(t)dt < \infty \quad \forall i, j \in \{1, \dots, d\}$;
- the spectral radius $\rho(K)$ of the matrix K satisfies $\rho(K) < 1$.

If $T_\infty = \lim_{n\to\infty} T_n$, we have that $T_\infty = \infty$ almost surely.

We need to introduce another notation that we will use in all the following results. Let ψ be the non-negative measurable function with values in the set of $d \times d$-matrices defined by

$$\psi = \sum_{i\geq 1} (\phi)^{*i}, \tag{1.4}$$

where the symbol $*$ denotes the convolution product:

$$\phi^{*1} = \phi, \quad (\phi)^{*i}(t) = \int_0^t (\phi)^{*(i-1)}(t-s)\phi(s)ds.$$

Now we can use the fact that Young inequality for convolution of positive functions becomes an equality, i.e. if f, g are positive functions:

$$\|f * g\|_{L^1} = \|f\|_{L^1}\|g\|_{L^1}.$$

So using this equality and Assumption 1, it follows that $\int_0^\infty (\phi)^{*i}(s)ds = K \times \cdots \times K$ (i times), and summing over $i \in \mathbb{N}$ we find $\|\psi\|_{L^1}$. Since the sum is convergent, we get that the function ψ is well defined as a function in L^1.

Finally, we denote by $M = (M_t)_{t\geq 0}$ the d-dimensional martingale defined by

$$M_t = N_t - \int_0^t \lambda_s ds,$$

where $\lambda = (\lambda^1, \dots, \lambda^d)$ is the the intensity process.

By this definition of N_t, it is clear that the Hawkes processes are just an extension of Poisson processes, since it is possible to recover a homogeneous Poisson process if $\phi = 0$ in (1.3), hence when the intensity process is a constant. The presence of the kernel ϕ in the definition of the intensity of a Hawkes process produces an autoregressive behaviour. Indeed, let us interpret the intensity process as the instantaneous probability to have a jump of the counting process in an infinitesimal time interval. Then, the fact that a jump occurs in one of the components of the process produces an increase of the probability to have a jump also in the other components. This property is denoted by the expression *mutually exciting processes*. On the other hand, this self-exciting behaviour is compensated by the fact that the regression kernel ϕ is taken as a positive decreasing function. In the sequel, we will use Hawkes processes to build some models in which both the dependence on the past and the self-exciting property will play an important role.

1.3 Law of large numbers

Let $N = (N_t)_{t\geq 0}$ be a multivariate Hawkes process. In this section we start to focus on the behaviour of Hawkes processes when we send the time parameter to infinity. A first theorem in this context is a law of large numbers for Hawkes processes, an useful result that will be applied in the following chapters. We start by presenting two technical formulas.

Lemma 1.2. *Assume that $N = (N_t)_{t\geq 0}$ satisfies Assumption 1. Then for all $t \geq 0$:*

1. $\mathbb{E}(N_t) = t\mu + \left(\int_0^t \psi(t-s)s\,ds\right)\mu$;
2. $N_t - \mathbb{E}(N_t) = M_t + \int_0^t \psi(t-s)M_s\,ds$.

Proof. Proof of (1). First, we prove that, for any finite stopping time τ, it holds

$$\mathbb{E}(N_\tau) = \mu\mathbb{E}(\tau) + \mathbb{E}\left(\int_0^\tau \phi(\tau-u)N_u\,du\right). \tag{1.5}$$

Denote by $(T_n)_{n\geq 1}$ the sequence of the jump times of N and set $S_n = \tau \wedge T_n$. Then, since $N_t - \int_0^t \lambda_s ds$ is a martingale, we have

$$\begin{aligned}\mathbb{E}(N_{S_n}) = \mathbb{E}\left(\int_0^{S_n} \lambda(t)dt\right) &= \mu\mathbb{E}(S_n) + \mathbb{E}\left(\int_0^{S_n}\int_0^t \phi(t-u)\,dN_u\,dt\right)\\ &= \mu\mathbb{E}(S_n) + \mathbb{E}\left(\int_0^{S_n}\int_0^{S_n-u} \phi(t)\,dt\,dN_u\right)\end{aligned}$$

and integrating by parts we find

$$\mathbb{E}(N_{S_n}) = \mu\mathbb{E}(S_n) + \mathbb{E}\left(\int_0^{S_n} \phi(S_n-t)N_t\,dt\right).$$

Finally, using that $T_n \to \infty$ and that τ is finite a.s., we have that $S_n \to \tau$ as $n \to \infty$ and $N_{S_n} \to N_\tau$ as $n \to \infty$, hence by monotone convergence we deduce (1.5).

Now we can derive *(1)*: taking the deterministic stopping time $\tau = t$, it is sufficient to apply Lemma A.1 of the appendix to the equation

$$\mathbb{E}(N_t) = t\mu + \int_0^t \phi(t-s)\mathbb{E}(N_s)\,ds$$

to find

$$\mathbb{E}(N_t) = t\mu + \int_0^t \psi(t-s)\mu s\, ds.$$

Notice that to apply Lemma A.1 we need the map $t \mapsto \mathbb{E}(N_t)$ to be locally bounded, which is guaranteed by the lemma that we present after the end of this proof.

Sketch of proof of (2). Define the processes $X_t = N_t - \mathbb{E}(N_t)$, then it easily follows that

$$X_t = M_t + \int_0^t \phi(t-s)X_s\, ds,$$

and applying again Lemma A.1 we conclude. □

Lemma 1.3. *For any finite stopping time τ, it holds*

$$\mathbb{E}(N_\tau) \le (Id - K)^{-1}\mu\mathbb{E}(\tau)$$

componentwise.

Proof. Using the notation of the previous proof, we can write:

$$\begin{aligned}\mathbb{E}(N_{S_n}) &= \mathbb{E}(S_n)\mu + \mathbb{E}\left(\int_0^{S_n} \phi(S_n - t)N_t\, dt\right)\\ &\le \mathbb{E}(S_n)\mu + K\mathbb{E}(N_{S_n})\end{aligned}$$

componentwise. By induction, we get, $\forall n \ge 1$:

$$\mathbb{E}(N_{S_n}) \le (Id + K + \cdots + K^{n-1})\mathbb{E}(S_n)\mu + K^n\mathbb{E}(N_{S_n})$$

componentwise. Using that $\sum_{n\ge 0} K^n = (Id - K)^{-1}$ and taking the limit on the right hand side, we find

$$\mathbb{E}(N_{S_n}) \le (Id - K)^{-1}\mathbb{E}(\tau)\mu.$$

Finally, using the convergence also on the left hand side, we get the result. □

We are now able to prove a law of large numbers for Hawkes processes.

Theorem 1.4. (Theorem 1 in [BDHM13b]) *Assume that $N = (N_t)_{t\ge 0}$ satisfies Assumption 1. Then $N_t \in L^2(P)$ for all $t \ge 0$ and it holds:*

$$\sup_{t\in[0,1]} \left\| \frac{1}{T}N_{Tt} - t(Id-K)^{-1}\mu \right\| \to 0, \quad \text{as } T \to \infty,$$

almost surely and in L^2.

Remark. Theorem 1.4 shows that Hawkes processes asymptotically behave as their expectation at large time scales. We want to present the result in two specific cases, to underline the idea behind this theorem.

- *Poisson processes.* Assume that the process N is a Poisson process with constant intensity μ. Then it holds:

$$\mathbb{E}\left(\frac{1}{T}N_{Tt}\right) = \frac{\mu T t}{T} = \mu t.$$

- *Hawkes processes in stationary regime.* Assume $d = 1$. Under Assumption 1, we can consider a stationary version of the Hawkes process N, which means that the counting process has independent increments and the distribution of the intensity does not depend on time. Taking the mean (denoted here by $\lambda^{(T)}$) in the definition of λ_{Tt}, we can write
$$\lambda^{(T)} = \mu + \lambda^{(T)} \int_0^{Tt} \phi(s)ds,$$
which gives
$$\lambda^{(T)} = \frac{\mu}{1 - \int_0^{Tt} \phi(s)ds} \xrightarrow[T\to\infty]{} \frac{\mu}{1 - \|\phi\|_{L^1}}.$$
Hence, using the martingale property between N and its compensator,
$$\mathbb{E}\left(\frac{1}{T} N_{Tt}\right) = \mathbb{E}\left(\frac{1}{T}\int_0^{Tt} \lambda_s ds\right) = t\lambda^{(T)} \xrightarrow[T\to\infty]{} \frac{\mu t}{1 - \|\phi\|_{L^1}}.$$

In both these cases it is clear that the term $t(Id - K)^{-1}\mu$ represents the limit of the expectation of the rescaled process. In the proof of the theorem we will show why this is true for general Hawkes processes.

Proof. We start by the proof of the L^2-convergence, which is divided into two parts. We propose the proof in the case $d = 1$, but the procedure is the same also for an arbitrary d. In terms of notation, the only difference is that we will write $(1 - \|\phi\|_{L^1})^{-1}$ instead of $(Id - K)^{-1}$.

Step 1. The goal is to prove the following deterministic convergence:
$$\frac{1}{T}\mathbb{E}(N_{Tt}) - \frac{\mu t}{1 - \|\phi\|_{L^1}} \xrightarrow[T\to\infty]{} 0, \tag{1.6}$$
uniformly in $t \in [0, 1]$. First, thanks to the stability condition introduced above, we have
$$\frac{1}{1 - \|\phi\|_{L^1}} = \sum_{n\geq 0} (\|\phi\|_{L^1})^n = 1 + \sum_{n\geq 1} (\|\phi\|_{L^1})^n = 1 + \int_0^{\infty} \psi(t)dt,$$
which implies, for any fixed $t \in [0, 1]$,
$$\frac{\mu t}{1 - \|\phi\|_{L^1}} = \mu t \left(1 + \int_0^{\infty} \psi(t)dt\right).$$
Hence, using the point *(1)* of Lemma 1.2, we get
$$\begin{aligned}\frac{\mu t}{1 - \|\phi\|_{L^1}} - \frac{1}{T}\mathbb{E}(N_{Tt}) &= \mu t \left(1 + \int_0^{\infty} \psi(u)du\right) - \frac{1}{T}\left(\mu T t + \mu \int_0^{Tt} \psi(Tt - s)s\, ds\right) \\ &= \mu \left(t \int_0^{\infty} \psi(u)du - \frac{1}{T}\int_0^{Tt} \psi(u)(Tt - u)du\right) \\ &= \mu \left(t \int_{Tt}^{\infty} \psi(u)du + \frac{1}{T}\int_0^{Tt} u\, \psi(u)du\right).\end{aligned}$$

For the first integral we can say that

$$\int_T^\infty \psi(tu)du \xrightarrow[T\to\infty]{} 0,$$

and the convergence holds uniformly in $t \in [0,1]$, since ψ is integrable. For the second integral, define the function $G(t) = \int_0^t \psi(s)ds$ and use integration by parts:

$$\frac{1}{T}\int_0^{Tt} u\,\psi(u)du = \frac{1}{T}\left(Tt\int_0^{Tt}\psi(u)du\right) - \frac{1}{T}\int_0^{Tt} G(u)du.$$

It is possible to prove that also this term converges to 0 as $T\to\infty$, uniformly in $t\in[0,1]$, hence we get the convergence in (1.6).

Step 2. It remains to prove that

$$\frac{1}{T}\sup_{t\in[0,1]}|N_{Tt} - \mathbb{E}(N_{Tt})| \xrightarrow[T\to\infty]{} 0 \quad \text{in } L^2(P). \tag{1.7}$$

We want to control this quantity using the martingale M_t in order to conclude applying Doob's inequality. In particular, thanks to the point *(2)* of Lemma 1.2, for any fixed $t\in[0,1]$:

$$\begin{aligned}
|N_{Tt} - \mathbb{E}(N_{Tt})| &= \left|M_{Tt} + \int_0^{Tt}\psi(Tt-s)M_s\,ds\right| \\
&\leq |M_{Tt}| + \sup_{s\in[0,Tt]}|M_s|\int_0^{Tt}|\psi(Tt-s)|\,ds \\
&\leq \sup_{s\in[0,T]}|M_s|\left(1+\int_0^T\psi(T-s)\,ds\right) \\
&\leq C_\phi \sup_{s\in[0,T]}|M_s|
\end{aligned}$$

where the last inequality is given by the boundedness of the L^1 norm of ψ. Using Doob's inequality for the martingale M_t:

$$\mathbb{E}\left[(|N_{Tt}-\mathbb{E}(N_{Tt})|)^2\right] \leq C_\phi^2\,\mathbb{E}[\sup_{s\in[0,T]}|M_s|^2] \leq CT.$$

Hence we deduce

$$\mathbb{E}\left[\frac{1}{T^2}(|N_{Tt}-\mathbb{E}(N_{Tt})|)^2\right] \to 0 \quad \text{as } T\to\infty.$$

We conclude the proof giving a remark on the use of Doob's inequality for M_t. Indeed the inequality gives:

$$\mathbb{E}[\sup_{s\in[0,T]}|M_s|^2] \leq 4\mathbb{E}(M_T^2).$$

By Ito isometry, if M_t was a continuous martingale, the term $\mathbb{E}(M_T^2)$ would give the brackets of the martingale, but in this case we are working with a martingale with jumps.

Hence it is necessary to give an estimate using the quadratic variation:

$$\begin{aligned}\mathbb{E}[\sup_{s\in[0,T]}|M_s|^2] \le 4\mathbb{E}(N_T) = 4\mathbb{E}\left(\int_0^T \lambda_s\,ds\right) \\ = 4\mu T + 4\int_0^T \phi(T-s)\mathbb{E}(N_s)\,ds \\ = 4\mu T + 4\int_0^T \psi(T-s)s\,ds \le CT,\end{aligned}$$

for C a proper constant.

Step 3. To get the convergence almost surely, it is sufficient to prove that

$$\frac{1}{T}\sup_{t\in[0,1]}|M_{Tt}| \xrightarrow[T\to\infty]{} 0 \quad \text{a.s.}$$

This is possible to achieve by showing that also the quadratic variation converges a.s. □

1.4 Functional central limit theorem

In this section, we present a functional central limit theorem. Before stating the theorem, we briefly recall the notion of Skorokhod topology, which is the standard topology used for càdlàg functions.

Denote with $D([0,1])$ the space of càdlàg functions, i.e. functions $x : [0,1] \to \mathbb{R}$ satisfying:

1. (*Right-continuity*) $x^+(t) = \lim_{s\downarrow t} x(s)$ exists and $x^+(t) = x(t)$, for all $0 \le t < 1$;
2. (*Existence of the left-limit*) $x^-(t) = \lim_{s\uparrow t} x(s)$ exists, for all $0 < t \le 1$.

Denote with Λ the class of maps $\lambda : [0,1] \to [0,1]$ which are strictly increasing, continuous, with $\lambda(0) = 0$ and $\lambda(1) = 1$. Then, given $x, y \in D([0,1])$, we can define the Skorokhod metric as

$$d(x,y) := \inf_{\lambda\in\Lambda}\{\|\lambda - Id\|_\infty \vee \|x - y\circ\lambda\|_\infty\}.$$

This metric defines the Skorokhod topology. A sequence $(x_n)_{n\ge1}$ of elements of $D([0,1])$ converges to $x \in D([0,1])$ in the Skorokhod topology if and only if there exists a sequence of functions $(\lambda_n)_{n\ge1}$ in Λ such that $\lim_{n\to\infty} x_n(\lambda_n(t)) = x(t)$ uniformly in t, and $\lim_{n\to\infty}\lambda_n(t) = t$ uniformly in t.

Theorem 1.5. (Theorem 2 in [BDHM13b]) *Assume that $N = (N_t)_{t\ge0}$ satisfies Assumption 1, then the following convergence in law with respect to the Skorokhod topology holds:*

$$\left(\frac{1}{\sqrt{T}}(N_{Tt} - \mathbb{E}(N_{Tt}))\right)_{t\in[0,1]} \xrightarrow[T\to\infty]{} \left((Id-K)^{-1}\Sigma^{\frac{1}{2}}W_t\right)_{t\in[0,1]}$$

where $(W_t)_{t\in[0,1]}$ is a standard d-dimensional Brownian motion and Σ is the diagonal matrix with $\Sigma_{ii} = ((Id-K)^{-1}\mu)_i$.

Remark. The main fact arising by this functional central limit theorem is that we pass from a point process in the microscopic regime to a Brownian diffusion in the macroscopic scale. This passage is fundamental for the applications that we will present in the following chapters, since we want to work with Hawkes-based models which essentially behave as a Brownian motion when we derive the the macroscopic limit.

Proof. (Sketch of the proof.) We propose the proof for the case $d = 1$. Consider W a 1-dimensional Brownian motion, $\Sigma = \frac{\mu}{1-\|\phi\|_{L^1}}$ and $\sigma = \Sigma^{\frac{1}{2}}$.

Step 1. Prove that the sequence of martingales

$$M^{(T)} := \left(\frac{1}{\sqrt{T}} M_{Tt}\right)_{t\in[0,1]}$$

converges in law for the Skorokhod topology to σW. This convergence is achieved applying the Theorem A.5, which is a result for the convergence in law of a sequence of local martingales, taken by [Jac75]. Thanks to this theorem, it is sufficient to notice that $M^{(T)}$, $\forall T > 0$, has uniformly bounded jumps, and

$$[M^{(T)}, M^{(T)}]_t = \sum_{s\leq t} \frac{1}{T} \left(M_{Ts} - M_{Ts}^{-}\right)^2 = \frac{1}{T} N_{Tt}.$$

So we need only that, for all $t \in [0, 1]$,

$$[M^{(T)}, M^{(T)}]_t \xrightarrow[T\to\infty]{} t\sigma^2 \quad \text{in probability}$$

which is a consequence of Theorem 1.4.

Step 2. Define $X_t^{(T)} = \frac{1}{\sqrt{T}}(N_{Tt} - \mathbb{E}(N_{Tt}))$. By step 1, $(1 - \|\phi\|_{L^1})^{-1} M^{(T)}$ converges in law to $(1 - \|\phi\|_{L^1})^{-1}\Sigma^{\frac{1}{2}} W$, hence to conclude the proof it is enough to prove that

$$\sup_{t\in[0,1]} \left| X_t^{(T)} - (1 - \|\phi\|_{L^1})^{-1} M_t^{(T)} \right| \xrightarrow[T\to\infty]{} 0$$

in probability. By Lemma 1.2, $X^{(T)}$ can be written as

$$X_t^{(T)} = M_t^{(T)} + \int_0^t T\psi(Tu) M_{t-u}^{(T)}\, du.$$

Hence our final goal is

$$\sup_{t\in[0,1]} \left| \int_0^t T\psi(Tu) M_{t-u}^{(T)}\, du - \left(\int_0^\infty \psi(s)ds\right) M_t^{(T)} \right| \xrightarrow[T\to\infty]{} 0$$

in probability.

The two main properties used to obtain this goal are the integrability of the function ψ and the C-tightness of the sequence $(M^{(T)})_{T>0}$ (see Appendix A.3 for detailed definitions). In particular, the tightness of the sequence $(M^{(T)})_{T>0}$ is a consequence of step 1, while the fact that $M^{(T)}$ has jumps whose maximum amplitude converges to 0 as $T \to \infty$ is a consequence of the boundedness of the jumps of M. □

We can introduce another assumption to achieve a formulation with the same terms of the law of large numbers.

Assumption 2. Assume that the function ϕ satisfies

$$\int_0^\infty \phi(t) t^{\frac{1}{2}} dt < \infty \quad \text{componentwise.}$$

Thanks to Assumption 2, one can prove that a convergence of type (1.6) holds, with a factor $\sqrt{T}$ in front of the left hand side. Hence we can substitute the term $\mathbb{E}(N_{Tt})$ by its limit in the statement of the previous theorem, and find the following corollary.

Corollary 1.6. *Under Assumptions 1 and 2, the following convergence in law with respect to the Skorokhod topology holds:*

$$\sqrt{T}\left(\frac{1}{T}N_{Tt} - t(Id-K)^{-1}\mu\right)_{t\in[0,1]} \xrightarrow[T\to\infty]{} \left((Id-K)^{-1}\Sigma^{\frac{1}{2}}W_t\right)_{t\in[0,1]}.$$

Proof. It follows by the fact that, under the new assumption, it holds

$$\sqrt{T}\left(\frac{1}{T}\mathbb{E}(N_{Tt}) - t(Id-K)^{-1}\mu\right) \xrightarrow[T\to\infty]{} 0$$

uniformly in $t \in [0,1]$, and that

$$\sqrt{T}\left(\frac{1}{T}N_{Tt} - \frac{1}{T}\mathbb{E}(N_{Tt})\right)_{t\in[0,1]} \xrightarrow[T\to\infty]{} \left((Id-K)^{-1}\Sigma^{\frac{1}{2}}W_t\right)_{t\in[0,1]}$$

in law, by the previous theorem. □

These theorems about the LLN and the CLT are the first two results focused on the limit behaviour of Hawkes processes developed in [BDHM13b]. They will be the starting point to develop further results in a more specific setting when we will present Hawkes-based models in the following chapters.

1.5 Hawkes based price model

We conclude this chapter with the price model presented in [BDHM13a] and [BDHM13b] to give a first idea of the importance of Hawkes processes in high-frequency financial modeling. We want to use Hawkes processes to reproduce exactly the movements of an asset price from a microscopic point of view accounting for the discreteness of price at fine scales.

Define the following process

$$P_t = N_t^+ - N_t^-, \quad t \geq 0,$$

where (N^+, N^-) is a 2-dimensional Hawkes process with intensity

$$\begin{pmatrix}\lambda_t^+\\ \lambda_t^-\end{pmatrix} = \begin{pmatrix}\mu\\ \mu\end{pmatrix} + \int_0^t \begin{pmatrix}0 & \phi(t-s)\\ \phi(t-s) & 0\end{pmatrix}\begin{pmatrix}dN_s^+\\ dN_s^-\end{pmatrix},$$

with $\mu > 0$, and ϕ a non negative measurable function which satisfies the stability condition

$$\int_0^\infty \phi(s)ds < 1.$$

With this construction, we can reproduce the tick-by-tick variations of the process from a microscopic point of view and describe exactly the movements of the price. Indeed, we assume that an asset price moves with jumps of unit size at tick-by-tick level, so N^+ counts the number of upward movements, while N^- counts the number of downward movements for the presence of the minus sign in front of it.

We can notice that the self-exciting property of Hawkes processes has an important role in this model. Indeed, it allows to reproduce the negative correlation of price increments at microstructure level, which is the fact that in the microscopic dynamics of an asset price it is more likely that an upward jump is followed by a downward jump, and vice versa. The autoregressive kernel of the intensity process gives exactly this behaviour.

Another useful property arises if we look at

$$V_t = N_t^+ + N_t^-, \quad t \geq 0.$$

V_t can be seen as the microscopic volatility, since it measures the total fluctuation of prices. In particular, using that both $N_t^+ - \int_{(0,t)} \lambda_s^+ ds$ and $N_t^- - \int_{(0,t)} \lambda_s^- ds$ are martingales, we have that the process

$$V_t - \int_0^t (\lambda_s^+ + \lambda_s^-)ds, \quad t \geq 0,$$

is still a martingale. Hence, we deduce that also the process $(V_t)_t$ is a Hawkes process with intensity

$$\begin{aligned} \lambda_t^V &= \lambda_t^+ + \lambda_t^- \\ &= 2\mu + \int_{(0,t)} \phi(t-s)d(N_s^+ + N_s^-) \\ &= 2\mu + \int_{(0,t)} \phi(t-s)dV_s. \end{aligned}$$

Thanks to these considerations it is clear that Hawkes processes represent an important tool to build financial models at microscopic level. On the other hand, we will show that they allow to find a Brownian diffusion passing to macroscopic scales, which is another important aspect. We will further discuss about this in the next chapter to better explain these properties. See Figures B.1 and B.2 for a numerical simulation of the processes at different time scales.

Chapter 2

Limit theorems for Hawkes processes

In this chapter we introduce the so called nearly unstable Hawkes processes, which are defined as in chapter 1, but adding the assumption that the L^1 norm of their kernel converges to one. This particular regime leads to new limit theorems proved in [JR15]. The idea of dealing with this particular regime is moved by empirical estimations in financial markets. Indeed, the parameter $\|\phi\|_{L^1}$ in the Hawkes based model represents the degree of endogeneity of the market and several studies show that its value is close to one. But this means that the Hawkes processes that we consider are almost unstable. For this reason, it is important to develop limit theorems that allow to describe the dynamics of Hawkes processes also in this regime. The main result of [JR15] is that we can find a CIR dynamics in the limit, which is the usual SDE used in finance to model the squared stochastic volatility. Moreover, we will work with the price process introduced in chapter 1 to reproduce the upward and downward movements of the market, and we will state that it converges to a Heston model at large scales.

2.1 Nearly unstable Hawkes processes

In this chapter we work with Hawkes processes in dimension $d = 1$. We are interested in working with a sequence and describing its limiting behaviour, hence we add an index T which will be the parameter to send to infinity when we want to focus on the asymptotic dynamics. Notice that to work with a countable indexing we should write T_n instead of T, where $(T_n)_{n\geq 0}$ is the discrete sequence of jump times, but we simplify the notation.

Consider a sequence of point processes $(N_t^T)_{t\in[0,T]}$ defined, for each $T > 0$, by $N_0^T = 0$ and by the intensity process

$$\lambda_t^T = \mu + \int_{(0,t)} \phi^T(t-s)dN_s^T, \quad \forall t \in [0,T],$$

where $\mu > 0$ is a constant and $\phi^T : \mathbb{R}_+ \to \mathbb{R}_+$ a non-negative measurable function satisfying the condition $\|\phi^T\|_{L^1} < \infty$. We assume to equip the probability space with a filtration $(\mathcal{F}_t^T)_{t\in[0,T]}$, where $\mathcal{F}_t^T = \sigma(\{N_s^T : s \leq t\})$. From chapter 1, we know that this construction completely characterizes the law of the Hawkes processes.

The goal of this part is presenting some convergence theorems for Hawkes processes which have a kernel that almost violates the stability assumption. This means that we want the $\|\phi^T\|_{L^1}$ to be really close to 1. To describe this regime, we introduce the following assumption on our processes.

Assumption 3. The kernel function ϕ^T is equal to

$$\phi^T(t) = a_T\phi(t), \quad \forall t \in \mathbb{R}_+,$$

where $(a_T)_{T>0}$ is a sequence of real numbers with $a_T \in (0,1), \forall T > 0$, and such that $a_T \to 1$ as $T \to \infty$, while ϕ is a non-negative measurable function such that

- $\|\phi\|_{L^1} = 1$;
- $m := \int_0^\infty s\phi(s)ds < \infty$;
- ϕ is differentiable, $\|\phi'\|_{L^1} < \infty$, $\|\phi'\|_\infty < \infty$.

Thanks to this hypotheses, the L^1 norm of the kernel is equal to a_T and it converges to 1 as $T \to \infty$, making the stability condition almost violated. For this reason, we will refer to this almost unstable situation with the expression *nearly unstable Hawkes processes.*

In this setting, we can distinguish two cases depending on the asymptotic regime as $T \to \infty$: we can assume that $T(1-a_T) \to \infty$ or that $T(1-a_T)$ converges to a finite limit. On one hand, if we consider the case $T(1-a_T) \to \infty$, it means that the stability condition is still preserved since the L^1 norm of the kernel converges slowly to 1. In this case we will provide a result which describes a limiting behaviour similar to the one of the law of large numbers seen in Theorem 1.4 in chapter 1, i.e. a deterministic dynamics. On the other hand, we can ask that $T(1-a_T)$ converges to a finite limit, which means that a_T converges more rapidly to 1. Then the regime is close to instability and the result is different than the previous case.

2.2 Deterministic convergence

Theorem 2.1. (Theorem 2.1 in [JR15]) *Assume that $T(1-a_T) \to \infty$. Under Assumption 3, the sequence of Hawkes processes $(N^T)_{T>0}$ satisfies the following convergence:*

$$\sup_{t\in[0,1]} \frac{1-a_T}{T} |N^T_{Tt} - \mathbb{E}(N^T_{Tt})| \xrightarrow[T\to\infty]{} 0 \quad \text{in } L^2.$$

Proof. First, by the convolution property $\|f * g\|_{L^1} = \|f\|_{L^1}\|g\|_{L^1}$ for positive functions, we can write

$$\begin{aligned}\|\psi^T\|_{L^1} &= \int_0^\infty \psi^T(t)dt = \sum_{k\geq 1}\int_0^\infty (\phi^T)^{*k}(t)dt \\ &= \sum_{k\geq 1}\left(\|\phi^T\|_{L^1}\right)^k = \frac{\|\phi^T\|_{L^1}}{1-\|\phi^T\|_{L^1}}\end{aligned}$$

where the last sum converges thanks to the stability condition $\|\phi^T\|_{L^1} < 1$. Using Lemma 1.2, we have:

$$\begin{aligned}\frac{1-\|\phi^T\|_{L^1}}{T}(N^T_{Tt} - \mathbb{E}(N^T_{Tt})) &= \frac{1-\|\phi^T\|_{L^1}}{T}\left(M^T_{Tt} + \int_0^{Tt} \psi^T(Tt-s)M^T_s\,ds\right)\\ &\leq \frac{1-\|\phi^T\|_{L^1}}{T}(1+\|\psi^T\|_{L^1})\sup_{t\in[0,T]}|M^T_t|\\ &\leq \frac{1}{T}\sup_{t\in[0,T]}|M^T_t|.\end{aligned}$$

Now, the process M^T is a square integrable martingale, with quadratic variation $[M^T, M^T]_t = N^T_t$, so thanks to Doob's inequality we find:

$$\begin{aligned}\mathbb{E}\left((\sup_{t\in[0,T]} M^T_t)^2\right) &\leq 4\sup_{t\in[0,T]}\mathbb{E}((M^T_t)^2) \leq 4\sup_{t\in[0,T]}\mathbb{E}(N^T_t)\\ &= 4\mathbb{E}(N^T_T) \leq 4\mu\frac{T}{1-\|\phi^T\|_{L^1}}\end{aligned}$$

where for the last inequality we have used Lemma 1.3. Finally,

$$\mathbb{E}\left(\sup_{t\in[0,1]}\left(\frac{1-\|\phi^T\|_{L^1}}{T}(N^T_{Tt} - \mathbb{E}(N^T_{Tt}))\right)^2\right) \leq \frac{4\mu}{T(1-\|\phi^T\|_{L^1})},$$

which goes to zero since $T(1-a_T) \to \infty$. This concludes the proof. □

2.3 Heuristic derivation of the asymptotic dynamics

In this section, we deal with the case in which $T(1-a_T)$ converges to a finite limit as $T \to \infty$ and we present a heuristic derivation of the main theorem that we will state below.

Using the same notation of chapter 1, recall that

$$M^T_t = N^T_t - \int_0^t \lambda^T_s ds, \quad \forall t \in [0,T]$$

is the martingale associated to N^T, and ψ^T is a non-negative function defined by

$$\psi^T(t) = \sum_{k=1}^{\infty}(\phi^T)^{*k}(t).$$

Note that ψ^T is well defined since $\|\phi^T\|_{L^1} < 1$.

Introduce now a scaling of the intensity. Recall that we observe the processes in the interval $[0,T]$, but we rescale the time parameter to add the variable $t \in [0,1]$ for all the processes, whereas for the magnitude in space we choose to multiply by the factor $(1-a_T)$. In particular, we define

$$C^T_t = (1-a_T)\,\lambda^T_{Tt}, \quad t \in [0,1].$$

We need to rewrite it in a convenient way to achieve its limiting behaviour, so we start with a formula which gives a useful way to rewrite the intensity λ^T in terms of ψ^T and the martingale M^T.

Proposition 2.2. *The intensity process can be written as*

$$\lambda_t^T = \mu + \int_0^t \psi^T(t-s)\mu\, ds + \int_0^t \psi^T(t-s)\, dM_s^T, \quad \forall t \geq 0.$$

Proof. Starting by the definition of λ_t^T, we have

$$\begin{aligned}\lambda_t^T &= \mu + \int_0^t \phi^T(t-s)dN_s^T \\ &= \mu + \int_0^t \phi^T(t-s)dM_s^T + \int_0^t \phi^T(t-s)\lambda_s^T ds.\end{aligned}$$

Now we can apply Lemma A.1, where the locally bounded function h that we use here is the function

$$h(t) = \mu + \int_0^t \phi^T(t-s)dM_s^T,$$

so we find

$$\lambda_t^T = \mu + \int_0^t \phi^T(t-s)dM_s^T + \int_0^t \psi^T(t-s)\left(\mu + \int_0^s \phi^T(s-u)dM_u^T\right) ds.$$

By Fubini theorem and thanks to the trick of writing $\psi^T * \phi^T = \psi^T - \phi^T$, we get

$$\begin{aligned}\int_0^t \psi^T(t-s)\left(\int_0^s \phi^T(s-u)dM_u^T\right) ds &= \int_0^t \left(\int_0^{t-u} \psi^T(t-u-s)\phi^T(s)ds\right) dM_u^T \\ &= \int_0^t (\psi^T * \phi^T(t-u))dM_u^T \\ &= \int_0^t \psi^T(t-u)dM_u^T - \int_0^t \phi^T(t-u)dM_u^T.\end{aligned}$$

Replacing this term in the intensity, we find the result. □

Thanks to this proposition, the process C^T depends on the function ψ^T, so to get its asymptotic behaviour a first issue is understanding the limit of the function $x \mapsto \psi^T(Tx)$. We provide an explanation introducing some random variables.

Define the following function:

$$\rho^T(x) = \frac{T}{\|\psi^T\|_{L^1}}\psi^T(Tx), \quad x \geq 0.$$

This is a density function, since its L^1-norm is equal to 1, and in particular it is the density function of the random variable

$$X^T = \frac{1}{T}\sum_{j=1}^{I^T} X_j$$

where $(X_j)_{j\geq 1}$ is a sequence of iid random variables with density ϕ and I^T is a geometric random variable with parameter $1-a_T$ independent of $(X_j)_{j\geq 1}$. Indeed, for any continuous and bounded function g, we have

$$\begin{aligned}
\mathbb{E}\left[g\left(\frac{1}{T}\sum_{j=1}^{I^T} X_j\right)\right] &= \sum_{k\geq 1}\mathbb{P}(I^T=k)\,\mathbb{E}\left[g\left(\frac{1}{T}\sum_{j=1}^{k} X_j\right)\right]\\
&= \sum_{k\geq 1}(1-a_T)(a_T)^{k-1}\int_0^\infty g\left(\frac{1}{T}y\right)(\phi^{*k}(y))dy\\
&= \frac{T(1-a_T)}{a_T}\int_0^\infty g(x)\sum_{k\geq 1}\phi^{*k}(Tx)(a_T)^k dx\\
&= \frac{T}{\|\psi^T\|_{L^1}}\int_0^\infty g(x)\sum_{k\geq 1}(a_T\phi(Tx))^{*k}dx\\
&= \int_0^\infty g(x)\frac{T\psi^T(Tx)}{\|\psi^T\|_{L^1}}dx.
\end{aligned}$$

Now we want to compute the characteristic function of the random variable X^T and provide its asymptotic behaviour. Denote by $\hat{\phi}$ the characteristic function of the random variable X_1 and with $\hat{\rho}^T$ the characteristic function of X^T. For $z\in\mathbb{R}$,

$$\begin{aligned}
\hat{\rho}^T(z) = \mathbb{E}[e^{izX^T}] &= \sum_{k\geq 1}\mathbb{P}(I^T=k)\,\mathbb{E}[e^{i\frac{z}{T}\sum_{j=1}^k X_j}]\\
&= \sum_{k\geq 1}(1-a_T)(a_T)^{k-1}\prod_{j=1}^{k}\mathbb{E}[e^{i\frac{z}{T}X_j}]\\
&= \sum_{k\geq 1}(1-a_T)(a_T)^{k-1}\left(\hat{\phi}(\frac{z}{T})\right)^k\\
&= \frac{\hat{\phi}(\frac{z}{T})}{1-\frac{a_T}{1-a_T}(\hat{\phi}(\frac{z}{T})-1)}.
\end{aligned}$$

To achieve the asymptotic behaviour, we can write the expansion of the term $\hat{\phi}(\frac{z}{T})$. Indeed, using that $\hat{\phi}$ is continuously differentiable and X_1 has density ϕ, we get

$$\frac{\partial\hat{\phi}}{\partial u}(u) = \mathbb{E}[iX_1e^{iuX_1}],$$

$$\frac{\partial\hat{\phi}}{\partial u}(u=0) = \mathbb{E}[iX_1] = im,$$

and, using that $a_T\to 1$ and $\hat{\phi}(\frac{z}{T})\to 1$ as $T\to\infty$, we can write

$$\hat{\phi}(\frac{z}{T}) - 1 \sim_{T\to\infty} im\frac{z}{T}.$$

So at the end we can rewrite the function $\hat{\rho}^T$ to make explicit the asymptotic behaviour:

$$\hat{\rho}^T(z) \sim_{T\to\infty} \frac{1}{1-\frac{imz}{T(1-a_T)}}.$$

At this point, we can notice that the only possibility to get a non trivial law by this characteristic function is to ask that the order of the observation time is $T \sim (1 - a_T)^{-1}$. So we consider the asymptotic setting given in the next assumption.

Assumption 4. There exists $\lambda > 0$ such that

$$T(1 - a_T) \xrightarrow[T \to \infty]{} \lambda.$$

Under this assumption, we get that $\hat{\rho}^T(z)$ converges pointwise, as $T \to \infty$, to

$$\frac{1}{1 - iz\frac{m}{\lambda}} = \frac{\frac{\lambda}{m}}{\frac{\lambda}{m} - iz} =: \hat{\rho}(z),$$

which is the characteristic function of an exponential random variable with parameter $\frac{\lambda}{m}$. So by Lévy theorem, we can get the convergence of the random variables. We write this result in the following proposition.

Proposition 2.3. *Under Assumptions 3 and 4, the sequence of random variables $(X^T)_{T>0}$ converges in law towards an exponential random variable with parameter $\frac{\lambda}{m}$.*

Remark. Notice that the hypothesis that m is finite given in Assumption 3 is fundamental to obtain this proposition, since it guarantees a non degenerate limit for the function ψ^T. In the next chapter we will work under a more general assumption on the kernel ϕ, leading to different results.

At this point we are able to provide the convergence behaviour of the function ψ^T. Introduce the quantity

$$u_T = \frac{T(1 - a_T)}{\lambda}$$

which in particular converges to 1 as $T \to \infty$. We can write

$$\psi^T(Tx) = \rho^T(x)\frac{\|\psi^T\|_{L^1}}{T} = \rho^T(x)\frac{a_T}{T(1 - a_T)} = \rho^T(x)\frac{a_T}{\lambda u_T}.$$

So, passing to the limit and thanks to the convergence proved for the function ρ^T, we get

$$\psi^T(Tx) \xrightarrow[T \to \infty]{} \frac{1}{m}e^{-\frac{\lambda}{m}x}.$$

We can now provide the convergence of the process C^T. Using Proposition 2.2 and the notation introduced so far, we have

$$
\begin{aligned}
C_t^T &= (1-a_T)\lambda_{Tt}^T \\
&= (1-a_T)\mu + (1-a_T)\mu \int_0^{Tt} \psi^T(Tt-s)ds + (1-a_T)\int_0^{Tt} \psi^T(Tt-s)\, dM_s^T \\
&= (1-a_T)\mu + \mu u_T \int_0^t \lambda\psi^T(Ts)ds + u_T \int_0^{Tt} \frac{\lambda}{T}\psi^T(Tt-s)\, dM_s^T \\
&= (1-a_T)\mu + \mu \int_0^t u_T\lambda\psi^T(Ts)ds + u_T \int_0^{t} \frac{\lambda}{T}\psi^T(T(t-u))\, dM_{Tu}^T \\
&= (1-a_T)\mu + \mu \int_0^t u_T\lambda\psi^T(Ts)ds + \int_0^{t} \sqrt{\lambda}\psi^T(T(t-s))\frac{\sqrt{u_T}}{\sqrt{T}}\sqrt{\frac{C_s^T}{\lambda_{Ts}^T}} dM_{Ts}^T \\
&= (1-a_T)\mu + \mu \int_0^t u_T\lambda\psi^T(Ts)ds + \int_0^{t} \sqrt{\lambda}\sqrt{C_s^T}\psi^T(T(t-s))\frac{\sqrt{u_T}}{\sqrt{T}}\frac{dM_{Ts}^T}{\sqrt{\lambda_{Ts}^T}} \\
&= (1-a_T)\mu + \mu \int_0^t u_T\lambda\psi^T(Ts)ds + \int_0^{t} \sqrt{\lambda}\sqrt{C_s^T}\psi^T(T(t-s))dB_s^T
\end{aligned}
$$

where we have introduced the process

$$B_t^T = \frac{\sqrt{u_T}}{\sqrt{T}} \int_0^{Tt} \frac{dM_s^T}{\sqrt{\lambda_s^T}}.$$

We use the following limits to deduce heuristically the convergence behaviour:

- $\psi^T(Tx) \to \frac{1}{m}e^{-\frac{\lambda}{m}x}$, as $T \to \infty$,
- $B^T \to B$ in law, as $T \to \infty$, where B is a standard Brownian motion,
- $u_T \to 1$ as $T \to \infty$.

Hence we deduce the dynamics of the limit C^∞ of the sequence $(C^T)_{T>0}$, as $T \to \infty$:

$$C_t^\infty = \mu(1-e^{-\frac{\lambda}{m}t}) + \frac{\sqrt{\lambda}}{m}\int_0^t e^{-\frac{\lambda}{m}(t-s)}\sqrt{C_s^\infty}dB_s. \tag{2.1}$$

Now we use the Itô's formula for semimartingales

$$X_tY_t = X_0Y_0 + \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle X, Y\rangle_t,$$

where in our case X is a semimartingale and Y is a finite variation process, so the last term of the formula is zero. So we get

$$C_t^\infty = \mu\frac{\lambda}{m}\int_0^t e^{-\frac{\lambda}{m}s}ds + \frac{\sqrt{\lambda}}{m}\left(\int_0^t \sqrt{C_s^\infty}dB_s - \int_0^t\left(\int_0^s \sqrt{C_u^\infty}dB_u\right)\frac{\lambda}{m}e^{-\frac{\lambda}{m}(t-s)}ds\right).$$

Consider the last integral, using Fubini theorem and the change of variable $s = t + u - r$, we find

$$\begin{aligned}\int_0^t \left(\int_0^s \sqrt{C_u^\infty} dB_u\right) e^{-\frac{\lambda}{m}(t-s)} ds &= \int_0^t \int_u^t \sqrt{C_u^\infty} e^{-\frac{\lambda}{m}(t-s)} ds\, dB_u \\ &= \int_0^t \int_u^t \sqrt{C_u^\infty} e^{-\frac{\lambda}{m}(r-u)} dr\, dB_u \\ &= \int_0^t \int_0^r \sqrt{C_u^\infty} e^{-\frac{\lambda}{m}(r-u)} dB_u\, dr\end{aligned}$$

and finally, inverting the relation (2.1) to make explicit the integral, we find

$$\int_0^t \int_0^r \sqrt{C_u^\infty} e^{-\frac{\lambda}{m}(r-u)} dB_u\, dr = \int_0^t \frac{m}{\sqrt{\lambda}} (C_r^\infty - \mu + \mu e^{-\frac{\lambda}{m} r}) dr.$$

So at the end we have

$$C_t^\infty = \mu \frac{\lambda}{m} \int_0^t e^{-\frac{\lambda}{m} s} ds + \frac{\sqrt{\lambda}}{m} \int_0^t \sqrt{C_s^\infty} dB_s - \frac{\lambda}{m} \int_0^t (C_r^\infty - \mu + \mu e^{-\frac{\lambda}{m} r}) dr, \quad t \in [0,1],$$

and simplifying,

$$C_t^\infty = \int_0^t (\mu - C_s^\infty) \frac{\lambda}{m} ds + \frac{\sqrt{\lambda}}{m} \int_0^t \sqrt{C_s^\infty} dB_s, \quad t \in [0,1], \tag{2.2}$$

which is the CIR equation.

In this way we have obtained the macroscopic dynamics of the rescaled intensity process through heuristic arguments. In the sequel, we will present the theorem which gives this result in a rigorous way.

2.4 CIR dynamics

We introduce a new hypothesis on the boundedness of the density function ρ^T and then we state the theorem.

Assumption 5. There exists a constant $K_\rho > 0$ such that, for all $x \geq 0$ and $T > 0$,

$$|\rho^T(x)| \leq K_\rho.$$

Theorem 2.4. (Theorem 2.2 in [JR15]) *Under Assumptions 3, 4 and 5, the sequence of renormalized Hawkes intensities $(C^T)_{T>0}$ converges in law, for the Skorokhod topology, towards the law of the unique strong solution of the following Cox-Ingersoll-Ross stochastic differential equation:*

$$X_t = \int_0^t (\mu - X_s) \frac{\lambda}{m} ds + \frac{\sqrt{\lambda}}{m} \int_0^t \sqrt{X_s} dB_s, \quad t \in [0,1].$$

Moreover, the sequence of renormalized Hawkes process

$$V_t^T = \frac{1 - a_T}{T} N_{Tt}^T$$

converges in law, for the Skorokhod topology, towards the process

$$\int_0^t X_s ds, \quad t \in [0,1].$$

Proof. The goal of the proof is to prove that the processes $(C_t^T)_{T>0}$, for $t \in [0,1]$, introduced above converge towards a CIR process. To do so, the idea is to rewrite the SDE satisfied by the process C_t^T and then find the limit of the sequence of SDEs, in order to conclude that the laws of the solutions of the sequence of SDEs converge to the law of the solution of the limiting SDE.

Step 1. First, we want to rewrite the equation satisfied by the process C_t^T in a convenient way, using the limiting behaviour we have developed in the heuristic construction. Consider the formula of previous section:

$$C_t^T = (1-a_T)\mu + \mu \int_0^t u_T \lambda \psi^T(Ts)ds + \int_0^t \sqrt{\lambda}\sqrt{C_s^T}\psi^T(T(t-s))dB_s^T$$

where

$$B_t^T = \frac{\sqrt{u_T}}{\sqrt{T}} \int_0^{Tt} \frac{dM_s^T}{\sqrt{\lambda_s^T}}.$$

We have already obtained the following limits:

- $u_T \xrightarrow[T\to\infty]{} 1$,
- $\psi^T(Tx) \xrightarrow[T\to\infty]{} \frac{1}{m}e^{-\frac{\lambda}{m}x}$,
- $\rho^T(x) \xrightarrow[T\to\infty]{} \rho(x) = \frac{\lambda}{m}e^{-\frac{\lambda}{m}x}$,

so we can add and subtract these terms in the expression of C_t^T, in order to prove the convergence in the following steps. In practice, we have

$$\begin{aligned}\int_0^t u_T\lambda\psi^T(Ts)ds &= \int_0^t \frac{T(1-a_T)}{\lambda}\lambda\psi^T(Ts)ds = \int_0^t \frac{1-a_T}{a_T}Ta_T\psi^T(Ts)ds \\ &= \int_0^t \frac{Ta_T}{\|\psi^T\|_{L^1}}\psi^T(Ts)ds = \int_0^t a_T\rho^T(s)ds\end{aligned}$$

which converges to

$$\int_0^t \frac{\lambda}{m}e^{-\frac{\lambda}{m}s}ds = 1 - e^{-\frac{\lambda}{m}t}.$$

So the expression of C_t^T becomes

$$C_t^T = R_t^T + \mu(1-e^{-\frac{\lambda}{m}t}) + \frac{\sqrt{\lambda}}{m}\int_0^t e^{-\frac{\lambda}{m}(t-s)}\sqrt{C_s^T}dB_s^T,$$

where

$$\begin{aligned}R_t^T =& (1-a_T)\mu - \mu\left((1-e^{-\frac{\lambda}{m}t}) - \int_0^t \frac{Ta_T}{\|\psi^T\|_{L^1}}\psi^T(Ts)ds\right) \\ &+ \sqrt{\lambda}\int_0^t \left(\psi^T(T(t-s)) - \frac{1}{m}e^{-\frac{\lambda}{m}(t-s)}\right)\sqrt{C_s^T}dB_s^T.\end{aligned}$$

Using integration by parts and a change of variable in a similar way of what we did for the heuristic argument, we finally derive the expression

$$C_t^T = U_t^T + \frac{\lambda}{m}\int_0^t (\mu - C_s^T)ds + \frac{\sqrt{\lambda}}{m}\int_0^t \sqrt{C_s^T}dB_s^T \tag{2.3}$$

with

$$U_t^T = R_t^T + \frac{\lambda}{m} \int_0^t R_s^T ds.$$

Moved by the fact that we want to find (2.2) developed in the heuristic argument as limiting dynamics, the goal is to prove that R_t^T converges to zero, so that also U_t^T goes to zero, in order to get exactly (2.2) when we take the limit as $T \to \infty$ in (2.3).

Step 2. Starting by the expression of R_t^T, recall that $a_T \to 1$ as $T \to \infty$ and the term

$$\int_0^t \frac{T a_T}{\|\psi^T\|_{L^1}} \psi^T(Ts) ds$$

converges to $1 - e^{-\frac{\lambda}{m}t}$. So it remains to look only at the following term

$$Y_t^T = \int_0^t \left(m\psi^T(T(t-s)) - e^{-\frac{\lambda}{m}(t-s)} \right) \frac{1}{T} dM_t^T.$$

It is sufficient to prove that $(Y_t^T)_{T>0}$, for $t \in [0,1]$, converges to zero in law with respect to Skorokhod topology. The standard method to achieve this result is to prove the finite dimensional convergence and tightness of the sequence $(Y_t^T)_{T>0}$, for $t \in [0,1]$. We refer to the proof presented in [JR15] for details.

Step 3. We prove the convergence in law of the sequence $(B_t^T)_{t\in[0,1]}$, with

$$B_t^T = \frac{\sqrt{u_T}}{\sqrt{T}} \int_0^{Tt} \frac{dM_s^T}{\sqrt{\lambda_s^T}},$$

to a standard Brownian motion $(B_t)_{t\in[0,1]}$, with respect to the Skorokhod topology. First, by definition $(B^T)_{T>0}$ is a sequence of martingales with uniformly bounded jumps, since we have

$$\lambda_t^T = \mu + \int_{(0,t)} \phi^T(t-s) dN_s^T = \mu + \sum_{i<N_t^T} \phi^T(t - T_i) \geq \mu.$$

Therefore,

$$\frac{1}{\sqrt{T}} \int_0^{Tt} \frac{1}{\sqrt{\lambda_s^T}} dM_s^T \leq \frac{1}{\sqrt{T\mu}} \left(N_{Tt}^T - \int_0^{Tt} \lambda_s^T ds \right)$$

and, for any $t \in [0,1]$,

$$\frac{1}{\sqrt{T\mu}} (N_{Tt}^T - N_{Tt^-}^T) \leq \frac{1}{\sqrt{T\mu}}.$$

Hence we deduce

$$\sup_{t\in[0,1]} (B_{Tt}^T - B_{Tt^-}^T) \leq \frac{c}{\sqrt{\mu}},$$

which gives the uniformly bounded jumps. Now it is sufficient to look at the quadratic variation of $(B_t^T)_{t\in[0,1]}$ and to prove the convergence to the quadratic variation of a Brownian motion. Indeed,

$$[B^T, B^T]_t = \frac{u_T}{T} \int_0^{Tt} \frac{1}{\lambda_s^T} dN_s^T = \frac{u_T}{T} \left(Tt + \int_0^{Tt} \frac{1}{\lambda_s^T} dM_s^T \right)$$

and it holds

$$\mathbb{E}\left[\left(\int_0^{Tt}\frac{1}{T\lambda_s^T}dM_s^T\right)^2\right]\leq\mathbb{E}\left[\int_0^T\frac{1}{T^2\lambda_s^T}ds\right]\leq\frac{1}{T\mu},$$

which goes to 0 as $T\to\infty$. Therefore we deduce the convergence for the quadratic variation

$$[B^T,B^T]_t\xrightarrow[T\to\infty]{}t\quad\text{in probability,}$$

where the limit is the quadratic variation of a standard Brownian motion. Therefore, applying Theorem A.5 we get that $(B_t^T)_{t\in[0,1]}$ converges in law for the Skorokhod topology towards a standard Brownian motion, as $T\to\infty$.

Step 4. We put together the previous steps to to provide a limit for the sequence of SDEs

$$C_t^T=U_t^T+\frac{\lambda}{m}\int_0^t(\mu-C_s^T)ds+\frac{\sqrt{\lambda}}{m}\int_0^t\sqrt{C_s^T}dB_s^T,$$

with

$$B_t^T=\frac{\sqrt{u_T}}{\sqrt{T}}\int_0^{Tt}\frac{dM_s^T}{\sqrt{\lambda_s^T}}.$$

We have that

$$(U_t^T,B_t^T)_{t\in[0,1]}\xrightarrow[T\to\infty]{}(0,B_t)_{t\in[0,1]}$$

in law for the Skorokhod topology. This implies that the limiting SDE is the following:

$$X_t=\int_0^t(\mu-X_s)\frac{\lambda}{m}ds+\frac{\sqrt{\lambda}}{m}\int_0^t\sqrt{X_s}dB_s \tag{2.4}$$

which describes the CIR dynamics. Finally, recall that this equation admits a unique strong solution on $[0,1]$. All these arguments allow us to conclude that the sequence of solutions of (2.3) converges in law to the solution of (2.4).

Step 5. We prove the second part of Theorem 2.4. Introduce the following notation

$$V_t^T=\frac{1-a_T}{T}N_{Tt}^T.$$

Using that

$$\int_0^t C_s^T ds=\int_0^t(1-a_T)\lambda_{Ts}^T ds=\frac{1-a_T}{T}\int_0^{Tt}\lambda_u^T du,$$

we get

$$V_t^T=\int_0^t C_u^T du+\hat{M}_t^T$$

with

$$\hat{M}_t^T=\frac{1-a_T}{T}\left(N_{Tt}^T-\int_0^{Tt}\lambda_u^T du\right).$$

Thanks to Doob's inequality,

$$\mathbb{E}[(\sup_{t\in[0,1]}\hat{M}_t^T)^2]\leq 4\mathbb{E}[(\hat{M}_1^T)^2]\leq 4\left(\frac{1-a_T}{T}\right)^2\mathbb{E}[N_T^T]\leq\frac{4\mu(1-a_T)}{T}$$

which goes to 0, as $T\to\infty$. Using the first part of the theorem and these arguments, we conclude the result of convergence of the integrals. □

2.5 Hawkes based price model

In this section, we come back to the price model introduced in chapter 1 and we use the arguments developed in this chapter to deduce a theorem in the same spirit of Theorem 2.4.

Let $(P_t^T)_{t\geq 0}$ be the price process

$$P_t^T = \frac{1}{T}(N_{Tt}^{T+} - N_{Tt}^{T-}), \quad t \geq 0, \tag{2.5}$$

where (N^{T+}, N^{T-}) is a two dimensional Hawkes process with intensity

$$\begin{pmatrix} \lambda_t^{T+} \\ \lambda_t^{T-} \end{pmatrix} = \begin{pmatrix} \mu \\ \mu \end{pmatrix} + \int_0^t \begin{pmatrix} \phi_1^T(t-s) & \phi_2^T(t-s) \\ \phi_2^T(t-s) & \phi_1^T(t-s) \end{pmatrix} \begin{pmatrix} dN_s^{T+} \\ dN_s^{T-} \end{pmatrix},$$

with ϕ_1^T and ϕ_2^T two non negative measurable functions which satisfy the stability condition

$$\int_0^{+\infty} \phi_1^T(s)ds + \int_0^{+\infty} \phi_2^T(s)ds < 1.$$

Given this model, our goal is to move to large scales and find its limiting behavior when the stability condition is almost violated to get a result in the same regime of Theorem 2.4. We look at the following rescaling.

Assumption 6. For $i = 1, 2$ and $t \geq 0$, define

$$\phi_i^T(t) = a_T \phi_i(t)$$

where $(a_T)_{T\geq 0}$ is a sequence of numbers in $(0, 1)$, for all $T > 0$, such that $a_T \to 1$ as $T \to \infty$, and ϕ_1, ϕ_2 are non negative measurable functions such that

$$\int_0^{\infty} (\phi_1(s) + \phi_2(s))ds = 1,$$
$$\int_0^{\infty} s(\phi_1(s) + \phi_2(s))ds =: m < \infty.$$

Moreover, ϕ_1 and ϕ_2 are differentiable with $\|\phi_1'\|_\infty < \infty$, $\|\phi_2'\|_\infty < \infty$.

Remark. In this case, Assumption 6 has a financial interpretation, which is exactly the reason why limit theorems for nearly unstable Hawkes processes were studied. Indeed, the fact that the L^1 norm of the Hawkes kernels converges to one reproduces the high degree of endogeneity of the market, which means that large amounts of orders are just sent by other orders. This is completely consistent with the empirical data, as shown in [FS12], [FS15], [HBB13].

We can now state the theorem which describes the macroscopic dynamics.

Theorem 2.5. (Theorem 3.1 in [JR15]) *Assume $T(1 - a_T) \to \lambda$ as $T \to \infty$ and let $\phi = \phi_1 - \phi_2$. Under Assumption 6, the sequence of Hawkes based price models $(P_t^T)_{T>0}$,* *for $t \in [0,1]$, converges in law, for the Skorokhod topology, towards a Heston type process P on $[0,1]$ defined by*

$$\begin{cases} dC_t = (\frac{2\mu}{\lambda} - C_t)\frac{\lambda}{m}dt + \frac{1}{m}\sqrt{C_t}dB_t^1, \;\; C_0 = 0 \\ dP_t = \frac{1}{1-\|\phi\|_1}\sqrt{C_t}dB_t^2, \;\; P_0 = 0 \end{cases}$$

with (B^1, B^2) a 2-dimensional Brownian motion.

Remark. The proof of Theorem 2.5 is based on working with the process

$$C_t^T := \frac{\lambda_{Tt}^{T+} + \lambda_{Tt}^{T-}}{T}.$$

This process can be expressed in terms of some self-normalized martingales that converge to Brownian motions, in the same way as in the proof of Theorem 2.4. The idea is to apply Theorem 2.4 to get the CIR dynamics for the squared volatility, together with the fact that the price process (2.5) converges to the stochastic integral of the volatility. This gives exactly the Heston dynamics of the statement (see [Hes93] for details on the Heston model). We refer to [JR15] for the detailed proof.

Chapter 3

Rough dynamics

In this chapter we discuss the main theorems of [JR16], where the authors present a generalization of the results achieved in [JR15]. Indeed, as explained in this paper, the assumption to work with a kernel ϕ which has not only the L^1 norm convergent to one, but also a heavy tailed property seems to be more in agreement with financial data. So it is necessary to develop new results for the limiting behaviour of a sequence of the so called nearly unstable heavy tailed Hawkes processes. In this new framework, the dynamics achieved is different from the one discussed in chapter 2, since it is less regular and no more based on semi-martingales, but driven by fractional Brownian motion. The limiting distribution is a fractional version of the CIR dynamics. This result gives the possibility of working with rough volatility models in financial applications.

3.1 Nearly unstable heavy tailed Hawkes processes

In this section we introduce a new assumption to our processes to generalize the case presented in chapter 2. Indeed, we define a sequence of processes in the same way but then we introduce a more general assumption on their kernel.

Consider a sequence of Hawkes processes $(N_t^T)_{t\in[0,T]}$ defined, for each $T > 0$, by $N_0^T = 0$ and by the intensity process

$$\lambda_t^T = \mu^T + \int_{(0,t)} \phi^T(t-s)dN_s^T, \quad \forall t \in [0,T],$$

where $(\mu^T)_{T>0}$ is a sequence of positive real numbers, and $\phi^T : \mathbb{R}_+ \to \mathbb{R}_+$ is a non-negative measurable function satisfying the condition $\|\phi^T\|_{L^1} < \infty$. We assume to equip the probability space with a filtration $(\mathcal{F}_t^T)_{t\in[0,T]}$, where $\mathcal{F}_t^T = \sigma(\{N_s^T : s \leq t\})$.

Assumption 7. The kernel function ϕ^T is equal to

$$\phi^T(t) = a_T\phi(t), \quad \forall t \geq 0,$$

where $(a_T)_{T>0}$ is a sequence of real numbers with $a_T \in (0,1), \forall T > 0$, and such that $a_T \to 1$ as $T \to \infty$, while ϕ is a non-negative measurable function such that $\|\phi\|_{L^1} = 1$. Moreover,

$$\lim_{x\to\infty} \alpha x^\alpha \int_x^\infty \phi(s)ds = K$$

where $\alpha \in (0,1)$ and $K > 0$.

Remark. In Assumption 7 we impose to work with a function ϕ which satisfies

$$\phi(x) \sim_{x\to\infty} \frac{K}{x^{1+\alpha}},$$

which means that it has a power law tail. Notice that the condition of chapter 2

$$\int_0^\infty s\phi(s)ds < \infty$$

is no longer satisfied, so we need to develop new arguments to get similar results on the limiting behaviour of the sequence of Hawkes processes. On the other hand, we still assume that the L^1 norm of the kernel ϕ^T converges to one, so we consider again the almost unstable case. For these reasons, under Assumption 7, we use the expression *nearly unstable heavy tailed Hawkes processes*.

Under Assumption 7 we will find a different scaling behaviour respect to the CIR dynamics obtained in the previous chapter. In particular, this regime allows for the description of a model that better reflects financial data and it will lead to the rough volatility framework.

Recall the usual notation: we denote by M^T the martingale associated to N^T

$$M_t^T = N_t^T - \int_0^t \lambda_s^T ds, \quad \forall t \in [0, T],$$

and by ψ^T the non-negative function

$$\psi^T(t) = \sum_{k=1}^{\infty} (\phi^T)^{*k}(t).$$

Notice that ψ^T is well defined since $\|\phi^T\|_{L^1} < 1$.

3.2 Heuristic derivation of the asymptotic dynamics

In this section we want to achieve the asymptotic dynamics of the sequence of nearly unstable heavy tailed Hawkes processes through heuristic arguments, following the same approach used in section 2.3.

First, recall that thanks to Proposition 2.2 we have the following expression for the intensity process:

$$\lambda_{Tt}^T = \mu^T + \int_0^{Tt} \psi^T(Tt - s)\mu^T\, ds + \int_0^{Tt} \psi^T(Tt - s)\, dM_s^T, \quad \forall t \geq 0.$$

Similarly to chapter 2, we can consider the rescaling of the intensity process

$$C_t^T = \frac{1 - a_T}{\mu^T} \lambda_{Tt}^T, \quad t \in [0, 1],$$

and doing exactly the same computations as in section 2.3 we find, for $t \in [0, 1]$,

$$C_t^T = (1 - a_T) + \int_0^t T(1 - a_T)\psi^T(Ts)ds + \sqrt{\frac{T(1 - a_T)}{\mu^T}} \int_0^t \psi^T(T(t - s))\sqrt{C_s^T} dB_s^T$$

where

$$B_t^T = \frac{1}{\sqrt{T}} \int_0^{Tt} \frac{dM_s^T}{\sqrt{\lambda_s^T}}.$$

Recall that, as $T \to \infty$, the process $(B_t^T)_{t\in[0,1]}$ converges in law for the Skorokhod topology towards a standard Brownian motion $(B_t)_{t\in[0,1]}$, as shown in the proof of Theorem 2.4.

To find the limiting behaviour of the process C^T it is necessary to study the behaviour of the function $x \mapsto \psi^T(Tx)$. Hence, we introduce the function

$$\rho^T(x) = \frac{T}{\|\psi^T\|_{L^1}} \psi^T(Tx), \quad x \geq 0,$$

which is the density function of the random variable

$$X^T = \frac{1}{T} \sum_{j=1}^{I^T} X_j$$

where $(X_j)_{j\geq 1}$ is a sequence of iid random variables with density ϕ and I^T is a geometric random variable with parameter $1-a_T$ independent of $(X_j)_{j\geq 1}$. To derive the convergence of this density, in this case we can look at the Laplace transform $\hat{\rho}^T(z)$ of the random variable X^T, and we find, for $z \geq 0$,

$$\hat{\rho}^T(z) = \mathbb{E}[e^{-zX^T}] = \frac{\hat{\phi}(\frac{z}{T})}{1 - \frac{a_T}{1-a_T}(\hat{\phi}(\frac{z}{T}) - 1)}, \tag{3.1}$$

where $\hat{\phi}$ denotes the Laplace transform of the function ϕ. At this point we need to compute the expansion for the function $\hat{\phi}$.

Remark. This is the point in which there is the main difference from chapter 2, since there is no longer the assumption that $m = \int_0^\infty s\phi(s)ds$ is finite, but only the power law tail for the function ϕ. So the expansion of $\hat{\phi}$ will be different and will lead to a different asymptotic dynamics for C^T.

Introduce the following function

$$F(x) = \int_0^x \phi(s)ds, \quad x \geq 0.$$

Integrating by parts, we have

$$\begin{aligned}\hat{\phi}(z) = \mathbb{E}[e^{-zX_1}] &= \int_0^\infty e^{-zt}\phi(t)dt \\ &= z\int_0^\infty e^{-zt}F(t)dt = 1 - z\int_0^\infty e^{-zt}(1-F(t))dt.\end{aligned}$$

The idea is to apply now Theorem A.9 to the integral $\int_0^\infty e^{-zt}(1-F(t))dt$. Indeed, by Assumption 7 it holds

$$1 - F(x) = \int_x^\infty \phi(t)dt \sim_{x\to\infty} \frac{K}{\alpha x^\alpha}$$

with $\alpha \in (0,1)$, which can be written in the notation of the theorem as

$$\frac{K}{\alpha x^\alpha} = cx^\rho \frac{L(x)}{\Gamma(1+\rho)}$$

if and only if $\rho = -\alpha > -1$, L is the constant function equal to 1 and $c = \frac{K\Gamma(1-\alpha)}{\alpha}$. Therefore, applying Theorem A.9, we get

$$\int_0^\infty e^{-zt}(1-F(t))dt = \frac{K\Gamma(1-\alpha)}{\alpha} z^{\alpha-1} + o(z^{\alpha-1}).$$

Replacing this expression in $\hat{\phi}$ we get

$$\hat{\phi}(z) = 1 - \frac{K\Gamma(1-\alpha)}{\alpha} z^\alpha + o(z^\alpha).$$

Define the following quantities

$$\delta = \frac{K\Gamma(1-\alpha)}{\alpha}, \quad v_T = \frac{1}{\delta} T^\alpha (1-a_T).$$

We can replace these expressions in (3.1) to get

$$\hat{\rho}^T(z) = \frac{1 - \delta\left(\frac{z}{T}\right)^\alpha + o(\left(\frac{z}{T}\right)^\alpha)}{1 + \frac{a_T}{1-a_T}\delta\left(\frac{z}{T}\right)^\alpha + o(\left(\frac{z}{T}\right)^\alpha)}$$

which gives

$$\hat{\rho}^T(z) \sim_{T\to\infty} \frac{v_T}{v_T + z^\alpha}.$$

In this way we have obtained the asymptotic behaviour of the Laplace transform $\hat{\rho}^T$, but we can also deduce the convergence of the density functions, since $\frac{v_T}{v_T+z^\alpha}$ is the Laplace transform of the function

$$v_T x^{\alpha-1} E_{\alpha,\alpha}(-v_T x^\alpha)$$

where $E_{\alpha,\alpha}$ is a Mittag-Leffler function (see Appendix A.5). So we can conclude that $\rho^T(x) \sim_{T\to\infty} v_T x^{\alpha-1} E_{\alpha,\alpha}(-v_T x^\alpha)$.

Coming back to the expression of C^T we can write, for $t \in [0,1]$,

$$C_t^T = (1-a_T) + \int_0^t a_T \rho^T(s)ds + \frac{a_T}{\sqrt{T\mu^T(1-a_T)}} \int_0^t \rho^T(t-s)\sqrt{C_s^T}dB_s^T$$

where

$$B_t^T = \frac{1}{\sqrt{T}} \int_0^{Tt} \frac{dM_s^T}{\sqrt{\lambda_s^T}}.$$

Using the convergence developed above, we have, for $t \in [0,1]$,

$$\begin{aligned} C_t^T \sim_{T\to\infty} \; & v_T \int_0^t s^{\alpha-1} E_{\alpha,\alpha}(-v_T s^\alpha)ds \\ & + \frac{v_T}{\sqrt{T\mu^T(1-a_T)}} \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-v_T(t-s)^\alpha)\sqrt{C_s^T}dB_s^T. \end{aligned}$$

Recall that, as $T \to \infty$, the process $(B_t^T)_{t\in[0,1]}$ converges in law for the Skorokhod topology towards a standard Brownian motion $(B_t)_{t\in[0,1]}$. Denote by v the limit of v_T, and γ the limit of $\frac{1}{\sqrt{T\mu^T(1-a_T)}}$. We finally can pass to the limit (not rigorously) and find, for $t \in [0,1]$,

$$C_t^\infty = v\int_0^t s^{\alpha-1}E_{\alpha,\alpha}(-vs^\alpha)ds + \gamma v\int_0^t (t-s)^{\alpha-1}E_{\alpha,\alpha}(-v(t-s)^\alpha)\sqrt{C_s^\infty}dB_s. \tag{3.2}$$

Equation (3.2) describes a fractional version of the dynamics that we have found in the non heavy tailed case. The term $x^{\alpha-1}E_{\alpha,\alpha}(-vx^\alpha)$ is the one that gives the roughness to the model and this will be clear when we will prove in a rigorous way the limiting dynamics. Indeed, the kernel $x^{\alpha-1}$ gives the same regularity of a fractional Brownian motion, that in the Mandelbrot-van Ness representation (see [MVN68]) can be written as

$$B_t^H = \frac{1}{\Gamma(H+\frac{1}{2})}\left(\int_0^t (t-s)^{H-\frac{1}{2}}dW_s + \int_{-\infty}^0 ((t-s)^{H-\frac{1}{2}} - (-s)^{H-\frac{1}{2}})dW_s\right), \tag{3.3}$$

setting in our case $\alpha = H + \frac{1}{2}$.

To state the following theorems and derive more rigorous arguments, we need the following assumption.

Assumption 8. There exist two positive constants λ and μ^* such that

$$\lim_{T\to\infty} T^\alpha(1-a_T) = \lambda\delta,$$
$$\lim_{T\to\infty} T^{1-\alpha}\mu^T = \frac{\mu^*}{\delta}.$$

Assumption 8 implies that $v_T \to \lambda$ as $T \to \infty$, and so the sequence $(X^T)_{T>0}$ of random variables converges in law towards the random variable with density

$$\lambda x^{\alpha-1}E_{\alpha,\alpha}(-\lambda x^\alpha),$$

which is the Mittag-Leffler density function denoted by $f^{\alpha,\lambda}$ in Appendix A.5.

3.3 Work with rescaled martingales

We define the following processes, for $t \in [0,1]$,

$$X_t^T = \frac{\delta(1-a_T)}{T^\alpha\mu^*}N_{Tt}^T,$$
$$\Lambda_t^T = \frac{\delta(1-a_T)}{T^\alpha\mu^*}\int_0^{Tt}\lambda_s^T ds,$$
$$Z_t^T = \sqrt{\frac{T^\alpha\mu^*}{\delta(1-a_T)}}(X_t^T - \Lambda_t^T).$$

X_t^T is the renormalized Hawkes process, Λ_t^T its integrated intensity, and Z_t^T is the associated martingale.

Remark. The reason why we take this rescaling is to work with processes whose expectation has magnitude one, for example in the case of stationary processes:

$$\begin{aligned}\mathbb{E}[X_1^T] &= \frac{\delta(1-a_T)}{T^\alpha \mu^*}\mathbb{E}[N_T^T] = \frac{\delta(1-a_T)}{T^\alpha \mu^*} T\mathbb{E}[\lambda_T^T] \\ &= \frac{\delta(1-a_T)}{T^\alpha \mu^*}\frac{T\mu^T}{1-a_T} \sim_{T\to\infty} 1.\end{aligned}$$

We state a first theorem that gives the convergence of the sequence $(Z^T, X^T)_T$ using martingale arguments.

Theorem 3.1. (Proposition 3.2 in [JR16]) *Under Assumptions 7 and 8, the following hold:*

1. *the sequence $(Z^T, X^T)_T$ is tight;*
2. *if (Z, X) is a limit point of $(Z^T, X^T)_T$, then Z is a continuous martingale and $[Z, Z] = X$.*

Proof. Proof of (1). The first goal is to prove that the sequences $(X^T)_T$ and $(\Lambda^T)_T$ are C-tight (see Definition A.1). By Theorem A.2, it is sufficient to prove the tightness of the processes and the fact that the amplitude of their jumps converges to zero in probability. First, notice that

$$\mathbb{E}[N_t^T] = \mu^T t + \mu^T \int_0^t \psi^T(t-s)sds \leq t\mu^T(1 + \|\psi^T\|_{L^1})$$

and

$$\|\psi^T\|_{L^1} \leq \frac{c}{1-a_T}.$$

So at the end we find

$$\mathbb{E}[X_1^T] = \mathbb{E}[\Lambda_1^T] \leq c.$$

Using that X_t^T is increasing, for any $\varepsilon > 0$, there exists a constant K such that

$$\mathbb{P}(\sup_{t\in[0,1]} |X_t^T| > K) = \mathbb{P}(|X_1^T| > K) \leq \frac{\mathbb{E}[X_1^T]}{K} \leq \varepsilon.$$

Using always the fact that these processes are increasing, we also have that for any $\varepsilon > 0$, $\eta > 0$, there exists $\theta > 0$ such that

$$\mathbb{P}(\omega'(X^T, \theta) \geq \eta) \leq \varepsilon$$

where

$$\omega'(X^T, \theta) = \inf\{\max_{i\leq I} \omega(X^T, [t_{i-1}, t_i)) : 0 = t_0 < \ldots < t_I = 1, \inf_{i\leq I}(t_i - t_{i-1}) \geq \theta\}$$

and

$$\omega(X^T, [t_{i-1}, t_i)) = \sup_{s,t\in[t_{i-1},t_i)} |X_t^T - X_s^T|.$$

The same holds for the process Λ^T, so we get the tightness of $(X^T)_T$ and $(\Lambda^T)_T$.

Looking at the jumps, we can use that

$$\frac{1-a_T}{T^\alpha} \xrightarrow[T\to\infty]{} 0$$

Since N^T has bounded jumps and Λ^T is continuous, we get that the maximum jump size of both X^T and Λ^T goes to zero as $T \to \infty$. This gives the C-tightness of the processes $(X^T)_T$ and $(\Lambda^T)_T$.

Finally, we can achieve the tightness of the sequence $(Z^T)_T$ applying Theorem A.3, that relates the tightness of a sequence of martingales to the C-tightness of the sequence of the brackets. In our case, the sequence $(\Lambda^T)_T$ is C-tight, so we can conclude that the sequence of martingales $(Z^T)_T$ is tight. From the marginal tightness of $(X^T)_T$ and $(Z^T)_T$, we deduce also the joint tightness of $(Z^T, X^T)_T$.

Proof of (2). Consider a subsequence $(Z^{T_n}, X^{T_n})_n$ converging to a limit (Z, X). We know that $(Z^{T_n})_n$ is a sequence of martingales with bounded jumps that converges in law towards Z by hypothesis, and the sequence $([Z^{T_n}, Z^{T_n}])_n = (X^{T_n})_n$ is tight by the previous point. Hence we can apply Theorem A.4 to conclude that

$$[Z^{T_n}, Z^{T_n}]_n \xrightarrow[n\to\infty]{} [Z, Z] \quad \text{in law},$$

and so $X = [Z, Z]$.

Finally, Z is a continuous martingale. Indeed, the continuity follows by the definition of C-tightness for the sequence $(Z^{T_n})_n$. The fact that Z is a martingale is given by the fact that it is the limit of a sequence of martingales with bounded jumps. This concludes the proof. □

3.4 Fractional CIR dynamics

Let (Z, X) be a couple of processes such that $(Z^T, X^T)_T$ converges in law to (Z, X). In the following theorem, we state in a formal way the heuristic results presented above.

Theorem 3.2. (Theorem 3.1 in [JR16]) *There exists a Brownian motion B such that, for $t \in [0, 1]$,*

$$Z_t = B_{X_t},$$

and, for any $\varepsilon > 0$, X is continuous with Hölder regularity $(1 \wedge 2\alpha) - \varepsilon$ on $[0, 1]$ and satisfies

$$X_t = \int_0^t s f^{\alpha,\lambda}(t-s)ds + \frac{1}{\sqrt{\mu^*\lambda}} \int_0^t f^{\alpha,\lambda}(t-s) B_{X_s} ds. \tag{3.4}$$

Proof. We propose the proof of the dynamics (3.4), whereas for the proof of the Hölder regularity see [JR16].

By Skorokhod's representation theorem, we can assume that the sequence $(Z^T, X^T)_T$ converges to (Z, X) almost surely. Since Z and X are continuous, we have that

$$\sup_{t\in[0,1]} |X_t^T - X_t| \xrightarrow[T\to\infty]{} 0, \quad \sup_{t\in[0,1]} |Z_t^T - Z_t| \xrightarrow[T\to\infty]{} 0. \tag{3.5}$$

We start by working with the integrated intensity. For all $t \geq 0$, we have:

$$\begin{aligned}\int_0^t \lambda_s^T ds &= t\mu^T + \int_0^t \int_0^s \phi^T(s-u) dM_u^T ds + \int_0^t \int_0^s \phi^T(s-u)\lambda_u^T du ds \\ &= t\mu^T + \int_0^t \int_0^{t-u} \phi^T(v) dv dM_u^T + \int_0^t \int_0^{t-u} \phi^T(v) dv \lambda_u^T du \\ &= t\mu^T + \int_0^t \phi^T(t-u) M_u^T du + \int_0^t \phi^T(t-u)(\int_0^u \lambda_v^T dv) du,\end{aligned}$$

where in the last computation we have used integration by parts. Similarly to section 2.3, we have that

$$\begin{aligned}\int_0^t \psi^T(t-s)\left(\int_0^s \phi^T(s-u) M_u^T du\right) ds &= \int_0^t \left(\int_0^{t-u} \psi^T(t-u-s)\phi^T(s) ds\right) M_u^T du \\ &= \int_0^t (\psi^T * \phi^T(t-u)) M_u^T du \\ &= \int_0^t \psi^T(t-u) M_u^T du - \int_0^t \phi^T(t-u) M_u^T du.\end{aligned}$$

Using this relation and Lemma A.1 with the function $h(t) = t\mu^T + \int_0^t \phi^T(t-u) M_u^T du$, we find

$$\int_0^t \lambda_s^T ds = t\mu^T + \int_0^t \psi^T(t-s) s\mu^T ds + \int_0^t \psi^T(t-s) M_s^T ds.$$

So we can obtain an expression for Λ_t^T:

$$\begin{aligned}\Lambda_t^T &= \frac{\delta(1-a_T)}{T^\alpha \mu^*} \int_0^{Tt} \lambda_s^T ds \\ &= \frac{\delta(1-a_T)}{T^\alpha \mu^*}\left(Tt\mu^T + \int_0^{Tt} \psi^T(Tt-s) s\mu^T ds + \int_0^{Tt} \psi^T(Tt-s) M_s^T ds\right) \\ &= (1-a_T)tu_T + T(1-a_T)u_T \int_0^t \psi^T(T(t-s)) s ds + \int_0^{Tt} \psi^T(Tt-s)\frac{\delta(1-a_T)}{T^\alpha \mu^*} M_s^T ds,\end{aligned}$$

where we have defined the term

$$u_T = \frac{\delta \mu^T}{T^{\alpha-1}\mu^*}.$$

Furthermore, using the definition of Z^T, we get

$$\begin{aligned}\Lambda_t^T = &(1-a_T)tu_T + T(1-a_T)u_T \int_0^t \psi^T(T(t-s)) s ds \\ &+ T^{1-\frac{\alpha}{2}} \sqrt{\frac{\delta(1-a_T)}{\mu^*}} \int_0^t \psi^T(T(t-s)) Z_s^T ds.\end{aligned}$$

At this point, we can look separately at the behaviour of the three addends, as $T \to \infty$. Using that $u_T \to 1$, the first term goes to zero:

$$T_1 := (1-a_T)tu_T \xrightarrow[T\to\infty]{} 0.$$

Looking at the second term, we have

$$\begin{aligned}
T_2 &:= T(1-a_T)u_T \int_0^t \psi^T(T(t-s))sds \\
&= a_T u_T \int_0^t \frac{T\psi^T(T(t-s))}{\|\psi^T\|_{L^1}} sds \\
&= a_T u_T \int_0^t \rho^T(t-s)sds.
\end{aligned}$$

Now, we can use that, as $T \to \infty$, the term $\rho^T(x)$ converges weakly towards the density

$$f^{\alpha,\lambda}(x) = \lambda x^{\alpha-1} E_{\alpha,\alpha}(-\lambda x^\alpha). \tag{3.6}$$

In particular,

$$F^T(t) := \int_0^t \rho^T(s)ds$$

converges uniformly to

$$F^{\alpha,\lambda}(t) = \int_0^t f^{\alpha,\lambda}(s)ds$$

(see [JR16] for details). Using this fact and integrating by parts in the expression of T_2, we get that the term T_2 converges uniformly to

$$\int_0^t F^{\alpha,\lambda}(t-s)ds = \int_0^t f^{\alpha,\lambda}(t-s)sds.$$

Finally we focus on the last term of Λ^T, that can be written as

$$T_3 := \frac{a_T\sqrt{\delta}}{\sqrt{T^\alpha(1-a_T)\mu^*}} \int_0^t \rho^T(t-s)Z_s^T ds.$$

Now, integrating by parts and using that X^T is piecewise constant, we get pathwise:

- $\displaystyle\int_0^t \rho^T(t-s)Z_s^T ds = \int_0^t F^T(t-s)dZ_s^T,$
- $\displaystyle\int_0^t f^{\alpha,\lambda}(t-s)Z_s^T ds = \int_0^t F^{\alpha,\lambda}(t-s)dZ_s^T.$

Looking at the two terms in the right hand side in the previous expressions, by Ito's formula we have

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^t (F^{\alpha,\lambda}(t-s) - F^T(t-s))dZ_s^T\right)^2\right] &= \mathbb{E}\left[\int_0^t (F^{\alpha,\lambda}(t-s) - F^T(t-s))^2 dX_s^T\right] \\
&\leq c\int_0^t (F^{\alpha,\lambda}(t-s) - F^T(t-s))^2 ds,
\end{aligned}$$

which goes to zero. Moreover,

$$\int_0^t f^{\alpha,\lambda}(t-s)|Z_s - Z_s^T|ds \xrightarrow[T\to\infty]{} 0,$$

thanks to (3.5). Hence, using that by assumption $T^\alpha(1-a_T) \to \lambda\delta$, and putting together the previous convergences, the term T_3 converges in law to

$$\frac{1}{\sqrt{\lambda\mu^*}} \int_0^t f^{\alpha,\lambda}(t-s) Z_s ds.$$

Hence we get that Λ_t^T, $t \in [0,1]$, converges in law towards

$$\int_0^t s f^{\alpha,\lambda}(t-s) ds + \frac{1}{\sqrt{\lambda\mu^*}} \int_0^t f^{\alpha,\lambda}(t-s) Z_s ds.$$

To find equation (3.4), we need the following theorem:

Theorem 3.3. (Dambis-Dubins-Schwarz) *Let M be a continuous local martingale in $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, P)$ vanishing at 0 and with brackets that satisfy $\langle M, M\rangle_\infty = \infty$. Let T_t be defined as*

$$T_t = \inf\{s : \langle M, M\rangle_s > t\}.$$

Then $B_t = M_{T_t}$ is a $(\mathcal{F}_{T_t})$-Brownian motion and $M_t = B_{\langle M,M\rangle_t}$.

In our case, Z is a continuous martingale vanishing at 0, with $[Z, Z]_\infty = X_\infty = \infty$, so we can apply Dambis-Dubins-Schwarz theorem to conclude that there exists a Brownian motion such that, for $t \geq 0$,

$$Z_t = B_{[Z,Z]_t} = B_{X_t}.$$

At the end, we get that Λ_t^T, $t \in [0,1]$, converges in law towards

$$\int_0^t s f^{\alpha,\lambda}(t-s) ds + \frac{1}{\sqrt{\lambda\mu^*}} \int_0^t f^{\alpha,\lambda}(t-s) B_{X_s} ds. \tag{3.7}$$

Now, notice that the sequence of martingales $X^T - \Lambda^T$ converges to zero in probability, uniformly on $[0,1]$, as a consequence of Doob's inequality:

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,1]} (X_t^T - \Lambda_t^T)^2\right] &= \left(\frac{\delta(1-a_T)}{T^\alpha \mu^*}\right)^2 \mathbb{E}\left[\sup_{t\in[0,1]} (M_{Tt}^T)^2\right] \\
&\leq 4\left(\frac{\delta(1-a_T)}{T^\alpha \mu^*}\right)^2 \mathbb{E}\left[(M_T^T)^2\right] \\
&\leq c\left(\frac{1-a_T}{T^\alpha}\right)^2 \mathbb{E}\left[N_T^T\right] \\
&\leq c\frac{1-a_T}{T^\alpha} \xrightarrow[T\to\infty]{} 0.
\end{aligned}$$

Therefore, thanks to this property, we get also the dynamics of the limit of the sequence X^T. In particular, since X^T converges towards X as $T \to \infty$, we finally deduce

$$X_t = \int_0^t s f^{\alpha,\lambda}(t-s) ds + \frac{1}{\sqrt{\lambda\mu^*}} \int_0^t f^{\alpha,\lambda}(t-s) B_{X_s} ds.$$

□

In the next result, we restrict to look at the parameter $\alpha \in (\frac{1}{2}, 1)$, and we state that under this assumption we can take the derivative of the process X. It is the main theorem that gives the roughness of the model.

Theorem 3.4. (Theorem 3.2 in [JR16]) *Let $(X_t)_t$ be a process satisfying (3.4) for $t \in [0,1]$ and assume that $\alpha > \frac{1}{2}$. Then X is differentiable on $[0,1]$ and its derivative Y satisfies*

$$Y_t = F^{\alpha,\lambda}(t) + \frac{1}{\sqrt{\mu^* \lambda}} \int_0^t f^{\alpha,\lambda}(t-s)\sqrt{Y_s} d\bar{B}_s, \tag{3.8}$$

with $\bar{B}$ a Brownian motion. Furthermore, for any $\varepsilon > 0$, Y has Hölder regularity $\alpha - \frac{1}{2} - \varepsilon$.

Remark. We can notice the following facts arising by this theorem:

- In our model, the process X represents the integrated volatility, and thanks to Theorem 3.4 we have that if $\alpha \in (\frac{1}{2}, 1)$, then we can take the derivative and find the dynamics of the spot volatility Y.
- The stochastic differential equation (3.8) reflects the dynamics obtained with heuristic arguments in equation (3.2) and it is the fractional version of the CIR dynamics.
- Theorem 3.4 gives also the roughness of the model since it gives the Hölder regularity of the volatility. In particular, setting $H = \alpha - \frac{1}{2}$, we can notice that the Hölder regularity is very low when H is close to one, leading to a very rough process. This is exactly the regime that arises in financial data, as shown in several works about rough volatility modeling (see [GJR18], [EEFR18]).

Proof. In this proof, we use the definitions of fractional integrals and derivatives, see Appendix A.6.

Assume that X satisfies equation (3.4). First, using Proposition A.11 with the martingale Z and the function $f^{\alpha,\lambda}$, we can say that, for any $\nu \in (0, \alpha)$, $D^\nu f^{\alpha,\lambda}$ exists and it holds

$$\int_0^t f^{\alpha,\lambda}(t-s) Z_s ds = \int_0^t D^\nu f^{\alpha,\lambda}(t-s) I^\nu Z_s ds.$$

So equation (3.4) becomes

$$X_t = \int_0^t s f^{\alpha,\lambda}(t-s) ds + \frac{1}{\sqrt{\lambda \mu^*}} \int_0^t D^\nu f^{\alpha,\lambda}(t-s) I^\nu Z_s ds. \tag{3.9}$$

Moreover, by Proposition A.10, we have that, for any $\nu > \frac{1}{2}$ fixed, $D^{1-\nu} Z$ is also well defined. Then, by definition we have

$$\begin{aligned} I^\nu Z_t &= \frac{1}{\Gamma(\nu)} \int_0^t \frac{Z_u}{(t-u)^{1-\nu}} du \\ &= \frac{1}{\Gamma(\nu)} \int_0^t \frac{d}{ds} \left(\int_0^s \frac{Z_u}{(s-u)^{1-\nu}} du \right) ds \\ &= \int_0^t D^{1-\nu} Z_s ds. \end{aligned}$$

So we can use this relation together with Fubini theorem to get

$$\begin{aligned}\int_0^t D^\nu f^{\alpha,\lambda}(t-s) I^\nu Z_s ds &= \int_0^t D^\nu f^{\alpha,\lambda}(t-s) \left(\int_0^s D^{1-\nu} Z_u du \right) ds \\ &= \int_0^t \int_u^t D^\nu f^{\alpha,\lambda}(t-s) D^{1-\nu} Z_u ds du,\end{aligned}$$

and doing the change of variables $t - s + u = r$,

$$\begin{aligned}\int_0^t D^\nu f^{\alpha,\lambda}(t-s) I^\nu Z_s ds &= \int_0^t \int_u^t D^\nu f^{\alpha,\lambda}(r-u) D^{1-\nu} Z_u dr du \\ &= \int_0^t \int_0^r D^\nu f^{\alpha,\lambda}(r-u) D^{1-\nu} Z_u du dr.\end{aligned}$$

Moreover, thanks to integration by parts in the first integral in (3.9), we find

$$\int_0^t s f^{\alpha,\lambda}(t-s) ds = \int_0^t \left(\int_0^{t-s} f^{\alpha,\lambda}(u) du \right) ds = \int_0^t F^{\alpha,\lambda}(s) ds.$$

So at the end we have

$$X_t = \int_0^t F^{\alpha,\lambda}(s) ds + \frac{1}{\sqrt{\lambda \mu^*}} \int_0^t \int_0^s D^\nu f^{\alpha,\lambda}(s-u) D^{1-\nu} Z_u du ds,$$

which can be written as

$$X_t = \int_0^t Y_s ds,$$

with

$$Y_s = F^{\alpha,\lambda}(s) + \frac{1}{\sqrt{\lambda \mu^*}} \int_0^s D^\nu f^{\alpha,\lambda}(s-u) D^{1-\nu} Z_u du.$$

In particular, it follows from previous results that Y has Hölder regularity $\alpha - \frac{1}{2} - \varepsilon$, and X is differentiable with derivative Y.

We can now prove equation (3.8). First, applying a stochastic version of Fubini theorem for martingales, we can write

$$\begin{aligned}D^{1-\nu} Z_s &= \frac{1}{\Gamma(\nu)} \frac{d}{ds} \int_0^s \frac{Z_v}{(s-v)^{1-\nu}} dv \\ &= \frac{1}{\Gamma(\nu)} \frac{d}{ds} \int_0^s \int_0^v \frac{1}{(s-v)^{1-\nu}} dZ_u dv \\ &= \frac{1}{\Gamma(\nu)} \frac{d}{ds} \int_0^s \int_u^s \frac{1}{(s-v)^{1-\nu}} dv dZ_u \\ &= \frac{1}{\Gamma(\nu+1)} \frac{d}{ds} \int_0^s (s-u)^\nu dZ_u.\end{aligned}$$

Hence Y_t satisfies

$$Y_t = F^{\alpha,\lambda}(t) + \frac{1}{\sqrt{\lambda \mu^*}} \int_0^t \frac{1}{\Gamma(\nu+1)} D^\nu f^{\alpha,\lambda}(t-s) \frac{d}{ds} \int_0^s (s-u)^\nu dZ_u ds.$$

We want now to use the property of convolution $f * (g') = (f * g)'$, with the functions $f(x) := D^\nu f^{\alpha,\lambda}(x)$ and $g(x) := \int_0^x (x-u)^\nu dZ_u$, so we get

$$Y_t = F^{\alpha,\lambda}(t) + \frac{1}{\sqrt{\lambda\mu^*}}\frac{d}{dt}\int_0^t \frac{1}{\Gamma(\nu+1)} D^\nu f^{\alpha,\lambda}(t-s)\int_0^s (s-u)^\nu dZ_u ds.$$

By Fubini theorem and using the definitions of fractional integrals and derivatives, we get

$$\begin{aligned}
Y_t &= F^{\alpha,\lambda}(t) + \frac{1}{\sqrt{\lambda\mu^*}}\frac{d}{dt}\int_0^t \int_u^t \frac{1}{\Gamma(\nu+1)} D^\nu f^{\alpha,\lambda}(t-s)(s-u)^\nu ds dZ_u \\
&= F^{\alpha,\lambda}(t) + \frac{1}{\sqrt{\lambda\mu^*}}\frac{d}{dt}\int_0^t I^{\nu+1}(D^\nu f^{\alpha,\lambda}(t-u)) dZ_u \\
&= F^{\alpha,\lambda}(t) + \frac{1}{\sqrt{\lambda\mu^*}}\frac{d}{dt}\int_0^t \int_0^v I^\nu D^\nu f^{\alpha,\lambda}(v-u) dZ_u dv \\
&= F^{\alpha,\lambda}(t) + \frac{1}{\sqrt{\lambda\mu^*}}\int_0^t f^{\alpha,\lambda}(t-u) dZ_u.
\end{aligned}$$

We can now apply A.6 to the continuous martingale Z. We know that

$$[Z,Z]_t = X_t = \int_0^t Y_s ds.$$

So by Theorem A.6, there exists a Brownian motion $\bar{B}$ such that

$$Z_t = \int_0^t \sqrt{Y_s} d\bar{B}_s.$$

Hence in the end we find the dynamics in (3.8):

$$Y_t = F^{\alpha,\lambda}(t) + \frac{1}{\sqrt{\lambda\mu^*}}\int_0^t f^{\alpha,\lambda}(t-s)\sqrt{Y_s} d\bar{B}_s.$$

□

Chapter 4

Market impact modeling

The last chapter of the thesis is dedicated to presenting some applications of the limit theorems developed in previous chapters to market impact modeling. The execution of large amounts of orders in the market affects price dynamics, since buy orders generally cause prices to rise and sell orders cause them to fall. This information can influence market participants in their investment strategies and profit predictions. Our aim is to explain this relationship rigorously.

We present a model that simulates market transactions, working under the general assumption of no-statistical arbitrage, as discussed in [Jai15] and [JR20]. The goal is to describe the graph of the market impact function with respect to the transaction execution time. To achieve this, we first build a high frequency price model that replicates the stylized facts of modern market microstructure and converges in the long run to the rough Heston model. This method allows us to merge the Hawkes based modeling at microscopic level with the rough volatility framework at macroscopic level, exploiting the potential of both the two approaches.

4.1 Modeling buy and sell orders in the market

In this section we need first to introduce some notions. We call *market impact* the connection between an incoming order and the consequent price change. We want to reproduce the market impact considering each transaction of the market. Indeed, we can say that on average a buy order causes an upward movement of the price, while a sell order a downward change. This is an intrinsic property of the market, and it is important to model it since it produces an execution cost that has to be estimated since it tends to decrease the profits of investment strategies. On the other hand, in practice it is better to consider a large amount of orders and look at their cumulative impact. We call *metaorder* a large amount of transactions, and we will work with it to find its market impact. In practice, we consider investors that want to buy or sell stocks, and the volume that they wish to execute is much larger than the liquidity that is available in the order book. Thus they need to split their orders into small transactions that they execute over a given period of time denoted by T. The set of all these transactions is called a metaorder.

A natural way to define the market impact of a metaorder with volume V is to take the average of the difference between the price before and after the execution of the metaorder. In practice, we will look at

$$MI(t) := \mathbb{E}[P_t - P_0], \quad t \geq 0. \tag{4.1}$$

Hence, the first tool we need is a model for the price process. In this section we present a Hawkes based approach to describe the transactions that occur in the market, following the one in [Jai15]. In particular, we want to simulate the flow of market orders reproducing exactly all the movements from a microscopic point of view. We have to say that in general the transactions of market participants can be of two types: limit orders, characterized by a price at which the participant wants to buy or sell, and market orders, that are instantaneous transactions at the best available price. We only consider market orders in our model.

Let us consider two independent Hawkes processes, N^a and N^b, on the same probability space with intensities λ^a and λ^b respectively, defined as

$$\lambda_t^a = \mu + \int_{(0,t)} \phi(t-s) dN_s^a,$$
$$\lambda_t^b = \mu + \int_{(0,t)} \phi(t-s) dN_s^b,$$

where μ is a positive constant, and $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ is a non-negative measurable function satisfying the stability condition

$$\|\phi\|_{L^1} = \int_0^\infty \phi(t) dt < 1.$$

These two independent Hawkes processes count exactly all the transactions in the market. Indeed, we impose that N^a counts the buy orders, while N^b the sell orders.

As usual, we denote by M^a and M^b the martingales associated respectively to N^a and N^b:

$$M_t^a = N_t^a - \int_0^t \lambda_s^a ds, \quad M_t^b = N_t^b - \int_0^t \lambda_s^b ds, \quad \forall t \geq 0.$$

We define ψ the non-negative function

$$\psi(t) = \sum_{k=1}^\infty (\phi)^{*k}(t).$$

Notice that ψ is well defined since $\|\phi\|_{L^1} < 1$.

We can now define a price process, starting with some natural assumptions. We observe the price process in the time interval $[0, \tau]$, that can be thought as the length of a day, and denote by V_t^a, V_t^b the cumulated volumes of market orders at the bid and ask side of the market respectively at time $t \in [0, \tau]$. First, to relate the volumes with the Hawkes processes, we can set

$$V_\tau^a = \sum_{i=1}^{N_\tau^a} v_i^a, \quad V_\tau^b = \sum_{i=1}^{N_\tau^b} v_i^b,$$

where N^a_τ (resp. N^b_τ) is the number of buy (resp. sell) metaorders per day, and v^a_i (resp. v^b_i) is the volume of the i-th buy (resp. sell) metaorder. We can assume that the price depends linearly on the volume of the transactions, which means that, if P_0 denotes the initial price and P_τ the final price, we have

$$P_\tau = P_0 + k\left(\sum_{i=1}^{N^a_\tau} v^a_i - \sum_{i=1}^{N^b_\tau} v^b_i\right) = P_0 + k(V^a_\tau - V^b_\tau),$$

for $k > 0$. Finally, we can ask that the price is a martingale to achieve that at time t we can find the price looking at $\mathbb{E}[P_\tau|\mathcal{F}_t]$, and then send τ to infinity, so we get

$$P_t = P_0 + k \lim_{\tau\to\infty} \mathbb{E}[V^a_\tau - V^b_\tau|\mathcal{F}_t], \quad t \geq 0. \tag{4.2}$$

We take (4.2) as definition for our price process. We state now a result that enables to find a formula based on the Hawkes processes that we have introduced above.

Proposition 4.1. (Proposition 3.2 in [Jai15]) *Assume that the volume of each market order is constantly equal to v. The price* (4.2) *can be written as*

$$P_t = P_0 + \int_0^t \zeta(t-s)(dN^a_s - dN^b_s), \quad t \geq 0, \tag{4.3}$$

where

$$\zeta(t) = kv\left(1 + \int_t^\infty \left(\psi(u) - \int_0^t \psi(u-s)\phi(s)ds\right)du\right).$$

Proof. Since the volume of each market order is constant equal to v, the cumulated volumes are just $V^a_s = vN^a_s$ and $V^b_s = vN^b_s$. Hence, starting by (4.2), we have

$$\begin{aligned} P_t &= P_0 + k \lim_{s\to\infty} \mathbb{E}[V^a_s - V^b_s|\mathcal{F}_t] \\ &= P_0 + kv \lim_{s\to\infty} \mathbb{E}[N^a_s - N^b_s|\mathcal{F}_t]. \end{aligned}$$

Using the definition of associated martingale to the Hawkes process and Proposition 2.2, we get

$$\begin{aligned} P_t &= P_0 + kv \lim_{s\to\infty} \mathbb{E}\left[M^a_s - M^b_s + \int_0^s (\lambda^a_u - \lambda^b_u)du \Big| \mathcal{F}_t\right] \\ &= P_0 + kv \lim_{s\to\infty} \mathbb{E}\left[M^a_s - M^b_s + \int_0^t (\lambda^a_u - \lambda^b_u)du + \int_t^s \int_0^u \psi(u-x)(dM^a_x - dM^b_x)du \Big| \mathcal{F}_t\right] \\ &= P_0 + kv\left(N^a_t - N^b_t + \lim_{s\to\infty} \mathbb{E}\left[\int_t^s \int_0^u \psi(u-x)(dM^a_x - dM^b_x)du \Big| \mathcal{F}_t\right]\right). \end{aligned}$$

Now, thanks to the fact that $dM_x = dN_x - \lambda_x dx = dN_x - (\mu + \int_0^x \phi(x-r)dN_r)dx$, we get

$$\begin{aligned} P_t &= P_0 + kv\left(N^a_t - N^b_t + \int_t^\infty \int_0^t \psi(u-x)(dM^a_x - dM^b_x)du\right) \\ &= P_0 + kv\left(N^a_t - N^b_t + \int_t^\infty \int_0^t \psi(u-x)(dN^a_x - dN^b_x)du\right) \\ &\quad - kv\left(\int_t^\infty \int_0^t \psi(u-x) \int_0^x \phi(x-r)(dN^a_r - dN^b_r)dxdu\right). \end{aligned}$$

We can look separately at the following three terms. First,

$$N_t^a - N_t^b = \int_0^t (dN_r^a - dN_r^b).$$

For the second term, we can invert the two integrals to get

$$\begin{aligned}\int_t^\infty \int_0^t \psi(u-x)(dN_x^a - dN_x^b)du &= \int_0^t \int_t^\infty \psi(u-x)du(dN_x^a - dN_x^b)\\ &= \int_0^t \int_{t-x}^\infty \psi(v)dv(dN_x^a - dN_x^b).\end{aligned}$$

Finally, for the last term, using Fubini's theorem to change the integrals, we have

$$\begin{aligned}&\int_t^\infty \int_0^t \psi(u-x) \int_0^x \phi(x-r)(dN_r^a - dN_r^b)dxdu =\\ &\qquad = \int_t^\infty \left(\int_0^t \int_r^t \psi(u-x)\phi(x-r)dx(dN_r^a - dN_r^b) \right) du\\ &\qquad = \int_0^t \left(\int_t^\infty \int_r^t \psi(u-x)\phi(x-r)dxdu \right) (dN_r^a - dN_r^b)\\ &\qquad = \int_0^t \left(\int_t^\infty \int_0^{t-r} \psi(u-x'-r)\phi(x')dx'du \right) (dN_r^a - dN_r^b)\\ &\qquad = \int_0^t \left(\int_{t-r}^\infty \int_0^{t-r} \psi(u'-x')\phi(x')dx'du' \right) (dN_r^a - dN_r^b).\end{aligned}$$

Hence, putting all together, we find

$$\begin{aligned}P_t &= P_0 + kv \int_0^t \left(1 + \int_{t-r}^\infty \psi(u)du - \int_{t-r}^\infty \int_0^{t-r} \psi(u-x)\phi(x)dxdu \right) (dN_r^a - dN_r^b)\\ &= P_0 + \int_0^t \zeta(t-r)(dN_r^a - dN_r^b).\end{aligned}$$

□

Remark. From equation (4.3), we can see that if there is a buy trade at time t (jump of N^a) then the price jumps upwards of $\zeta(0)$, while if there is a sell trade at time t (jump of N^b) then the price jumps downwards of $\zeta(0)$.

The process (4.3) is the initial point to study market impact and in the next sections we will adapt the results of the previous chapters to our model.

4.2 The market impact function

In this section, we want to add the presence of a metaorder in the market and give a setting to apply the results of chapter 3, following [JR20]. First, we can establish some natural assumptions in financial modeling.

Starting by Definition 4.1, we call *permanent market impact* the following limit:

$$PMI = \lim_{t\to\infty} MI(t).$$

The idea is that we look at the long time behaviour of the function MI to see what is the exact increment of the price once it is stabilized after the execution of a metaorder.

Assumption 9. We assume the *no-statistical arbitrage principle*, that is the absence of round strategies with positive average profit and loss. This implies that:

- the PMI is linear with respect to the volume;
- the price P_t is a martingale.

We look at the time interval $[0,T]$, with the goal of sending $T \to \infty$ in the sequel. For this reason we add the index T in all the processes as in the previous chapters. We denote the Hawkes processes by $N^{a,T}$ and $N^{b,T}$, with intensities, for $t \in [0,T]$,

$$\lambda_t^{a,T} = \mu^T + \int_{(0,t)} \phi^T(t-s) dN_s^{a,T},$$
$$\lambda_t^{b,T} = \mu^T + \int_{(0,t)} \phi^T(t-s) dN_s^{b,T},$$

where $(\mu^T)_{T>0}$ is a sequence of positive real numbers, and $\phi^T : \mathbb{R}_+ \to \mathbb{R}_+$ is a non-negative measurable function satisfying the condition $\|\phi^T\|_{L^1} < \infty$. We assume the kernel function ϕ^T to be equal to

$$\phi^T(t) = a_T \phi(t), \quad \forall t \geq 0,$$

where $(a_T)_{T>0}$ is a sequence of real numbers with $a_T \in (0,1), \forall T > 0$, while ϕ is a non-negative measurable function such that $\|\phi\|_{L^1} = 1$.

Remark. We do not assume that $a_T \to 1$ as $T \to \infty$, but we will find this property as a result in the next theorems.

First, we can start by the price process (4.3) and write it in a more convenient way. Notice that in this setting, we have that $\|\phi^T\|_{L^1} < 1$ and $\psi^T(t) = \sum_{k=1}^{\infty} (\phi^T)^{*k}(t)$. Moreover, we set $kv = 1$, so the price becomes

$$P_t^T = P_0 + \int_0^t \zeta^T(t-s)(dN_s^{a,T} - dN_s^{b,T}), \quad t \in [0,T], \tag{4.4}$$

where

$$\zeta^T(t) = 1 + \int_t^{\infty} \psi^T(u) du - \int_t^{\infty} \int_0^t \psi^T(u-s) \phi^T(s) ds du.$$

Notice that

$$\zeta^T(0) = 1 + \int_0^{\infty} \psi^T(u) du = 1 + \|\psi^T\|_{L^1}.$$

Since the price is assumed to be a martingale (see [Jai15]), we can write

$$\zeta^T(t) = \zeta^T(0) + \int_0^t \frac{d}{ds} \zeta^T(s) ds. \tag{4.5}$$

We have

$$\begin{aligned}\frac{d}{dt}\zeta^T(t) &= -\psi^T(t) + \int_0^t \psi^T(t-s)\phi^T(s)ds - \int_t^\infty \psi^T(u-t)\phi^T(t)du \\ &= -\psi^T(t) + \psi^T * \phi^T(t) - \phi^T(t)\int_0^\infty \psi^T(v)dv \\ &= -\psi^T(t) + \psi^T(t) - \phi^T(t) - \phi^T(t)\|\psi^T\|_{L^1} \\ &= -(1+\|\psi^T\|_{L^1})\phi^T(t) = -\zeta^T(0)\phi^T(t).\end{aligned}$$

Using (4.5), we get

$$\begin{aligned}\zeta^T(t) &= \zeta^T(0) - \int_0^t \zeta^T(0)\phi^T(s)ds \\ &= \zeta^T(0)\left(1 - \int_0^t \phi^T(s)ds\right) \\ &= \left(1 + \int_0^\infty \psi^T(s)ds\right)\left(1 - \int_0^t \phi^T(s)ds\right).\end{aligned}$$

Another useful way to rewrite this function is

$$\begin{aligned}\zeta^T(t) &= \left(1 + \int_0^\infty \psi^T(s)ds\right)\left(1 - \int_0^t \phi^T(s)ds\right) \\ &= \frac{1}{1-a_T}\left(1 - \int_0^t \phi^T(s)ds\right) \\ &= \frac{1}{1-a_T}\left(1 - a_T\int_0^t \phi(s)ds\right) \\ &= \frac{1}{1-a_T}\left(1 - a_T + \int_t^\infty a_T\phi(s)ds\right) \\ &= 1 + \left(1 + \int_0^\infty \psi^T(s)ds\right)\int_t^\infty \phi^T(s)ds =: \xi^T(t).\end{aligned}$$

At this point, we can focus on modeling the presence of a buy metaorder in the market counting the number of its shares. Following [JR20], consider a non-homogeneous Poisson process $(n_t^T)_{t\in[0,T]}$ with the following intensity

$$\nu^T(t) = I^T f\left(\frac{t}{T}\right), \quad t \in [0,T],$$

where

$$I^T = \gamma\frac{\mu^T}{1-a_T}$$

with $\gamma \in (0,1)$, while f is a non-negative continuous function supported on $[0,1]$ with norm equal to one, so it is non-zero only for $t \in [0,T]$. This choice for the sequence I^T gives that the magnitude of the duration of the metaorder is of order T and depends on the long-term average intensity of the Hawkes processes, while the function f affects the splitting of the metaorder: for example, if f is an indicator function, then the metaorder is split into equivalent parts, in the sense that the interval of time between the executions of two following transactions is an exponential random variable that does not depend on time.

To look at the impact that a buy metaorder produces, the idea is to add the presence of the metaorder in the Hawkes based price model introduced before. In practice, we consider

$$P_t^T = P_0 + \int_0^t \xi^T(t-s)(dN_s^{a,T} - dN_s^{b,T} + dn_s^T), \quad t \in [0,T], \tag{4.6}$$

where

$$\xi^T(t) = 1 + \left(1 + \int_0^\infty \psi^T(s)ds\right) \int_t^\infty \phi^T(s)ds.$$

For any T fixed, we can look at the function

$$MI^T(t) := \mathbb{E}[P_t^T - P_0].$$

Thanks to expression (4.6), we get

$$MI^T(t) = \int_0^t \xi^T(t-s)\mathbb{E}[dn_s^T], \quad t \in [0,T],$$

where we have used the expression given in Proposition 1.2 to rewrite both the terms $\mathbb{E}[N_s^{a,T}]$ and $\mathbb{E}[N_s^{b,T}]$ and to simplify them. Thanks to the definition of ξ^T, we get

$$\begin{aligned} MI^T(t) &= \mathbb{E}[n_t^T] + \int_0^t \left(1 + \int_0^\infty \psi^T(u)du\right) \int_{t-s}^\infty \phi^T(u)du\, \mathbb{E}[dn_s^T] \\ &= \mathbb{E}[n_t^T] + \int_0^t \frac{1}{1-a_T} \int_{t-s}^\infty \phi^T(u)du\, \mathbb{E}[dn_s^T]. \end{aligned}$$

If we define the function

$$\Gamma^T(t) := \frac{1}{1-a_T} \int_t^\infty \phi^T(u)du,$$

then we can write

$$MI^T(t) = \mathbb{E}[n_t^T] + \int_0^t \Gamma^T(t-s)\, \mathbb{E}[dn_s^T].$$

This shows that we can decompose the market impact function into two parts

$$MI^T(t) = PMI^T(t) + TMI^T(t),$$

where $PMI^T(t)$ is called *permanent market impact part* and it is equal to

$$PMI^T(t) = \mathbb{E}[n_t^T], \quad t \in [0,T],$$

and $TMI^T(t)$ is called *transient market impact part* and it is equal to

$$TMI^T(t) = \int_0^t \Gamma^T(t-s)\, \mathbb{E}[dn_s^T], \quad t \in [0,T].$$

Remark. The function $TMI^T(t)$ is called transient market impact because it is vanishing as $t \to \infty$. Indeed, as $t \to \infty$, the function $\Gamma^T(t)$ goes to zero, and the intensity of the process n_t^T is eventually null, hence

$$\lim_{t\to\infty} TMI^T(t) = 0.$$

Hence in the limit as $t \to \infty$, only the permanent part remains.

4.3 Asymptotic dynamics of the market impact

At this point we need to introduce a rescaling, following always the one presented in [JR20], in order to have a proper regime to get the following theorems. Recall that we work with the market impact function

$$MI^T(f,t) = \mathbb{E}[n_t^T] + \int_0^t \Gamma^T(t-s)\,\mathbb{E}[dn_s^T] = PMI^T(f,t) + TMI^T(f,t), \quad t \in [0,T],$$

where we write explicitly the dependence on the function f, given by the intensity of the process n_t^T. We consider the following rescaling

$$\overline{MI}^T(f,t) := \frac{MI^T(f,Tt)}{T\beta^T}, \quad t \in [0,1],$$

where $\beta^T = \frac{\mu^T}{1-a_T}$. We can compute separately the two parts. Indeed, the permanent part gives

$$\begin{aligned}
\overline{PMI}^T(f,t) &= \frac{PMI^T(f,Tt)}{T\beta^T} = \frac{1}{T\beta^T}\mathbb{E}[n_{Tt}^T] \\
&= \frac{1}{T\beta^T}\int_0^{Tt} \nu^T(s)ds = \frac{1}{T\beta^T}\int_0^{Tt} \gamma\beta^T f\left(\frac{s}{T}\right) ds \\
&= \gamma \int_0^t f(x)dx.
\end{aligned}$$

For the transient part, we have

$$\begin{aligned}
\overline{TMI}^T(f,t) &= \frac{TMI^T(f,Tt)}{T\beta^T} = \frac{1}{T\beta^T}\int_0^{Tt} \Gamma^T(Tt-s)\,\mathbb{E}[dn_s^T] \\
&= \frac{1}{T\beta^T}\int_0^{Tt} \frac{1}{1-a_T}\left(\int_{Tt-s}^{\infty} \phi^T(u)du\right)\mathbb{E}[dn_s^T] \\
&= \frac{1}{T\beta^T}\int_0^{Tt} \frac{1}{1-a_T}\left(\int_{Tt-s}^{\infty} a_T\phi(u)du\right)\nu^T(s)ds \\
&= \frac{\gamma}{T}\frac{a_T}{1-a_T}\int_0^{Tt}\left(\int_{Tt-s}^{\infty}\phi(u)du\right) f\left(\frac{s}{T}\right) ds \\
&= \frac{\gamma}{T}\frac{a_T}{1-a_T}\int_0^{Tt}\left(\int_{x}^{\infty}\phi(u)du\right) f\left(t-\frac{x}{T}\right) dx.
\end{aligned}$$

Remark. The rescaled permanent part does not depend on T. Hence when we will look at the limit as $T \to \infty$ of the sequence $\overline{MI}^T$, it will be equivalent for the sequence $\overline{TMI}^T$.

Also in this case, our goal is to look at the limit as $T \to \infty$. We can work under the following natural assumption.

Assumption 10. If $f = \mathbb{1}_{[0,s]}$ for some $s \in (0,1]$, the scaling limit of the market impact function exists pointwise and it is non-increasing after time s. There exists $t > s$ such that the value of the limiting function at time t is smaller than that at time s.

Under Assumption 10, for $f = \mathbb{1}_{[0,s]}$, $s \in (0,1]$, we can define the pointwise limits, for any $t \geq 0$,

$$\widehat{MI}(f,t) = \lim_{T\to\infty} \overline{MI}^T(f,t),$$
$$\widehat{TMI}(f,t) = \lim_{T\to\infty} \overline{TMI}^T(f,t).$$

In the sequel, we will refer to this function also with the expression macroscopic (resp. permanent or transient) market impact function. Our goal is to study their asymptotic behaviour for any continuous function f.

Theorem 4.2. (Theorem 2.2 in [JR20]) *Under Assumptions 9 and 10, for any non-negative function f defined on $\mathbb{R}^+$, continuous on $[0,1]$, and supported on $[0,1]$, the macroscopic market impact function and its transient part exist. Indeed, there exist a parameter $\alpha \in (0,1]$ and a constant $K > 0$ such that, for any $t > 0$,*

$$\begin{cases} \lim_{T\to\infty} \overline{TMI}^T(f,t) = \gamma K(1-\alpha)\int_0^t f(t-u)u^{-\alpha}du & \text{if } \alpha < 1, \\ \lim_{T\to\infty} \overline{TMI}^T(f,t) = \gamma K f(t) & \text{if } \alpha = 1. \end{cases}$$

Moreover, the Hawkes kernel ϕ satisfies

$$\int_0^t \int_s^\infty \phi(u)duds = t^{1-\alpha}L(t), \tag{4.7}$$

where L is a slowly varying function (see the definition A.2). Finally, it holds

$$\frac{L(T)}{(1-a_T)T^\alpha} \xrightarrow[T\to\infty]{} K, \tag{4.8}$$

which implies that $a_T \to 1$ as $T \to \infty$.

Remark. We can notice some facts:

- in this case $a_T \to 1$ as $T \to \infty$ is obtained as a consequence of the model, and not assumed by hypothesis as in previous chapters. Also in this context, it is interpreted as the property that the market is highly endogenous.
- if $f = \mathbb{1}_{[0,1]}$, the permanent part is a linear function of the time, while the transient part is a power-law of the form

$$\begin{cases} t^{1-\alpha} & \text{if } 0 \leq t \leq 1, \\ t^{1-\alpha} - (t-1)^{1-\alpha} & \text{if } t > 1. \end{cases}$$

  See Figure B.3 for the plot of the functions in this case.

Proof. Step 1. First we derive the shape of the transient part. Consider $f = \mathbb{1}_{[0,s]}$ for some $s \in (0, 1]$. Thanks to Assumption 10, the pointwise limit of the market impact function exists, so in particular also the one of the transient market impact function, as the permanent part does not depend on T. Hence we can look at

$$\overline{TMI}^T(f,t) = \gamma \frac{a_T}{1-a_T} \int_0^t f(t-y) \int_{Ty}^{\infty} \phi(u) du dy$$

Using that ϕ is non-negative and integrable, we have that also $\overline{TMI}^T(f,t)$ is non-negative, and we can take the derivative to get

$$\begin{aligned} \frac{\partial}{\partial t}\overline{TMI}^T(f,t) &= \gamma \frac{a_T}{1-a_T} \left(f(0) \int_{Tt}^{\infty} \phi(u) du + \int_0^t f'(t-y) \int_{Ty}^{\infty} \phi(u) du dy \right) \\ &= \gamma \frac{a_T}{1-a_T} \left(\int_{Tt}^{\infty} \phi(u) du - \mathbb{1}_{[0,s]}(t) \int_{T(t-s)}^{\infty} \phi(u) du \right). \end{aligned}$$

So we deduce that $\overline{TMI}^T(f,t)$ is non-decreasing and concave in $[0, s]$, non-increasing after the time s. Hence it has its maximum in the point s. Also the function $\widehat{TMI}(f,t)$ has the same properties since it is the pointwise limit, and $\widehat{TMI}(f,s) > 0$.

Step 2. We can now show (4.7) and (4.8). Define the following functions, for $t \in (0, 1]$,

$$\begin{aligned} g(t) &= \frac{1}{\gamma}\widehat{TMI}(\mathbb{1}_{[0,t]}, t), \\ R(t) &= \int_0^t \int_y^{\infty} \phi(u) du dy. \end{aligned}$$

We have that

$$\frac{\overline{TMI}^T(\mathbb{1}_{[0,t]}, t)}{\overline{TMI}^T(\mathbb{1}_{[0,t]}, 1)} = \frac{\int_0^t \int_{Ty}^{\infty} \phi(u) du dy}{\int_0^1 \int_{Ty}^{\infty} \phi(u) du dy} = \frac{R(Tt)}{R(T)}$$

which converges to

$$\frac{\widehat{TMI}(\mathbb{1}_{[0,t]}, t)}{\widehat{TMI}(\mathbb{1}_{[0,t]}, 1)} = \frac{g(t)}{g(1)}$$

as $T \to \infty$. Hence we have, for all $t > 0$,

$$\frac{R(Tt)}{R(T)} \xrightarrow[T\to\infty]{} \frac{g(t)}{g(1)}.$$

Thanks to Theorem A.8, there exist a constant $\beta \in [0, 1]$ (to have the concavity of $g(t)$ if $t \in (0, 1]$) and L a slowly varying function such that

$$g(t) = g(1)t^{\beta},$$

and

$$R(t) = L(t)t^{\beta}.$$

Choosing $\beta = 1-\alpha$, the last relation gives (4.7). Moreover, taking $s = t = 1$ in the expression of $\overline{TMI}^T$, we have

$$\frac{1}{T}\frac{a_T}{1-a_T}\int_0^T\int_y^\infty \phi(u)dudy = \frac{1}{T}\frac{a_T}{1-a_T}R(T) = \frac{1}{T}\frac{a_T}{1-a_T}T^\beta L(T).$$

Taking the limit in the left hand side, we can find the expression of g, so we find that the limit is

$$g(1) =: K$$

as $T \to \infty$. Choosing again $\beta = 1-\alpha$, we find (4.8).

Step 3. We conclude with a sketch of proof of the limit of $\overline{TMI}^T$. First we can define the function

$$\bar{\Gamma}^T(y) = \frac{a_T}{1-a_T}\int_{Ty}^\infty \phi(u)du.$$

We know that

$$\lim_{T\to\infty}\int_0^t \bar{\Gamma}^T(y)dy = Kt^\beta.$$

The idea is now to use an approximation argument to move from the function g to any non-negative measurable function h. Indeed, it is easy to show that

$$\lim_{T\to\infty}\int_0^t g(u)\bar{\Gamma}^T(u)du = K\beta\int_0^t g(u)u^{\beta-1}du,$$

(see [JR20] for the exact proof). From this, we deduce also that for any non-negative function h defined on $\mathbb{R}^+$, continuous on $[0,1]$, and supported on $[0,1]$ it holds

$$\lim_{T\to\infty}\int_0^t h(t-u)\bar{\Gamma}^T(u)du = K\beta\int_0^t h(t-u)u^{\beta-1}du.$$

Adding the factor γ, from the left hand side we can find $\widehat{TMI}(h,t)$, and the conclusion follows choosing $\beta = 1-\alpha$. □

4.4 Limit theorems for the price

In this section we come back to focus on the dynamics of the price process introduced above, and we present a theorem that gives its macroscopic dynamics. Moreover, we provide some numerical simulations of the processes at different time scales in Figures B.1 and B.2.

First, we need to introduce a rescaling. Recall that

$$P_t^T = P_0 + \int_0^t \xi^T(t-s)(dN_s^{a,T} - dN_s^{b,T}), \quad t\in[0,T],$$

where

$$\xi^T(t) = \left(1+\int_0^\infty \psi^T(s)ds\right)\left(1-\int_0^t \phi^T(s)ds\right).$$

We want to work under the following regime.

Assumption 11. There exists $\delta > 0$ such that

$$\mu^T(1-a_T)T \xrightarrow[T\to\infty]{} \delta.$$

Under this assumption, we define a rescaling of the price, choosing $P_0 = 0$ and

$$\bar{P}_t^T := \frac{1}{T\beta^T} P_{Tt}^T = \frac{1-a_T}{T\mu^T} \int_0^t \xi^T(T(t-s))(dN_{Ts}^{a,T} - dN_{Ts}^{b,T}), \quad t \in [0,1].$$

Moreover, we consider the parameter α and the constant $K > 0$ given by theorem 4.2, and we define

$$\lambda = \frac{1}{K\Gamma(2-\alpha)}.$$

We can now state the main theorem which gives the macroscopic dynamics of the price.

Theorem 4.3. (Theorem 3.2 in [JR20]) *Under Assumptions 9, 10 and 11:*

1. *The sequence of rescaled price processes $(\bar{P}^T)_{T\geq 0}$ converges in law for the Skorokhod topology towards the process $\hat{P}$ defined by*

$$\hat{P}_t = \frac{1}{\sqrt{\delta}}\left(B^a_{X^a_t} - B^b_{X^b_t}\right), \quad t \in [0,1], \tag{4.9}$$

   where B^a and B^b are two independent Brownian motions such that $B^a_{X^a_t}$ and $B^b_{X^b_t}$ are two martingales. X^a is an increasing process solution of

$$X^a_t = \int_0^t F^{\alpha,\lambda}(s)ds + \frac{1}{\sqrt{\delta\lambda}} \int_0^t F^{\alpha,\lambda}(t-s)dB^a_{X^a_s}, \tag{4.10}$$

   and X^b is solution of the same equation with the index b.

2. *In particular, there exists a Brownian motion W such that the integrated variance of $\hat{P}$ defined as*

$$X = \frac{1}{\delta}(X^a + X^b)$$

   is solution of the stochastic rough Volterra equation:

$$X_t = \frac{2}{\delta}\int_0^t F^{\alpha,\lambda}(s)ds + \frac{1}{\delta\sqrt{\lambda}}\int_0^t F^{\alpha,\lambda}(t-s)dW_{X_s}. \tag{4.11}$$

3. *For any $\varepsilon > 0$, the process X has Hölder regularity $1 \wedge (2\alpha - \varepsilon)$, it is continuously differentiable for $\alpha > \frac{1}{2}$, and not continuously differentiable for $\alpha \leq \frac{1}{2}$.*

Theorem 4.3 establishes that the process X is continuously differentiable when $\alpha > \frac{1}{2}$. This process can be seen as the integrated variance of the price process, and when we take its derivative we can describe the spot volatility of the model, that we will denote by Y. To get the dynamics of Y, we work separately with the processes Y^a and Y^b, which are the derivatives of X^a and X^b respectively, and then we will pass to

$$Y_t = \frac{1}{\delta}(Y^a_t + Y^b_t).$$

Corollary 4.4. (Corollary 3.3 in [JR20]) *If $\alpha \in (\frac{1}{2}, 1]$, the process X is differentiable almost surely and its derivative Y is the unique solution of the stochastic rough Volterra equation*

$$Y_t = \frac{\lambda}{\Gamma(\alpha)} \left(\int_0^t (t-s)^{\alpha-1} \left(\frac{2}{\delta} - Y_s \right) ds + \frac{1}{\delta\sqrt{\lambda}} \int_0^t (t-s)^{\alpha-1} \sqrt{Y_s} dW_s \right), \tag{4.12}$$

where W is a Brownian motion. The dynamics of the price $\hat{P}$ is given by

$$d\hat{P}_t = \frac{1}{\sqrt{\delta}} \left(\sqrt{Y_t^a} dB_t^a - \sqrt{Y_t^b} dB_t^b \right). \tag{4.13}$$

Theorem 4.3 and Corollary 4.4 completely describe the macroscopic dynamics of the price model. From (4.9), we can see that the limit $\hat{P}$ of a Hawkes based process, built to reproduce the instantaneous transactions in the market, is a diffusive process, and in (4.13) we can see the dependence only on the volatility and Brownian motions. In (4.12) we get the dynamics of the spot volatility. These two equations together form the rough Heston model.

The classical Heston model is a stochastic volatility model in which the volatility is a Brownian semi-martingale, and it is largely used in portfolio theory for several reasons. Indeed, it reproduces important stylized facts of low frequency price data, it generates very reasonable shapes and dynamics for the implied volatility surface, and there is an explicit formula for the characteristic function of the asset log-price. Recent studies about the smoothness of the volatility (for example [CR98]) explain that it is better to describe the volatility with a fractional Brownian motion, leading to the introduction of the fractional stochastic volatility models. Moreover, in [GJR18], [EEFR18], [LMPR18] the authors show that the fractional Brownian motion has to be chosen with a Hurst exponent H of order 0.1, to have a model that is not only consistent with the empirically observed properties of the volatility time series but also with the shape of the volatility surface (see Figure B.4). Therefore, the Heston model has been adapted to the new setting of rough volatility models (see [EEGR19], [EER19]). The *rough Heston model* of a 1-dimensional asset price S is defined, for $t \geq 0$, by

$$\begin{cases} dS_t = S_t \sqrt{V_t} dW_t \\ V_t = V_0 + \dfrac{1}{\Gamma(\alpha)} \left(\lambda \displaystyle\int_0^t (t-u)^{\alpha-1} (\theta(u) - V_u) du + \nu \int_0^t (t-u)^{\alpha-1} \sqrt{V_u} dB_u \right) \end{cases}$$

where λ, ν, V_0 are positive coefficients, $\alpha \in (\frac{1}{2}, 1)$, and the correlation coefficient between the two Brownian motions $(B_t)_{t\geq 0}$ and $(W_t)_{t\geq 0}$ is $\rho \in (-1, 1)$. The roughness of the model is given by the kernel $(t-u)^{\alpha-1}$ which is the same that appears in expression (3.3) for the fractional Brownian motion. Indeed, the parameter α is related to the Hurst exponent of a fractional Brownian motion by the relation $\alpha = H + \frac{1}{2}$. In this model, the volatility sample paths have Hölder regularity $H - \varepsilon$, for any $\varepsilon > 0$.

Thanks to Theorem 4.3 and Corollary 4.4, we get that in the case $\alpha > \frac{1}{2}$ the dynamics of the price $\hat{P}$ together with the volatility Y is the one of a rough Heston model. On the other hand, if $\alpha < \frac{1}{2}$, the model has a higher roughness since the model is even less regular.

4.5 Proof of Theorem 4.3

The proof of Theorem 4.3 is divided in several steps, we present some of the main important aspects, whereas for some results we refer to [JR20].

First, we consider the processes

$$
\begin{aligned}
X_t^T &= \frac{1-a_T}{T\mu^T} N_{Tt}^T, \\
\Lambda_t^T &= \frac{1-a_T}{T\mu^T} \int_0^{Tt} \lambda_s^T ds, \\
Z_t^T &= \sqrt{\frac{T\mu^T}{1-a_T}} (X_t^T - \Lambda_t^T),
\end{aligned}
$$

for $t \in [0,1]$, where X_t^T is the renormalized Hawkes process, Λ_t^T its integrated intensity, and Z_t^T is the associated martingale. In fact, in this first part we develop general arguments without specifying the superscripts a and b.

We want to use the results of chapter 3. Indeed, we have the same setting and hypotheses on the sequence of Hawkes processes, the only difference is the assumption regarding the Hawkes kernel ϕ, since:

- Assumption in chapter 3: $\displaystyle\int_t^{\infty} \phi(s)ds = Kt^{-\alpha}$;
- Assumption in this chapter: $\displaystyle\int_0^t \int_s^{\infty} \phi(u)duds = L(t)t^{1-\alpha}$.

Hence in this case we work in a more general setting, but this does not change very much the results. In particular, the main difference is in the proof of the following lemma, which is the analogous result of the convergence expressed in (3.6).

Lemma 4.5. *The sequence of functions $\rho^T(t) = \frac{T(1-a_T)}{a_T}\psi^T(Tt)$ converges weakly towards $f^{\alpha,\lambda}$. Furthermore, $\int_0^t \rho^T(s)ds$ converges uniformly towards $F^{\alpha,\lambda} = \int_0^t f^{\alpha,\lambda}(s)ds$.*

Proof. We proceed in a similar way as in the analogous proof of section 3.2. First, note that since ρ^T can be interpreted as a density function, we can look at its Laplace transform $\hat{\rho}^T$ and find the convergence. Our goal is to prove that $\hat{\rho}^T$ converges to the Laplace transform of $f^{\alpha,\lambda}$. Using (3.1), we have that for $z > 0$

$$
\hat{\rho}^T(z) = \frac{\hat{\phi}(\frac{z}{T})}{1 - \frac{a_T}{1-a_T}(\hat{\phi}(\frac{z}{T}) - 1)},
$$

where $\hat{\phi}$ denotes the Laplace transform of the function ϕ. We want to find the asymptotic expansion of this function. Define

$$
R(t) = \int_0^t \int_s^{\infty} \phi(u)duds,
$$

This function is equal to

$$
R(t) = t^{1-\alpha} L(t)
$$

thanks to Theorem 4.2. We can apply Theorem A.9 to get that

$$\hat{R}(z) = \int_0^\infty e^{-zs} R(s) ds \sim_{z\to 0^+} z^{\alpha-2} L\left(\frac{1}{z}\right)\Gamma(2-\alpha).$$

Integrating by parts the right hand side, we find

$$\begin{aligned}\hat{R}(z) &= \int_0^\infty e^{-zs} R(s) ds = \int_0^\infty \frac{e^{-zs}}{z}\int_s^\infty \phi(u) du ds \\ &= \frac{1}{z^2}\int_0^\infty \phi(u) du - \frac{1}{z^2}\int_0^\infty e^{-zs}\phi(s) ds \\ &= \frac{1}{z^2}(1-\hat{\phi}(z)).\end{aligned}$$

So we find

$$\hat{R}(z) = \frac{1}{z^2}(1-\hat{\phi}(z)) \sim_{z\to 0^+} z^{\alpha-2} L\left(\frac{1}{z}\right)\Gamma(2-\alpha),$$

and choosing $\frac{z}{T}$ as argument of the functions, it implies

$$\frac{a_T}{1-a_T}\left(1-\hat{\phi}\left(\frac{z}{T}\right)\right) \sim_{T\to\infty} \frac{a_T}{1-a_T}\left(\frac{z}{T}\right)^\alpha L\left(\frac{T}{z}\right)\Gamma(2-\alpha).$$

Using (4.8) and the definition of slowly varying function, we finally deduce that

$$\lim_{T\to\infty}\frac{a_T}{1-a_T}\left(1-\hat{\phi}\left(\frac{z}{T}\right)\right) = z^\alpha K\Gamma(2-\alpha).$$

Coming back into the expression of $\hat{\rho}^T$, we get

$$\lim_{T\to\infty}\hat{\rho}^T(z) = \frac{1}{1+K\Gamma(2-\alpha)z^\alpha} = \frac{\lambda}{\lambda+z^\alpha}$$

choosing $\lambda = \frac{1}{K\Gamma(2-\alpha)}$. As the right hand side is the Laplace transform of $f^{\alpha,\lambda}$, we conclude that, for $z > 0$,

$$\rho^T(z) \xrightarrow[T\to\infty]{weakly} f^{\alpha,\lambda}(z).$$

□

Given this lemma, all the results developed in chapter 3 are still valid, under Assumption 11. In particular, we can apply theorems 3.1 and 3.2 to find the dynamics of any limit point X of the sequence $(X^T)_{T\geq 0}$:

$$X_t = \int_0^t s f^{\alpha,\lambda}(t-s) ds + \frac{1}{\sqrt{\delta\lambda}}\int_0^t f^{\alpha,\lambda}(t-s) B_{X_s} ds.$$

Thanks to integration by parts, this also writes as

$$X_t = \int_0^t F^{\alpha,\lambda}(t-s) ds + \frac{1}{\sqrt{\delta\lambda}}\int_0^t F^{\alpha,\lambda}(t-s) dB_{X_s}.$$

Furthermore, using that $Z_t = B_{X_t}$, we can derive the following useful equation, that we will use later:

$$X_t = \int_0^t f^{\alpha,\lambda}(t-s)\left(s + \frac{1}{\sqrt{\delta\lambda}} Z_s\right) ds.$$

Notice that this will be the equation satisfied by both X^a and X^b when we will come back to distinguish the two processes in the sequel.

At this point, we need to prove that the sequence $(X^T, Z^T)_{T\geq 0}$ converges in law for the Skorokhod topology, which is true if

1. $(X^T, Z^T)_{T\geq 0}$ is tight;
2. all the limit points have the same law.

We need to prove only point (2), since the tightness is the same of chapter 3. Hence our goal is to prove the uniqueness of the law of any limit point (X, Z).

An important observation is that the dynamics

$$X_t = \int_0^t f^{\alpha,\lambda}(t-s)\left(s + \frac{1}{\sqrt{\delta\lambda}} Z_s\right) ds \tag{4.14}$$

is equivalent to

$$D^\alpha X_t + \lambda X_t - \lambda t = \sqrt{\frac{\lambda}{\delta}} Z_t, \tag{4.15}$$

where D^α is the fractional derivative. We show how to derive this last equation in Appendix A.6, following [SKM93]. Equation (4.15) allows to say that the law of (X, Z) is uniquely determined by the law of X, so it is sufficient to prove the uniqueness in law of the limit points of $(X_T)_{T\geq 0}$ to deduce the convergence in law of $(X^T, Z^T)_{T\geq 0}$.

We use the following lemma, that we are not going to prove (see [JR20] for details).

Lemma 4.6. *Let X be a limit point of $(X^T)_{T\geq 0}$ and $h : \mathbb{R}^+ \to \mathbb{R}$ a continuously differentiable function such that $h(0) = 0$. Define the following characteristic function:*

$$\nu(h,t) = \mathbb{E}[\exp((ih * dX)(t))], \quad t \geq 0.$$

Then it satisfies

$$\nu(h,t) = \exp\left(\int_0^t g(s) ds\right),$$

where g is the unique continuous solution of the rough Volterra Riccati equation

$$g = f^{\alpha,\lambda} * \left(\frac{1}{2\delta} g^2 + ih\right). \tag{4.16}$$

Thanks to this lemma, we focus on the characteristic functions of the limit point X and we get that it is a functional of the unique solution of (4.16). Hence we can obtain the uniqueness in law for the limit points of $(X_T)_{T\geq 0}$ from the uniqueness of continuous solution of equation (4.16). This ends the proof of the convergence in law of the sequence $(X^T, Z^T)_{T\geq 0}$, and we will use it to get the convergence of the price process.

Now we are ready to derive (4.9), (4.10) and (4.11). We start to use the superscripts a and b to distinguish the processes related to buy orders and the ones for sell orders. First, we can achieve a useful way to rewrite the price process.

Proposition 4.7. *The process $\bar{P}^T$ can be written as*

$$\bar{P}_t^T = \frac{1-a_T}{T\mu^T}\left(1+\int_0^\infty \psi^T(s)ds\right)(M_{Tt}^{a,T} - M_{Tt}^{b,T}), \quad t \in [0,1],$$

where $M^{a,T}$ and $M^{b,T}$ are the martingales

$$M_t^{a,T} = N_t^{a,T} - \int_0^t \lambda_s^{a,T} ds, \quad M_t^{b,T} = N_t^{b,T} - \int_0^t \lambda_s^{b,T} ds.$$

Proof. Starting by the definition of the process P^T, we know that

$$P_t^T = \int_0^t \left(1+\int_0^\infty \psi^T(s)ds\right)\left(1-\int_0^{t-u}\phi^T(s)ds\right)(dN_u^{a,T} - dN_u^{b,T}).$$

We work on the term with the two integrals, and we find

$$\begin{aligned}
&\int_0^t \left(1+\int_0^\infty \psi^T(s)ds\right)\int_0^{t-u}\phi^T(s)ds(dN_u^{a,T} - dN_u^{b,T}) = \\
&\qquad = \left(1+\int_0^\infty \psi^T(s)ds\right)\int_0^t\int_u^t \phi^T(s-u)ds(dN_u^{a,T} - dN_u^{b,T}) \\
&\qquad = \left(1+\int_0^\infty \psi^T(s)ds\right)\int_0^t\int_0^s \phi^T(s-u)(dN_u^{a,T} - dN_u^{b,T})ds \\
&\qquad = \left(1+\int_0^\infty \psi^T(s)ds\right)\int_0^t (\lambda_s^{a,T} - \mu^T - \lambda_s^{b,T} + \mu^T)ds.
\end{aligned}$$

Replacing this expression in P^T we get

$$\begin{aligned}
P_t^T &= \left(1+\int_0^\infty \psi^T(s)ds\right)\int_0^t (dN_u^{a,T} - \lambda_u^{a,T}du - dN_u^{b,T} + \lambda_u^{b,T}du) \\
&= \left(1+\int_0^\infty \psi^T(s)ds\right)(M_t^{a,T} - M_t^{b,T}).
\end{aligned}$$

Finally,

$$\bar{P}_t^T = \frac{1-a_T}{T\mu^T}P_{Tt}^T = \frac{1-a_T}{T\mu^T}\left(1+\int_0^\infty \psi^T(s)ds\right)(M_{Tt}^{a,T} - M_{Tt}^{b,T}).$$

□

Thanks to the definitions of the processes $(X^{a,T}, Z^{a,T})_{T\geq 0}$ related to $(N^{a,T})_{T\geq 0}$, and $(X^{b,T}, Z^{b,T})_{T\geq 0}$ related to $(N^{b,T})_{T\geq 0}$, we also have

$$\begin{aligned}
\bar{P}_t^T &= \frac{1-a_T}{T\mu^T}\left(1+\int_0^\infty \psi^T(s)ds\right)(M_{Tt}^{a,T} - M_{Tt}^{b,T}) \\
&= \frac{1}{1-a_T}\frac{1-a_T}{T\mu^T}\left(N_{Tt}^{a,T} - \int_0^{Tt}\lambda_s^{a,T}ds - N_{Tt}^{b,T} + \int_0^{Tt}\lambda_s^{b,T}ds\right) \\
&= \frac{1}{1-a_T}\sqrt{\frac{1-a_T}{T\mu^T}}(Z_t^{a,T} - Z_t^{b,T}) \\
&= \frac{1}{\sqrt{T\mu^T(1-a_T)}}(Z_t^{a,T} - Z_t^{b,T}).
\end{aligned}$$

Now we have that $T\mu^T(1-a_T) \to \delta$ as $T \to \infty$. Moreover, thanks to previous arguments, we have that $(Z^{a,T})_{T\geq 0}$ and $(Z^{b,T})_{T\geq 0}$ converge in law for the Skorokhod topology, so also $(\bar{P}^T)_{T\geq 0}$ converges in law towards a process $\hat{P}$ for the Skorokhod topology. By Theorem 3.2, there exist two independent Brownian motions B^a and B^b such that

$$Z_t^a = B_{X_t^a}^a, \quad Z_t^b = B_{X_t^b}^b.$$

Hence the limit of $\bar{P}^T$ is the process

$$\hat{P}_t = \frac{1}{\sqrt{\delta}}(B_{X_t^a}^a - B_{X_t^b}^b),$$

where X^a is the limit of the sequence $(X^{a,T})_{T\geq 0}$ and it solves

$$X_t^a = \int_0^t F^{\alpha,\lambda}(t-s)ds + \frac{1}{\sqrt{\delta\lambda}}\int_0^t F^{\alpha,\lambda}(t-s)dB_{X_s^a}^a.$$

The same holds for X^b. Finally, let us look at

$$X = \frac{1}{\delta}(X^a + X^b).$$

This process satisfies

$$X_t = \frac{2}{\delta}\int_0^t F^{\alpha,\lambda}(t-s)ds + \frac{1}{\delta\sqrt{\lambda}}\int_0^t F^{\alpha,\lambda}(t-s)\frac{1}{\sqrt{\delta}}d(B_{X_s^a}^a + B_{X_s^b}^b).$$

If we define the Brownian motion

$$W_{X_t} = \frac{1}{\sqrt{\delta}}(B_{X_s^a}^a + B_{X_s^b}^b),$$

then we find

$$X_t = \frac{2}{\delta}\int_0^t F^{\alpha,\lambda}(s)ds + \frac{1}{\delta\sqrt{\lambda}}\int_0^t F^{\alpha,\lambda}(t-s)dW_{X_s}.$$

4.6 Proof of Corollary 4.4

This corollary is an application of the results of chapter 3 and a consequence of a theorem of [EEFR18].

We start to work with X^a. By Theorem 4.3, it satisfies

$$X_t^a = \int_0^t F^{\alpha,\lambda}(s)ds + \frac{1}{\sqrt{\delta\lambda}}\int_0^t F^{\alpha,\lambda}(t-s)dB_{X_s^a}^a$$

and doing integration by parts, we have seen in the previous proof that this is equal to

$$X_t^a = \int_0^t sf^{\alpha,\lambda}(t-s)ds + \frac{1}{\sqrt{\delta\lambda}}\int_0^t f^{\alpha,\lambda}(t-s)B_{X_s^a}^a ds.$$

Hence, the process X^a has the dynamics described in Theorem 3.2. It follows that we can apply Theorem 3.4: if $\alpha > \frac{1}{2}$, then the dynamics of the derivative Y^a is

$$Y_t^a = F^{\alpha,\lambda}(t) + \frac{1}{\sqrt{\delta\lambda}}\int_0^t f^{\alpha,\lambda}(t-s)\sqrt{Y_s^a}dB_s^a. \tag{4.17}$$

Notice that from this equation it is clear the roughness of the process for the presence of the kernel $f^{\alpha,\lambda}$, but it is more clear once we derive the following result, taken by [EEFR18].

Proposition 4.8. (Proposition 4.9 in [EEFR18]) *Let λ, ν, θ be positive constants and $\alpha \in (\frac{1}{2}, 1)$, let B be a Brownian motion. The process V is solution of the following rough SDE*

$$V_t = \theta F^{\alpha,\lambda}(t) + \nu \int_0^t f^{\alpha,\lambda}(t-s)\sqrt{V_s}dB_s$$

if and only if it is solution of

$$V_t = \frac{\lambda}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1}(\theta - V_s)ds + \frac{\lambda\nu}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1}\sqrt{V_s}dB_s.$$

Furthermore, both equations admit a unique strong solution.

Hence, starting by (4.17), we get

$$Y_t^a = \frac{\lambda}{\Gamma(\alpha)} \left(\int_0^t (t-s)^{\alpha-1}(1 - Y_s^a)ds + \frac{1}{\sqrt{\delta\lambda}} \int_0^t (t-s)^{\alpha-1}\sqrt{Y_s^a}dB_s^a \right)$$

(the same dynamics holds for Y^b, but with the index b). From this, we get also (4.12).

Finally, we can find (4.13). Starting by

$$\hat{P}_t = \frac{1}{\sqrt{\delta}} \left(B^a_{X_t^a} - B^b_{X_t^b} \right), \quad t \in [0,1],$$

we know that $B^a_{X_t^a} = Z_t^a$ with $[Z^a, Z^a]_t = X_t^a$, and $B^b_{X_t^b} = Z_t^b$ with $[Z^b, Z^b]_t = X_t^b$. This means that

$$[Z^a, Z^a]_t = X_t^a = \int_0^t Y_s^a ds,$$
$$[Z^b, Z^b]_t = X_t^b = \int_0^t Y_s^b ds.$$

Hence, by Theorem A.6, we find

$$Z_t^a = \int_0^t \sqrt{Y_s^a}dB_s^a,$$
$$Z_t^b = \int_0^t \sqrt{Y_s^b}dB_s^b,$$

and finally,

$$d\hat{P}_t = \frac{1}{\sqrt{\delta}} \left(\sqrt{Y_t^a}dB_t^a - \sqrt{Y_t^b}dB_t^b \right).$$

Appendix A

Complements

A.1 Technical lemma

Lemma A.1. *Let $h : \mathbb{R}_+ \to \mathbb{R}^d$ be a Borel and locally bounded function and consider the equation:*

$$f(t) = h(t) + \int_0^t \phi(t-s)f(s)ds, \quad \forall t \geq 0. \tag{A.1}$$

Then there exists a unique locally bounded solution $f_h : \mathbb{R}_+ \to \mathbb{R}^d$ given by

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

Proof. First we want to prove that the candidate solution f_h satisfies equation (A.1). Notice that f_h is locally bounded since $\psi \in L^1$ and h is locally bounded. Hence we can compute:

$$\begin{aligned}
\int_0^t \phi(t-s)f_h(s)ds &= \int_0^t \phi(t-s)h(s)ds + \int_0^t \phi(t-s)\left(\int_0^s \psi(s-u)h(u)du\right)ds \\
&= \int_0^t \phi(t-s)h(s)ds + \int_0^t \int_0^{t-u} \phi(t-u-s)\psi(s)ds\, h(u)du \\
&= \int_0^t \phi(t-s)h(s)ds + \int_0^t (\psi(t-u) - \phi(t-u))\, h(u)du \\
&= \int_0^t \psi(t-u)h(u)du,
\end{aligned}$$

where we have used the useful relation

$$(\psi * \phi)(t) = \psi(t) - \phi(t).$$

Hence it follows that f_h satisfies equation (A.1).

For the uniqueness, if f_h and g_h are two functions that satisfy (A.1), define the function

$$w^i(t) = |f_h^i(t) - g_h^i(t)| \quad \forall i \in \{1, \ldots, d\}.$$

Applying Fubini theorem, we get, for all $i \in \{1, \dots, d\}$,

$$\int_0^\infty w^i(t)dt \le \int_0^\infty \int_0^t \sum_{j=1}^d \phi^{ij}(t-s)w^j(s)ds\,dt = \sum_{j=1}^d \|\phi^{ij}\|_{L^1} \int_0^\infty w^j(t)dt$$

and, if $w = (w^1, \dots, w^d)$, this implies that

$$(Id - K) \int_0^\infty w(t)dt \le 0.$$

Thanks to the assumption on the spectral radius of K, we get that $f_h = g_h$ almost everywhere, and so

$$\int_0^t \phi(t-s)f_h(s)ds = \int_0^t \phi(t-s)g_h(s)ds.$$

But both f_h and g_h satisfy (A.1), hence $f_h = g_h$. □

A.2 Population interpretation of Hawkes processes

In this section we want to discuss the representation of Hawkes processes in term of clusters, following the interpretation presented in [EER19]. Consider a d-dimensional Hawkes process $N = (N^1, \dots, N^d)$ with intensity given by

$$\lambda_t^i = \mu^i + \int_{(0,t)} \sum_{j=1}^d \phi^{ij}(t-s)dN_s^j, \quad \forall t \ge 0, \forall i \in \{1, \dots, d\}, \tag{A.2}$$

which satisfies the stability assumption. We can interpret the law of this process through a population dynamics. Consider that there exist d types of individuals and, for each type, an individual can be a migrant or a descendant of a migrant. Starting form time $t = 0$, we have the following dynamics:

1. migrants of type $k \in \{1, \dots, d\}$ arrive according to a non-homogeneous Poisson process with rate μ^k;
2. each migrant of type $k \in \{1, \dots, d\}$ generates children of type $j \in \{1, \dots, d\}$ according to a non-homogeneous Poisson process with rate $\phi^{jk}(t)$;
3. each child of type $k \in \{1, \dots, d\}$ generates children of type $j \in \{1, \dots, d\}$ according to a non-homogeneous Poisson process with rate $\phi^{jk}(t)$.

This population model describes exactly the same point process of (A.2), where the process N_t counts the number of individuals of the population up to time t.

We can achieve another useful formula, which explain the cluster representation of the population. Indeed, let us call N^0 the non-homogeneous Poisson process with intensity μ which describes the arrivals of the migrants into the population set, and $\tilde{N}^i$ the Hawkes process which counts the number of descendants of the migrant i. We call cluster the set of the offspring of a migrant. Then, it holds

$$N_t = N_t^0 + \sum_{1 \le i \le N_t^0} \tilde{N}_{t-T_i}^i \tag{A.3}$$

where the sequence $(T_i)_{i\geq 1}$ are the arrival times of the migrants.

Consider now the Hawkes based model for an asset price defined in section 4.1 to reproduce the transactions of the market counting the buy and sell orders. Then according to the population interpretation, the parameter μ describes the intensity of the exogenous orders of the market, whereas the function ϕ is related to the endogenous orders produced by other orders.

A.3 Tightness and convergence in law

We recall some definitions about tightness and we state some useful theorems that are used in several proofs above.

Consider a sequence of random variables $(X^n)_{n\geq 1}$ with values in a space E, where, for any $n \geq 1$, X^n is defined on the space $(\Omega^n, \mathcal{F}^n, P^n)$. We say that the sequence $(X^n)_{n\geq 1}$ converges in law to a random variable X if $Law(X^n) \to Law(X)$ weakly in $\mathcal{P}(E)$, that is

$$\mathbb{E}_{P^n}[f(X^n)] \xrightarrow[n\to\infty]{} \mathbb{E}_P[f(X)]$$

for all bounded and continuous functions f on E. At the same way, we say that the sequence $(X^n)_{n\geq 1}$ is tight if the sequence of distributions $(Law(X^n))_{n\geq 1}$ is tight, i.e. for all $\varepsilon > 0$ there exists a compact set K in E such that $P^n(X^n \notin K) \leq \varepsilon$.

Consider now a sequence $(X^n)_n$ of $\mathbb{R}^d$-valued càdlàg processes, where we always equip the space of càdlàg functions with the Skorokhod topology.

Definition A.1. A sequence $(X^n)_n$ of processes is said C-tight if it is tight and all the limit points of the sequence $\{Law(X^n)\}_n$ are laws of continuous processes.

Theorem A.2. (Proposition VI-3.26 in [JS03]) *The sequence $(X^n)_n$ is C-tight if and only if the sequence $(X^n)_n$ is tight and, for all $N \in \mathbb{N}, \varepsilon > 0$, it holds:*

$$\lim_{n\to\infty} P^n\left(\sup_{t\leq N} |\Delta X_t^n| > \varepsilon\right) = 0.$$

Remark. Theorem A.2 is an important characterization that guarantees the C-tightness of a sequence of càdlàg processes proving its tightness and the fact that the amplitude of jumps converges to zero in probability.

We suppose now to work with a sequences of locally square integrable martingales $Z^n - Z_0^n$.

Theorem A.3. (Theorem VI-4.13 in [JS03]) *Suppose that $(Z^n - Z_0^n)_{n\geq 1}$ is a sequence of locally square-integrable martingales, and denote by $(X^n)_{n\geq 1}$ the sequence of quadratic variations, i.e. $X^n = [Z^n, Z^n]$ for any $n \geq 1$. If*

1. *the sequence $(Z_0^n)_{n\geq 1}$ is tight,*

2. *the sequence $(X^n)_{n\geq 1}$ is C-tight,*

then the sequence $(Z^n)_{n\geq 1}$ is tight.

We state a theorem that relates the limit of a sequence of martingales to the limit of the sequence of their brackets. This result is a version of theorem VI-6.26 in [JS03], since we state it looking directly at tightness of the brackets.

Theorem A.4. (Theorem VI-6.26 in [JS03]) *Given a sequence of local martingales* $(Z^n)_n$ *with* $|\Delta Z^n| \leq k$*, for* $k > 0$*, assume that*

1. $Z^n \xrightarrow[n\to\infty]{} Z^\infty$ *in law;*
2. *for any* $t > 0$*, the sequence* $([Z^n, Z^n]_t)_n$ *is tight.*

Then we have

$$(Z^n, [Z^n, Z^n]) \xrightarrow[n\to\infty]{} (Z^\infty, [Z^\infty, Z^\infty]) \quad \text{in law},$$
$$[Z^n, Z^n] \xrightarrow[n\to\infty]{} [Z^\infty, Z^\infty] \quad \text{in law}.$$

We present now a theorem for the convergence in law of sequences of local martingales.

Theorem A.5. (Theorem VIII-3.11 in [JS03]) *Assume that* Z *is a continuous Gaussian martingale with characteristics* $(0, C, 0)$*, and that* $(Z^n)_{n\geq 1}$ *is a sequence of local martingales with bounded jumps, i.e.* $|\Delta Z^n| \leq k$*,* $\forall n \geq 1$*. Then the following properties are equivalent:*

1. $Z^n \xrightarrow[n\to\infty]{} Z$ *in law,*
2. $[Z^n, Z^n] \xrightarrow[n\to\infty]{} C$ *in law,*

where $[Z^n, Z^n]$ *denotes the quadratic variation of* Z^n*.*

We conclude this appendix reporting a theorem of representation of martingales, taken by [RY99].

Theorem A.6. (Proposition V-3.8 in [RY99]) *If* M *is a continuous local martingale such that the measure* $d\langle M, M\rangle_t$ *is a.s. equivalent to the Lebesgue measure, there exist a process* f_t *predictable and strictly positive* $dt \otimes dP$*-a.s. and a Brownian motion* B *such that, for all* $t \geq 0$*,*

$$d\langle M, M\rangle_t = f_t \, dt$$

and

$$M_t = M_0 + \int_0^t f_s^{\frac{1}{2}} \, dB_s.$$

A.4 Tauberian theorem

We present some useful results concerning slowly varying functions and Tauberian theorem, taken by [BGT89].

Definition A.2. A measurable function $L : \mathbb{R}_+ \to \mathbb{R}$ is called *slowly varying* if, for all $s > 0$,

$$\frac{L(st)}{L(t)} \xrightarrow[t\to\infty]{} 1.$$

Proposition A.7. *Let L be a slowly varying function and $\alpha > 0$, then*

$$t^{-\alpha} L(t) \xrightarrow[t\to\infty]{} 0.$$

Theorem A.8. *Let U be a positive measurable function on $\mathbb{R}^+$ such that for all $s > 0$ it satisfies*

$$\frac{U(ts)}{U(t)} \xrightarrow[t\to\infty]{} g(s) > 0,$$

for some function g. Then there exist a constant $\alpha \in \mathbb{R}$ and a slowly varying function L such that

$$g(t) = t^\alpha$$
$$U(t) = t^\alpha L(t).$$

Theorem A.9. (Tauberian theorem.) *Let U be a measurable non-negative function, $c \geq 0$, and $\rho > -1$. Assume that $\hat{U}(z) = \int_0^\infty e^{-zs} U(s) ds$ is finite for any $z > 0$, and*

$$U(t) \sim_{t\to\infty} ct^\rho \frac{L(t)}{\Gamma(1+\rho)}$$

for L a slowly varying function. Then

$$\hat{U}(z) \sim_{z\to 0^+} cz^{-\rho-1} L\left(\frac{1}{z}\right).$$

A.5 Mittag-Leffler functions

Definition A.3. Let α, β non-negative constants. The function

$$E_{\alpha,\beta}(z) = \sum_{n\geq 0} \frac{z^n}{\Gamma(\alpha n + \beta)}, \quad z \in \mathbb{C},$$

is called Mittag-Leffler function. Moreover, if $\alpha \in (0, 1]$ and $\lambda \in \mathbb{R}^+$, we define the Mittag-Leffler density function as

$$f^{\alpha,\lambda}(t) = \lambda t^{\alpha-1} E_{\alpha,\alpha}(-\lambda t^\alpha), \quad t > 0,$$

and its cumulative distribution function

$$F^{\alpha,\lambda}(t) = \int_0^t f^{\alpha,\lambda}(s) ds, \quad t \geq 0.$$

The Laplace transform of $f^{\alpha,\lambda}$ is

$$\hat{f}^{\alpha,\lambda}(z) = \frac{\lambda}{\lambda + z^\alpha}.$$

A.6 Fractional integrals and derivatives

We define fractional integrals and derivatives, following [SKM93].

Definition A.4. Consider a parameter $\alpha \in (0, 1)$ and a λ-Hölder function f (with $\lambda > \alpha$). Then we define the fractional integral as

$$I^\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_0^x \frac{f(t)}{(x-t)^{1-\alpha}} dt,$$

and the fractional derivative as

$$D^\alpha f(x) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dx} \int_0^x \frac{f(t)}{(x-t)^{\alpha}} dt.$$

We also state two results about fractional integrals and derivatives. Let us denote by $\mathcal{H}^\lambda$ the space of Hölder continuous functions with exponent $\lambda > 0$.

Proposition A.10. (Proposition A.1 in [JR16]) *If $f \in \mathcal{H}^\lambda$ and $f(0) = 0$, then for any $\alpha < \lambda$, f admits a fractional derivative of order α and $D^\alpha f \in \mathcal{H}^{\lambda-\alpha}$.*

Proposition A.11. (Corollary A.2 in [JR16]) *Let ϕ be continuous and ψ such that $x^\mu \psi(x) \in \mathcal{H}^\lambda$ for some $\mu > 0$. Then, for any $\alpha < \min(1-\mu, \lambda)$, $D^\alpha \psi$ exists, belongs to L^r for some $r > 1$ and*

$$\int_0^t \phi(t-s)\psi(s)ds = \int_0^t I^\alpha \phi(t-s) D^\alpha \psi(s) ds.$$

We present now a result taken by Example 42.2 in [SKM93], which explains how we can derive equation (4.15). It is an application of theorem 42.1 in [SKM93], see it for detailed arguments.

Consider the following Cauchy problem:

$$\begin{cases} D^\alpha y(x) - \lambda y(x) = h(x) & \text{for } n-1 < \alpha \le n, \\ D^{\alpha-k} y(x)|_{x=0} = b_k & \text{for } k = 1, \dots, n, \end{cases}$$

where h is a given function, λ, b_k are fixed.

To find the solution, following the idea developed in the proof of theorem 42.1 in [SKM93], we can define the following sequence recursively:

$$\begin{cases} y_0(x) = \displaystyle\sum_{k=1}^n b_k \frac{x^{\alpha-k}}{\Gamma(\alpha-k+1)}, \\ y_m(x) = y_0(x) + \displaystyle\frac{\lambda}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} y_{m-1}(t) dt + \frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} h(t) dt, \quad m \in \mathbb{N}, \end{cases}$$

which can be written also as

$$y_m(x) = \sum_{k=1}^n b_k \sum_{j=1}^{m+1} \frac{\lambda^{j-1} x^{j\alpha-k}}{\Gamma(j\alpha-k+1)} + \sum_{j=1}^m \frac{\lambda^{j-1}}{\Gamma(j\alpha)} \int_0^x h(u)(x-u)^{j\alpha-1} du,$$

for $m \in \mathbb{N}$. Using the result about the convergence of this sequence proved in the aforementioned theorem, it is sufficient to pass to the limit as $m \to \infty$ to find the solution of the Cauchy problem. This gives

$$y(x) = \sum_{k=1}^{n} b_k x^{\alpha-k} E_{\alpha,1+\alpha-k}(\lambda x^\alpha) + \int_0^x (x-u)^{\alpha-1} h(u) E_{\alpha,\alpha}(\lambda(x-u)^\alpha) du.$$

In particular, if $\alpha \in (0,1]$ and $k = 1$, we get

$$y(x) = b_1 x^{\alpha-1} E_{\alpha,\alpha}(\lambda x^\alpha) + \int_0^x (x-u)^{\alpha-1} h(u) E_{\alpha,\alpha}(\lambda(x-u)^\alpha) du. \tag{A.4}$$

Now we can come back in the proof of Theorem 4.3. Indeed, starting by equation (4.14), we have

$$\begin{aligned} X_t &= \int_0^t f^{\alpha,\lambda}(t-s) \left(s + \frac{1}{\sqrt{\delta\lambda}} Z_s \right) ds \\ &= \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-\lambda(t-s)^\alpha) \lambda \left(s + \frac{1}{\sqrt{\delta\lambda}} Z_s \right) ds, \end{aligned}$$

with $X_0 = 0$. This expression has the form of (A.4), hence we conclude that X_t solves the equation

$$D^\alpha X_t + \lambda X_t = \lambda \left(t + \frac{1}{\sqrt{\delta\lambda}} Z_t \right)$$

which gives (4.15).

Appendix B

Pictures

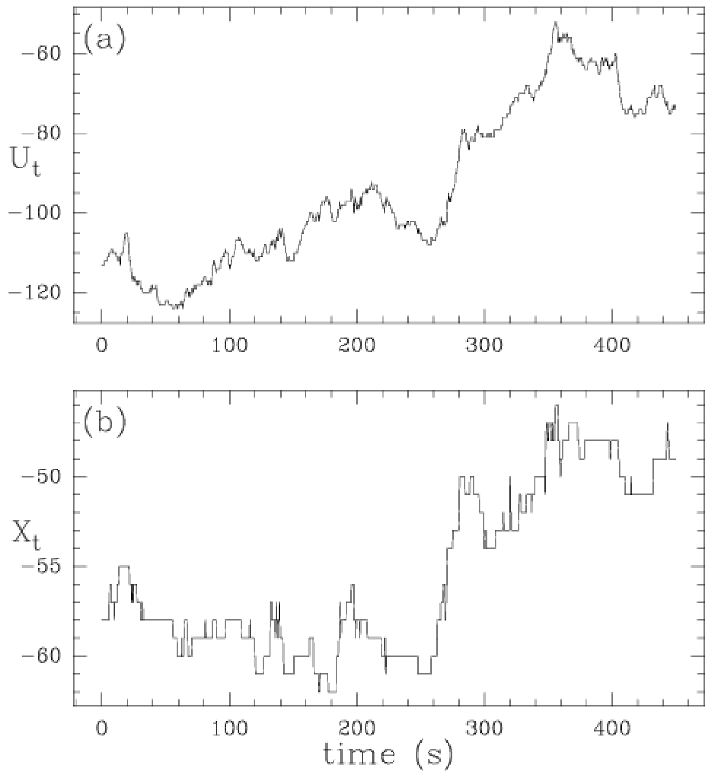

Figure B.1: Example of sample paths at microscopic scale of the processes U_t and X_t. $U_t = T_t^+ - T_t^-$ is the cumulated trade process, where T^+ and T^- are counting processes for the number of buy market orders and sell market orders respectively. $X_t = N_t^+ - N_t^-$ is the price process and it is built with the processes N^+, which counts the number of upward jumps, and N^-, which counts the number of downward jumps. The simulation is given for a model with exponential kernels and on a few minutes time interval. It is evident the fact that the price moves on a discrete grid.
Reference: Figure 1 in [BM14].

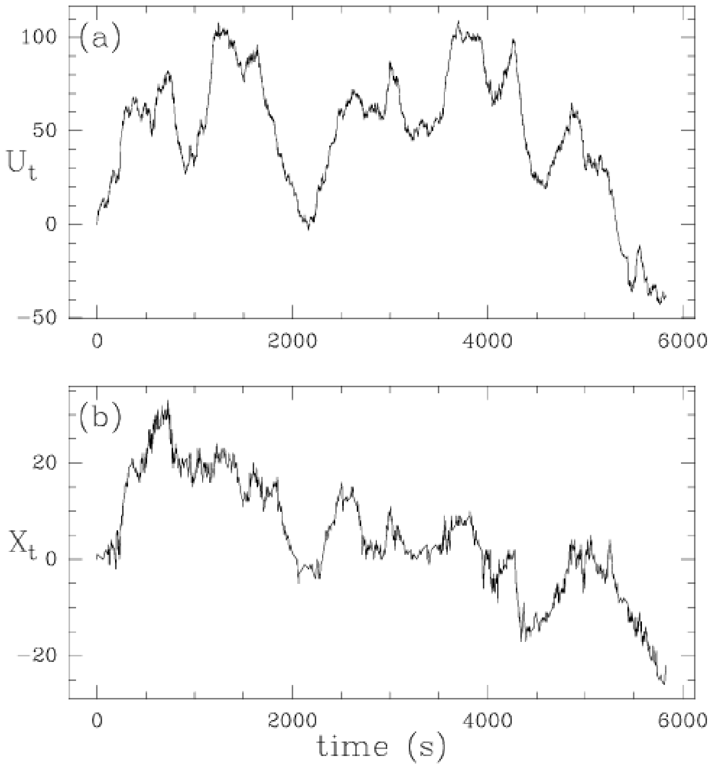

Figure B.2: Example of sample paths at macroscopic scale of the processes U_t and X_t, with the same notation of Figure B.1, but in a wider range of time. The discreteness of the movements of the previous case is now replaced by a diffusive dynamics.
Reference: Figure 2 in [BM14].

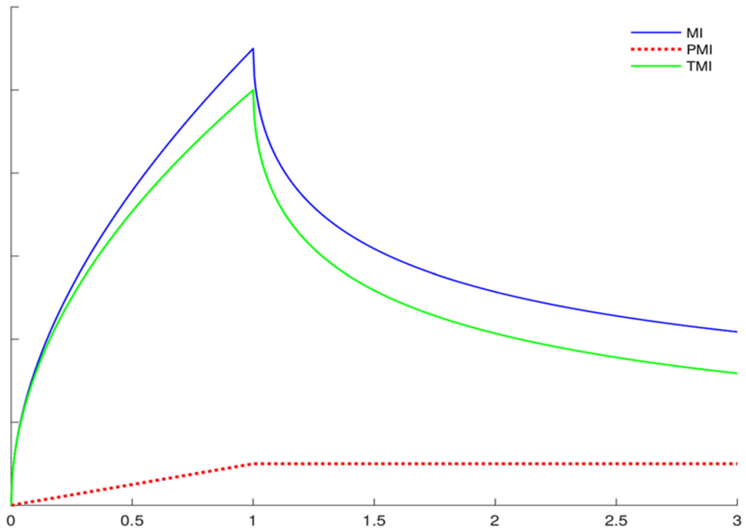

Figure B.3: Plot of market impact, transient market impact and permanent market impact for a metaorder executed uniformly, with $f = \mathbb{1}_{[0,1]}$ and $\alpha = \frac{1}{2}$.
Reference: Figure 1 in [JR20].

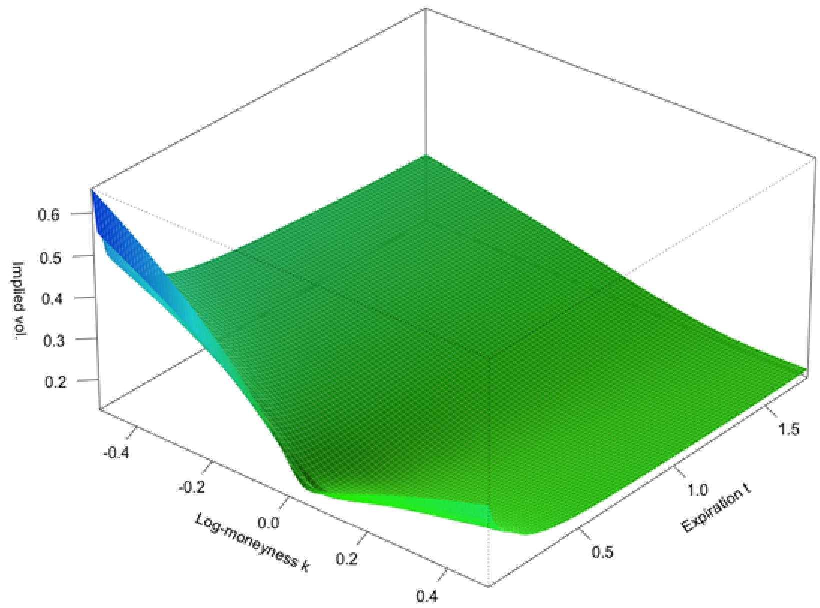

Figure B.4: The S&P volatility surface as of June 20, 2013.
The implied volatility of an option is the value of the volatility parameter in the Black-Scholes formula required to match the market price of the option. The volatility surface is the plot of the implied volatility as a function of the strike price and time to expiry. Volatility models driven by a fractional Brownian motion with Hurst exponent H close to zero manage to generate a volatility surface with the correct shape.
Reference: Figure 1.1 in [GJR18].

Bibliography

[BDHM13a] E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy. Modelling microstructure noise with mutually exciting point processes. *Quantitative Finance*, 13(1):65–77, 2013.

[BDHM13b] E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy. Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Processes and their Applications*, 123(7):2475–2499, 2013.

[BGT89] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*. Cambridge university press, 1989.

[BM14] E. Bacry and J.-F. Muzy. Hawkes model for price and trades high-frequency dynamics. *Quantitative Finance*, 14(7):1147–1166, 2014.

[CR98] F. Comte and E. Renault. Long memory in continuous-time stochastic volatility models. *Mathematical finance*, 8(4):291–323, 1998.

[EEFR18] O. El Euch, M. Fukasawa, and M. Rosenbaum. The microstructural foundations of leverage effect and rough volatility. *Finance and Stochastics*, 22:241–280, 2018.

[EEGR19] O. El Euch, J. Gatheral, and M. Rosenbaum. Roughening Heston. *Risk Magazine*, 2019.

[EER19] O. El Euch and M. Rosenbaum. The characteristic function of rough Heston models. *Mathematical Finance*, 29(1):3–38, 2019.

[FS12] V. Filimonov and D. Sornette. Quantifying reflexivity in financial markets: Toward a prediction of flash crashes. *Physical Review E*, 85(5):056108, 2012.

[FS15] V. Filimonov and D. Sornette. Apparent criticality and calibration issues in the Hawkes self-excited point process model: application to high-frequency financial data. *Quantitative Finance*, 15(8):1293–1314, 2015.

[GJR18] J. Gatheral, T. Jaisson, and M. Rosenbaum. Volatility is rough. *Quantitative Finance*, 18(6):933–949, 2018.

[Haw71a] A. G. Hawkes. Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 33(3):438–443, 1971.

[Haw71b] A. G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.

[HBB13] S. J. Hardiman, N. Bercot, and J.-P. Bouchaud. Critical reflexivity in financial markets: a Hawkes process analysis. *The European Physical Journal B*, 86:1–9, 2013.

[Hes93] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of financial studies*, 6(2):327–343, 1993.

[HO74] A. G. Hawkes and D. Oakes. A cluster process representation of a self-exciting process. *Journal of applied probability*, 11(3):493–503, 1974.

[Jac75] J. Jacod. Multivariate point processes: predictable projection, Radon-Nikodym derivatives, representation of martingales. *Probability Theory and Related Fields*, 31(3):235–253, 1975.

[Jai15] T. Jaisson. Market impact as anticipation of the order flow imbalance. *Quantitative Finance*, 15(7):1123–1135, 2015.

[JR15] T. Jaisson and M. Rosenbaum. Limit theorems for nearly unstable Hawkes processes. *Annals of Applied Probability*, 25(2):600–631, 2015.

[JR16] T. Jaisson and M. Rosenbaum. Rough fractional diffusions as scaling limits of nearly unstable heavy tailed Hawkes processes. *The Annals of Applied Probability*, 26(5):2860–2882, 2016.

[JR20] P. Jusselin and M. Rosenbaum. No-arbitrage implies power-law market impact and rough volatility. *Mathematical Finance*, 30(4):1309–1336, 2020.

[JS03] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes.* Springer-Verlag, Berlin, 2003.

[LMPR18] G. Livieri, S. Mouti, A. Pallavicini, and M. Rosenbaum. Rough volatility: evidence from option prices. *IISE transactions*, 50(9):767–776, 2018.

[MVN68] B. B. Mandelbrot and J. W. Van Ness. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, 10(4):422–437, 1968.

[RY99] D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* Springer, 1999.

[SKM93] S. Samko, A.A. Kilbas, and O. Marichev. *Fractional Integrals and Derivatives: Theory and applications.* Gordon and Breach, Yverdon, 1993.