

# Hawkes Process : Fast Calibration, Application to Trade Clustering and Diffusive Limit \*

José Da Fonseca<sup>†</sup>

Riadh Zaatour<sup>‡</sup>

August 4, 2013

## Abstract

This paper provides explicit formulas for the moments and the autocorrelation function of the number of jumps over a given interval for the Hawkes process. These computations are possible thanks to the affine property of this process. Using these quantities an implementation of the method of moments for parameter estimation that leads to an fast optimization algorithm is developed. The estimation strategy is applied to trade arrival times for major stocks that show a clustering behaviour, a feature the Hawkes process can effectively handle. As the calibration is fast, the estimation is rolled to determine the stability of the estimated parameters. Lastly, the analytical results enable the computation of the diffusive limit in a simple model for the price evolution based on the Hawkes process. It determines the connection between the parameters driving the high frequency activity to the daily volatility.

**JEL-Classification:** C13; C32; C58.

**Keywords:** Hawkes process, Calibration, High-frequency data, Trade clustering, Diffusive limit.

---

\*We are thankful to the participants of the Recent Advances in Algo and HF Trading conference (University College London, April 2013). All remaining errors are ours.

<sup>†</sup>Auckland University of Technology, Business School, Department of Finance, Private Bag 92006, 1142 Auckland, New Zealand. Phone: ++64 9 9219999 extn. 5063. Email: jose.dafonseca@aut.ac.nz

<sup>‡</sup>Corresponding author: Chair of Quantitative Finance, Ecole Centrale Paris, Grande Voie des Vignes, 92290 Châtenay-Malabry, France. Email: zaatour.riadh@yahoo.fr

# Introduction

Trading activity leads to time series of irregularly spaced points that show a clustering behaviour. This stylized property suggests the use of the Hawkes process, a point process mathematically defined by Hawkes (1971), which is an extension of the classical Poisson process that possesses this clustering property. It explains the large number of works on trading activity and more generally high-frequency econometrics based on this process as a modeling framework. To name only a few let us quote Hewlett (2006), Bowsher (2007), Large (2007), Bacry et al. (2013a) or Muni Toke and Pomponio (2011)<sup>1</sup>.

There are other stochastic processes possessing this clustering property, they are often more sophisticated than the Hawkes process in the sense that their dynamic involves several lags and strong nonlinearities. They are actively studied in the econometrics literature, see Hautsch (2012) for a general overview. As for these sophisticated stochastic processes the Hawkes process has a likelihood function which is known in closed-form. As a consequence, most of the existing literature focuses on the estimation of the dynamics. However, despite or because of its simplicity the Hawkes process has several advantages from an analytical point of view, that we will develop in this paper, that allow for very interesting applications.

First, we show how to compute in closed-form the moments of any order of the number of jumps over a given time interval. This analytical tractability even extends to the autocorrelation function of the number of jumps. As such, we can develop an estimation strategy based on these quantities which, compared with the likelihood estimation strategy, is extremely fast. This aspect is crucial when it comes to applications such as high frequency trading activity which requires a fast estimation procedure. The maximization of the likelihood function, that can take several minutes, cannot be used in real applications. What is more, with the estimation being immediate, we can roll the estimation procedure and study the parameter stability which is an essential aspect in practice.

Second, thanks to its analytical tractability we can explicitly compute the impulse function of the Hawkes process and therefore the trading activity, modeled by this process, can easily be analysed.

---

<sup>1</sup>For non financial applications see Vere-Jones (1970), Veen and Schoenberg (2008), Lewis and Mohler (2011) and Mohler et al. (2011).

Third, within a simple toy model for a stock based on the Hawkes process we compute the diffusive limit for the asset and therefore make the link between the microscopic activity (i.e. the trading activity at high frequency) to the macroscopic activity (i.e. the daily volatility as used in the Black-Scholes model) explicit. To perform such analysis the analytical tractability of the Hawkes process turns out to be essential and underlines the advantages related to the simplicity of this process. Our work falls in a new trend of the literature developed by Cont et al. (2010), Cont and De Larrard (2011), Cont and De Larrard (2012), Bacry et al. (2013a), Abergel and Jedidi (2013), Bacry et al. (2013b), Kirilenko et al. (2013) aiming at connecting these two scales (the high frequency quantities and the daily quantities).

The structure of the paper is as follows. In the first section, we describe the analytical framework which comprises the basic properties of the Hawkes process as well as the Dynkin formula that will be our main mathematical tool. Using these results, the computation of the moments and the autocorrelation function of the number of jumps over a given time interval is provided. It also contains the usual optimization algorithms used in the literature and the method of moments based on the analytical results. In a second part, we present the data, various estimation results and an impulse response analysis allowed by the model. This second part is completed with a toy model for a stock for which we derive the limit properties. Finally, we conclude and provide some technical results, tables and figures which are gathered in the appendix.

## 1 The Analytical Framework

### 1.1 Dynamics and affine structure of the moment-generating function

The Hawkes process was defined in Hawkes (1971) and is a self-excited point process whose intensity depends on the path followed by the point process. More precisely, the point process is determined by the intensity process  $(\lambda_t)_{t \geq 0}$  through the relations:

$$\mathbb{P}[N_{t+h} - N_t = 1 | \mathcal{F}_t] = \lambda_t h + o(h) \quad (1)$$

$$\mathbb{P}[N_{t+h} - N_t > 1 | \mathcal{F}_t] = o(h) \quad (2)$$

$$\mathbb{P}[N_{t+h} - N_t = 0 | \mathcal{F}_t] = 1 - \lambda_t h + o(h), \quad (3)$$

where  $(\mathcal{F}_t)_{t \geq 0}$  is a filtration on the underlying probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  containing the filtration generated by  $(N_t)_{t \geq 0}$ . The intensity follows the dynamic:

$$d\lambda_t = \beta (\lambda_\infty - \lambda_t) dt + \alpha dN_t. \quad (4)$$

A jump of  $N_t$  at a given time will increase the intensity which increases the probability of another jump thanks to equation (1) and justifies the use of the term “self-exciting” to qualify this process. The jumps tend to cluster but the process does not blow up because the drift becomes negative whenever the intensity is above  $\lambda_\infty > 0$  ( $\beta$  is by hypothesis positive) and prevents any explosion. Furthermore, applying Ito’s lemma to  $e^{\beta t} \lambda_t$  yields:

$$\lambda_t = e^{-\beta t} (\lambda_0 - \lambda_\infty) + \lambda_\infty + \int_0^t \alpha e^{-\beta(t-s)} dN_s. \quad (5)$$

From (5) we also observe that the impact on the intensity of a jump dies out exponentially as time passes. For the existence and uniqueness results we refer to chapter 14 of Daley and Jones (2008) and references therein, of particular interest is Brémaud and Massoulié (1994).

As  $t$  gets larger the impact of  $\lambda_0$ , the initial value for the intensity, vanishes leaving us with:

$$\lambda_t \sim \lambda_\infty + \int_0^t \alpha e^{-\beta(t-s)} dN_s.$$

Our presentation differs slightly from the usual one found in the literature where the Hawkes intensity is written as:

$$\lambda_t = \lambda_\infty + \int_{-\infty}^t \alpha e^{-\beta(t-s)} dN_s. \quad (6)$$

The equation (6) leads to a stochastic differential equation similar to (4), the process starts infinitely in the past and is at its stationary regime. In our case we have a dependency with respect to the initial position  $\lambda_0$  in equation (5) but, as mentioned above, for  $t$  large enough its impact will vanish.

Our presentation for the Hawkes process follows closely Errais et al. (2010) and is motivated by the fact that we want to perform stochastic differential calculus.

The process  $X_t = (\lambda_t, N_t)$  is a Markov process in the state space  $D = \mathbb{R}_+ \times \mathbb{N}$ . This is a key property that will give us very powerful tools to investigate the distributional properties of the process. Among

these tools is the infinitesimal generator. Consider a sufficiently regular function  $f : D \rightarrow \mathbb{R}$ , the infinitesimal generator of the process, denoted  $\mathcal{L}$ , is the operator acting on  $f$  such that:

$$\mathcal{L}f(x) = \lim_{h \rightarrow 0} \frac{\mathbb{E}_t^x [f(X_{t+h})] - f(x)}{h}$$

with  $\mathbb{E}_t^x[\cdot] = \mathbb{E}^x[\cdot | \mathcal{F}_t]$  and  $X_t = x$  ( $\mathbb{E}_0[\cdot] = \mathbb{E}[\cdot]$ ). In the case of a Hawkes process, this writes:

$$\mathcal{L}f(x) = \beta(\lambda_\infty - \lambda_t) \frac{\partial f}{\partial \lambda}(x) + \lambda_t [f(\lambda_t + \alpha, N_t + 1) - f(x)]. \quad (7)$$

For every function  $f$  in the domain of the infinitesimal generator, the process:

$$M_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_u) du$$

is a martingale relative to its natural filtration (see for example Proposition 1.6 of chapter VII in Revuz and Yor (1999)), thus for  $s > t$  we have:

$$\mathbb{E}_t \left[ f(X_s) - \int_0^s \mathcal{L}f(X_u) du \right] = f(X_t) - \int_0^t \mathcal{L}f(X_u) du$$

by the martingale property and we finally obtain the important Dynkin formula:

$$\mathbb{E}_t [f(X_s)] = f(X_t) + \mathbb{E}_t \left[ \int_t^s \mathcal{L}f(X_u) du \right]. \quad (8)$$

This formula allows the computation of conditional expectations of functions of the Markov process  $(\lambda_t, N_t)$  which turns out to be very useful when the expectation on the right hand side can be easily computed. In the following subsection we will rely heavily on this formula to compute some distributional properties of Hawkes process.

As underlined in Errais et al. (2010) the process  $X_t = (\lambda_t, N_t)$  is a Markov process which is *affine* which implies that a closed-form solution for the moment-generating function is available. Let  $u = (u_1, u_2)^\top \in \mathbb{R}^2$ , the conditional moment-generating function of  $X_T = (\lambda_T, N_T)$  is defined as  $f(t, X_t) = \mathbb{E}_t^x [e^{u^\top X_T}] = \mathbb{E}_t^x [e^{u_1 \lambda_T + u_2 N_T}]$ . Clearly,  $f(t, X_t)$  must be a martingale and the function  $f$  satisfies:

$$\frac{\partial f}{\partial t}(t, X_t) + \mathcal{L}f(t, X_t) = 0 \quad (9)$$

with boundary condition  $f(T, X_T) = e^{u^\top X_T}$ . As  $X_t = (\lambda_t, N_t)$  is a Markov affine point process we guess the solution of (9) is an exponential affine form of the state variable, that is to say:

$$f(t, X_t) = e^{a(t) + b(t)\lambda_t + c(t)N_t}. \quad (10)$$

Setting this guess into equation (9) we obtain the system of ordinary differential equations:

$$\frac{\partial a}{\partial t} = -\beta\lambda_\infty b(t) \quad (11)$$

$$\frac{\partial b}{\partial t} = \beta b(t) + 1 - e^{\alpha b(t) + c(t)} \quad (12)$$

$$\frac{\partial c}{\partial t} = 0 \quad (13)$$

with terminal conditions  $a(T) = 0$ ,  $b(T) = u_1$  and  $c(T) = u_2$ .

The above system of ODE fully characterizes the moment-generating function and the Laplace transform of the process which completely determines its distribution. However, an explicit solution for the equation (12) is usually not available. From the moment-generating function (10) we can retrieve the moments of the process *after* differentiating it with respect to  $u_1$  or  $u_2$ , depending on which variable is considered, and evaluating the resulting function for  $(u_1, u_2) = 0$ . This computation leads to the differentiation of the system of ODE with respect to  $u_1$  or  $u_2$ . Therefore, we can compute the different moments of the number of jumps by successively differentiating the system of ordinary differential equations. In our particular case (i.e. for a one-dimensional Hawkes process) the computations are feasible but for a multidimensional Hawkes process they soon become tedious. Also, whenever we wish to compute moments for the intensity process  $\lambda_t$  then the resolution of (12) is not possible. We provide in the appendix an example illustrating the difficulties raised by the computations of the moments.

Lastly, the computation of the autocovariance function of the number of jumps increments, that is to say:

$$\mathbb{E}_t^x [(N_{t_4} - N_{t_3})(N_{t_2} - N_{t_1})] \quad (14)$$

with  $t < t_1 < t_2 < t_3 < t_4$  can be obtained from (10) by performing successive conditioning. In that case it will introduce the intensity process, which appears on the right hand side of (10), and implies that the joint moment-generating function (i.e.  $(\lambda_t, N_t)$ ) has to be evaluated. The resulting system of ODE and its differentiation become more complicated. Because the quantity (14) is essential, it carries the clustering property of the Hawkes process, we need to develop a simpler approach to perform the computation.

In Errais et al. (2010), the authors rely on this approach to compute the expected number of jumps and higher moments, but they rapidly become tedious beyond the first moment. To overcome these

computational difficulties, that obviously increase with the dimension of the process or when the autocovariance property of the process is needed, we develop another approach that we now present.

## 1.2 Computing the moments and the autocovariance function

Our aim in this section is to compute the moments of the process  $X_t = (\lambda_t, N_t)$  and also the autocovariance of the number of jumps over a period  $\tau$ . To achieve this we rely on the infinitesimal generator of the process given by (7) and Dynkin's formula (8). In order to obtain the expected number of jumps and the expected intensity we use the following lemma:

**Lemma 1** *Given a Hawkes process  $X_t = (\lambda_t, N_t)$  with dynamic given by (4) then the expected number of jumps  $\mathbb{E}[N_t]$  and the expected intensity  $\mathbb{E}[\lambda_t]$  satisfy the set of ODE:*

$$d\mathbb{E}[N_t] = \mathbb{E}[\lambda_t] dt \quad (15)$$

$$d\mathbb{E}[\lambda_t] = (\beta\lambda_\infty + (\alpha - \beta)\mathbb{E}[\lambda_t]) dt. \quad (16)$$

Note that (16) can be explicitly computed and once obtained another integration leads to the expression for  $\mathbb{E}[N_t]$ . If these quantities are known, a similar procedure will give higher order moments as we have:

**Lemma 2** *Given a Hawkes process  $X_t = (\lambda_t, N_t)$  with dynamic given by (4) then  $\mathbb{E}[\lambda_t^2]$ ,  $\mathbb{E}[\lambda_t N_t]$  and  $\mathbb{E}[N_t^2]$  satisfy the set of ODE:*

$$d\mathbb{E}[N_t^2] = 2\mathbb{E}[\lambda_t N_t] dt + \mathbb{E}[\lambda_t] dt \quad (17)$$

$$d\mathbb{E}[\lambda_t N_t] = \beta\lambda_\infty \mathbb{E}[N_t] dt + (\alpha - \beta)\mathbb{E}[\lambda_t N_t] dt + \mathbb{E}[\lambda_t^2] dt + \alpha\mathbb{E}[\lambda_t] dt \quad (18)$$

$$d\mathbb{E}[\lambda_t^2] = (\alpha^2 + 2\beta\lambda_\infty)\mathbb{E}[\lambda_t] dt + 2(\alpha - \beta)\mathbb{E}[\lambda_t^2] dt. \quad (19)$$

These two lemmas allow the computation of different quantities useful to perform the estimation of the process. In fact, we have:

**Proposition 3** *Given a Hawkes process  $X_t = (\lambda_t, N_t)$  with dynamic given by (4) then we have the following equalities:*

$$\lim_{t \rightarrow \infty} \mathbb{E}[N_{t+\tau} - N_t] = \frac{\beta\lambda_\infty}{\beta - \alpha} \tau = \Lambda\tau \quad (20)$$

with  $\Lambda = \frac{\lambda_\infty}{1-\alpha/\beta}$  (the stationary regime expected intensity) gives the long run expected value of the number of jumps during a time interval of length  $\tau$ . The variance is:

$$\begin{aligned} V(\tau) &= \lim_{t \rightarrow \infty} \mathbb{E} [(N_{t+\tau} - N_t)^2] - \mathbb{E} [N_{t+\tau} - N_t]^2 \\ &= \Lambda \left( \tau \kappa_-^2 + (1 - \kappa_-^2) \frac{(1 - e^{-\tau \gamma_-})}{\gamma_-} \right) \end{aligned} \quad (21)$$

where:

$$\Lambda = \frac{\lambda_\infty}{1 - \alpha/\beta}, \quad \kappa_- = \frac{1}{1 - \alpha/\beta} \quad \text{and} \quad \gamma_- = \beta - \alpha.$$

The covariance is given by:

$$\begin{aligned} \text{Cov}(\tau, \delta) &= \lim_{t \rightarrow \infty} \mathbb{E} [(N_{t+\tau} - N_t)(N_{t+2\tau+\delta} - N_{t+\tau+\delta})] - \mathbb{E} [(N_{t+\tau} - N_t)] \mathbb{E} [(N_{t+2\tau+\delta} - N_{t+\tau+\delta})] \\ &= \frac{\lambda_\infty \beta \alpha (2\beta - \alpha) (e^{(\alpha-\beta)\tau} - 1)^2}{2(\alpha - \beta)^4} e^{(\alpha-\beta)\delta} \end{aligned} \quad (22)$$

for  $\delta > 0$ .

The strategy of taking the limit to simplify the dependency of the results with respect to the initial value of the process, which is unknown, is borrowed from Aït-Sahalia et al. (2010) who used the Hawkes process for modelling contagion effects between stocks. These results require the stability condition  $\frac{\alpha}{\beta} < 1$  which allows us to put the process in its long run stationary regime.

The autocovariance function contains information regarding the self-exciting or clustering property of the Hawkes process but it is more convenient to derive the autocorrelation function. We do not provide a proof for the following result as it is straightforward to obtain from the previous proposition.

**Proposition 4** *Given a Hawkes process  $X_t = (\lambda_t, N_t)$  with dynamic given by (4) then the autocorrelation function of the number of jumps over a given interval  $\tau$  is:*

$$\begin{aligned} \text{Acf}(\tau, \delta) &= \lim_{t \rightarrow \infty} \frac{\mathbb{E} [(N_{t+\tau} - N_t)(N_{t+2\tau+\delta} - N_{t+\tau+\delta})] - \mathbb{E} [(N_{t+\tau} - N_t)] \mathbb{E} [(N_{t+2\tau+\delta} - N_{t+\tau+\delta})]}{\sqrt{\text{var}(N_{t+\tau} - N_t) \text{var}(N_{t+2\tau+\delta} - N_{t+\tau+\delta})}} \\ &= \frac{e^{-2\beta\tau} (e^{\alpha\tau} - e^{\beta\tau})^2 \alpha(\alpha - 2\beta)}{2(\alpha(\alpha - 2\beta)(e^{(\alpha-\beta)\tau} - 1) + \beta^2\tau(\alpha - \beta))} e^{(\alpha-\beta)\delta} \end{aligned} \quad (23)$$

where the lag is  $\delta$ .



The above expression is always positive when  $\alpha < \beta$  which is the stability condition of the process and decays exponentially with the lag  $\delta$ . The half-life depends on the difference  $\alpha - \beta$ . Note also that the background intensity  $\lambda_\infty$  is not involved in the autocorrelation, a property that could have been expected. The complete expressions for the third and fourth moments are provided, without proof, in the appendix.

The computation performed above allows us to determine the moments up to the second order of  $(X_t)_{t \geq 0}$  as well as the autocorrelation function for the number of jumps over an interval  $\tau$ . Following this approach we can compute higher order moments. The key ingredient underlying the computations is the stability of the polynomial functions with respect to the infinitesimal generator of the Hawkes process. More precisely, the expected value of a polynomial function of the process  $(X_t)_{t \geq 0}$  (i.e.  $\sum_{i \leq n, j \leq m} a_{ij} x^i x^j$ ) can be expressed as a function of polynomial functions of same or lower degree. This property is a consequence of the affine structure of the Hawkes process and has been used for the classical standard affine model of Duffie and Kan (1996) in Cuchiero et al. (2012) and Filipović et al. (2013). To be more precise, if we denote  $Z_t = (\mathbb{E}[\lambda_t], \mathbb{E}[N_t], \mathbb{E}[\lambda_t^2], \mathbb{E}[\lambda_t N_t], \mathbb{E}[N_t^2])^\top$  then using the ODE obtained above the vector  $Z_t$  satisfies the ODE:

$$\frac{dZ_t}{dt} = AZ_t + B \quad (24)$$

with:

$$A = \begin{pmatrix} \alpha - \beta & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ \alpha^2 + 2\beta\lambda_\infty & 0 & 2(\alpha - \beta) & 0 & 0 \\ \alpha & \beta\lambda_\infty & 1 & \alpha - \beta & 0 \\ 1 & 0 & 0 & 2 & 0 \end{pmatrix} \quad B = \begin{pmatrix} \beta\lambda_\infty \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

The solution of this ODE is given by:

$$Z_t = e^{At} Z_0 + \int_0^t e^{A(t-s)} B ds \quad (25)$$

where the exponential of the matrix is computed using classical algorithms, see Golub and Van Loan (1996).

### 1.3 Inference strategies

In this subsection, we first present the classical Maximum Likelihood approach usually used to calibrate the Hawkes process and underline the numerical difficulties. Then using the explicit expression for the moments and the autocorrelation function computed in the previous subsection we develop a method of moment estimation strategy whose computational speed appears to be very fast compared to the existing alternatives.

#### 1.3.1 Maximum likelihood estimation

Let  $(X_t)_{t \geq 0}$  be a simple point process on  $[0, T]$  and  $t_1 \dots t_{N_T}$  denote a realization of  $(N_t)_{t \geq 0}$  over  $[0, T]$ , then, as established in proposition 7.2.III of Daley and Jones (2002), the log-likelihood of  $(X_t)_{t \geq 0}$  is of the form:

$$\begin{aligned} L &= \int_0^T (1 - \lambda_s) ds + \int_0^T \ln(\lambda_s) dN_s \\ &= \int_0^T (1 - \lambda_s) ds + \sum_{i=1}^{N_T} \ln(\lambda_{t_i}). \end{aligned}$$

In the case of a Hawkes process we have:

$$\begin{aligned} L &= \int_0^T (1 - \lambda_t) dt + \int_0^T \ln(\lambda_t) dN_t \\ &= \int_0^T \left( 1 - \lambda_\infty + \int_0^t \alpha e^{-\beta(t-s)} dN_s \right) dt + \int_0^T \ln \left( \lambda_\infty + \int_0^t \alpha e^{-\beta(t-s)} dN_s \right) dN_t \\ &= T - T\lambda_\infty + \sum_{i=1}^{N_T} \int_0^T \alpha e^{-\beta(t-t_i)} \times 1_{\{t \geq t_i\}} dt + \sum_{i=1}^{N_T} \ln \left( \lambda_\infty + \sum_{t_j \leq t_i} \alpha e^{-\beta(t_i-t_j)} \right) \end{aligned}$$

and simplifying the above integral

$$\begin{aligned} \int_0^T \alpha e^{-\beta(t-t_i)} \times 1_{\{t \geq t_i\}} dt &= \left[ -\frac{\alpha}{\beta} e^{-\beta(t-t_i)} \times 1_{\{t \geq t_i\}} \right]_0^T - \int_0^T -\frac{\alpha}{\beta} e^{-\beta(t-t_i)} \times \delta_{\{t=t_i\}} dt \\ &= \frac{\alpha}{\beta} - \frac{\alpha}{\beta} e^{-\beta(T-t_i)} \end{aligned}$$

we end with:

$$L = T - T\lambda_\infty - \sum_{i=1}^{N_T} \frac{\alpha}{\beta} \left( 1 - e^{-\beta(T-t_i)} \right) + \sum_{i=1}^{N_T} \ln(\lambda_\infty + \alpha A(i)) \quad (26)$$

where  $A(i) = \sum_{t_j \leq t_i} e^{-\beta(t_i-t_j)}$ .

The estimation leads to a non linear optimization algorithm such as Nelder-Mead to find the maximum of this function. We stress that for each set of parameters the evaluation of this function requires a

loop over the observations which for the problem at hand, trade clustering, is very large.

Some authors, such as Ozaki (1979), pointed out that  $A(i)$  used in (26) satisfies a recursive relation. Indeed, defining  $A(1) = 0$  then for  $i \geq 1$ ,  $A(i+1) = e^{-\beta(t_{i+1}-t_i)} \times (1 + A(i))$ , simplifies the calculation of the likelihood function and speeds up evaluation. However, the calibration still takes a few minutes and a large number of function calls are performed. Any simplification of the calibration procedure is therefore of interest.

### 1.3.2 Fast Hawkes process calibration

Even with the improvement previously presented the parameter estimation procedure based on the maximum likelihood function is still very time consuming. Having computed explicitly the moments as well as the autocorrelation function for the Hawkes process a natural estimation strategy is the generalized method of moments, see Hall (2004) for an exhaustive treatment, and Bollerslev and Zhou (2002) and Bollerslev and Zhou (2004) for an application to finance. The inference problem now writes as:

$$\hat{\theta} = \operatorname{argmin} \left\{ (M - f(\theta))^{\top} W (M - f(\theta)) \right\} \quad (27)$$

where  $M$  is the vector of empirical moments (eventually related to the autocorrelation function),  $f(\theta)$  is the vector of corresponding theoretical moments and  $W$  is a symmetric positive definite weighting matrix. The optimization problem can then be solved very quickly by Levenberg-Marquardt algorithm (we use the implementation provided by Lourakis (Jul. 2004)).

This method of estimation is known to be consistent and asymptotically normal. It can also be shown that with a suitable choice of the weighting matrix  $W$ , the estimator is asymptotically efficient in the sense that it has the smallest covariance matrix (the usual iterative algorithm to select this matrix can be used). As we are primarily interested in speed we fix ex-ante this matrix  $W$  to reduce the discrepancies between the components of the vector involved in the objective function and in that case the optimization problem turns out to be the simple least squares method (LS in the sequel). Some numerical experiments lead us to the following conclusions: the optimization problem (27) based on the mean and variance of number jumps during an interval  $\tau$  (i.e. equations (20) and (21)), and autocorrelation function (23) gives good results if calibration quality and speed are taken into account. Computation time remains negligible and is faster than the MLE.

From a numerical point of view we found it simpler and more robust to work with normalized quantities, in which case the optimization problem is:

$$\hat{\theta} = \operatorname{argmin} \left\{ \left( 1 - \frac{f(\theta)}{M} \right)^\top W \left( 1 - \frac{f(\theta)}{M} \right) \right\} \quad (28)$$

where the components of the vector  $(1 - \frac{f(\theta)}{M})$  are  $(1 - \frac{f_i(\theta)}{M_i})$  and involve the relative ratio of the theoretical moment to its empirically estimated counterpart. This is made possible because all the considered moments are different from zero<sup>2</sup>. In that case we can choose for  $W$  the identity, thus simplifying the specification of this matrix tremendously. Let us emphasize that the evaluation of the empirical moments is only made *once* during the optimization procedure and explains why this estimation strategy is intrinsically faster than the MLE.

This gives us a very appealing estimation procedure. Not only is it instantaneous and this point is crucial if the objective is to apply a model to a high-frequency problem (optimal execution, price impact analysis of a trade), but it also has the advantage of being robust against data pollution, an aspect which is very common in such data and this point was already underlined in Bacry et al. (2013a).

Lastly, let us also mention that our closed-form solutions make it possible to perform forecast experiments as they are usually done for the realized volatility (see Meddahi (2003)).

### 1.3.3 Robustness check of the fast calibrations

To assess the quality of our estimation procedure we perform the following Monte Carlo experiment. We randomly generate the parameters  $\lambda_\infty$ ,  $\alpha$  and  $\beta$ , with  $\beta > \alpha$ . We choose them in the interval  $[0, 1]$  because these are the values obtained in our empirical study. We simulate a Hawkes process for a duration of  $T = 8$  hours, corresponding to a trading day, using the thinning algorithm as presented in Ogata (1981). We carry out a calibration using both the MLE and the LS estimators. For the MLE, calibration is conducted using Nelder-Mead algorithm as implemented in the open-source library NL-opt<sup>3</sup>. The LS estimations are conducted using the Levenberg-Marquardt algorithm

---

<sup>2</sup>A similar scaling appears in Aït-Sahalia et al. (2010) where the authors use the matrix  $W$  to control the discrepancies between the moments.

<sup>3</sup>See <http://ab-initio.mit.edu/wiki/index.php/NLopt>

provided by Lourakis (Jul. 2004). The initial guess for the parameters is their true value multiplied by a uniform random variable on the interval  $[0.5 \ 1.5]$  (both algorithms have the same starting point). The estimation test is performed 25000 times for each algorithm. Table 1 contains the results. The mean relative error value expressed in percent is reported for each parameter. For the standard deviation we compute the root mean squared relative error value that we also express in percent.

[ Insert Table 1 here ]

We perform two fast calibrations. The first is based on the least squares optimization problem (28) with  $W$  the identity, the first two moments given by (20) and (21) computed with  $\tau = 60$  seconds as well as the autocorrelation function (23) with  $\tau = 60$  seconds and lags ranging from 0 to 600 seconds (by step of 60 seconds). For the second fast calibration, the objective function (28) only depends on the autocorrelation function (23) with  $\tau = 60$  seconds and lags ranging from 0 to 600 seconds (by step of 60 seconds). To obtain  $\lambda_\infty$ , because the autocorrelation function does not depend on it, we use (20) with  $\tau = 60$  seconds. That is, we suppose this equation is satisfied without error.

The MLE leads to an error of 0.33% for  $\lambda_\infty$  and similar magnitudes for  $\alpha$  and  $\beta$  as we have 0.16% and 0.23%, respectively. For example, the difference between the estimated  $\lambda_\infty$  and the true  $\lambda_\infty$  is 0.33% of the true value for this estimator. Overall, the estimation procedure is satisfactory as mean values and standard deviations are small. For the first fast calibration, the results are reported in column titled “Fast calibration I”, the mean values are also small, they are comparable to those obtained with the MLE, and the standard deviation values are roughly 1.5 times those for the MLE but remain acceptable. In conclusion, the fast calibration algorithm performs well and the computational cost is much lower. For the second fast calibration (i.e. column “Fast calibration II”) similar conclusions are obtained. Namely, the mean values and the standard deviations are small. Overall, the fast calibration algorithms perform well.

## 2 Applications

### 2.1 Data

We rely on tick-by-tick data from TRTH (Thomson Reuters Tick History). The data consist of trades and quotes files timestamped in milliseconds. We studied two stocks in particular: BNP Paribas and

Sanofi, as well as the futures on the Eurostoxx and the Dax. For each studied day we consider the futures with the shortest maturity date. The data cover the period between 2010/01/01 to 2011/12/31.

All the trading days begin at 9:00 and end at 17:30. We neglect the first and last 15 minutes in order to avoid the open and close auctions. We end with 8 trading hours per day, between 9:15 and 17:15.

Most of our study deals with trade time arrivals and statistics on the number of trades occurring on intervals of fixed length. The fact that the timestamps have a bounded precision, the millisecond in our case, is another appealing feature of our method of estimation. Indeed, many trades will have the same time to the nearest millisecond even if they did not take place at the same time. This millisecond will count as a unique entry in the ML estimation procedure; in the moment based inference all the trades will be taken into account when computing the moments. This, among other reasons, will make the method of moments more robust to data imprecision, a fact typical to high frequency data.

## 2.2 Trade clustering

As outlined in the introduction, trading activity is not a completely random and memory-less process. Consequently, the Poisson process is not suitable for modelling trade arrival times. As shown in Figure 1, a qqplot of interarrival times of trades against an exponential distribution clearly rejects the Poisson process as the data generating process for the order flow.

[ Insert Figure 1 here ]

In fact, trades tend to cluster and an illustration is given in Figure 2 where we plot an histogram of the number of trades occurring every minute during a trading day for the Eurostoxx. The clustering is graphically clear.

[ Insert Figure 2 here ]

Numerous reasons can explain this clustering of trade arrival times, among them liquidity takers splitting their orders so as to minimize their market impact, or insider traders reacting rapidly to take advantage from information they have before it is widespread in the market : these justify a one sided

trade clustering (i.e. either buy or sell initiated trades). On the other hand, heterogeneity of market participants is responsible for the two-sided trade clustering. A complete study can be found in Sarkar and Schwartz (2006).

To quantify this clustering in time we compute the correlation of the number of trades occurring during different time intervals of fixed length separated by a time lag. A plot of this autocorrelation as a function of the lag gives information about the degree of clustering. If the data were generated by a Poisson process this autocorrelation would be equal to zero.

We want to compute:

$$C(\tau, \delta) = \frac{\mathbb{E}[(N_{t+\tau} - N_t)(N_{t+2\tau+\delta} - N_{t+\tau+\delta})] - \mathbb{E}[(N_{t+\tau} - N_t)]\mathbb{E}[(N_{t+2\tau+\delta} - N_{t+\tau+\delta})]}{\sqrt{\text{var}(N_{t+\tau} - N_t)\text{var}(N_{t+2\tau+\delta} - N_{t+\tau+\delta})}}. \quad (29)$$

To this end, we count the number of trades occurring during two sliding non overlapping intervals of length  $\tau = 60$  seconds separated by a certain time lag  $\delta$ . We change this time lag  $\delta$  from 1 second to 20 minutes by a step of 1 second. This gives 1200 number of trades autocorrelation points for the function  $\delta \rightarrow C(\tau, \delta)$  that we report in Figure 3 for different time intervals  $\tau$  for a given symbol (in that case the Eurostoxx). We clearly see that the autocorrelation is positive and significant and that it decreases with the time lag. It is also remarkable that this is true for all the time interval lengths considered, and that independently from  $\tau$  this memory effect seems to take around 10 minutes to become insignificant.

[ Insert Figure 3 here ]

Figure 4 confirms that the same phenomenon is observed for all the tested symbols. The absolute value of the autocorrelation is higher for the two futures, which are far more liquid than the stocks, but the same decreasing shape is observed and the time life of this autocorrelation seems to be very close for all the symbols.

[ Insert Figure 4 here ]

Overall, these stylized facts justify the use of the Hawkes process as modelling framework.

Thanks to the closed-form solutions of the previous section we estimate the parameters using the fast calibration algorithm presented in the inference subsection. To further reduce computational cost the objective function (28) is such that it depends *only* on the empirical autocorrelation function (29) and its analytical counterpart (23)<sup>4</sup>. As the autocorrelation function of the number of trades is independent of  $\lambda_\infty$  we rely on equation (20) to obtain  $\lambda_\infty$  from the other parameters<sup>5</sup>. Based on Figure 5 we choose to fit the autocorrelation function (23) for  $\tau = 60$  seconds and  $\delta$  ranging from 0 to 600 seconds by step of 60 seconds because Figure 4 convinces us that this choice is suitable for all the analysed stocks. For each symbol we perform a daily calibration using the two-year sample described in the data subsection. In Table 2 we report the mean and median estimated values as well as the standard deviations. The table also contains in column “Poisson- $\lambda$ ” the intensity we obtain if a Poisson process is calibrated.

[ Insert Figure 4 here ]

[ Insert Table 2 here ]

A striking fact is that  $\lambda_\infty$  is much smaller than the Poisson- $\lambda$  as the former amounts to only 5% to 10% of the latter.  $\lambda_\infty$  represents the background activity of the Hawkes process whilst from Proposition 3 the long term expected intensity, that will be close to the Poisson intensity, is  $\Lambda = \frac{\lambda_\infty}{1-\alpha/\beta}$ . As such, the discrepancy between  $\lambda_\infty$  and the Poisson intensity is precisely due to  $\alpha$  which controls the self-exciting property of the Hawkes process. Our results suggest that the fraction of events due to the background process activity is small in comparison of the part due to the branching structure of the process (i.e. the self-exciting property captured by  $\alpha$ ). For the four symbols the parameters have the same size and the standard deviations are small. Note also that the stability condition ( $\alpha < \beta$ ) is satisfied although the optimization algorithm is performed without constraints. The quality of the fit can be assessed graphically, we report an example of such fit in Figure 6; it suggests that the Hawkes process captures well the empirical property of trade arrival times.

[ Insert Figure 6 here ]

For robustness, we also conduct calibrations of the model on a variety of traded assets ranging from interest rate futures to commodity, energy and foreign exchange futures. Results are reported in Table 3.

---

<sup>4</sup>Compared with the robustness check performed previously it means that we drop the first and second moments.

<sup>5</sup>From a statistical point of view it means that we suppose equation (20) is satisfied without error.



[ Insert Table 3 here ]

For all the assets, the standard deviation values are small compared to the mean parameter values and the similarity between the mean and median values, denoting the relative stability of the calibration and absence of outliers, confirm the ability of the Hawkes process to fit the data. As for the four previous assets all the calibrations lead to solutions that satisfy the stability constraint  $\alpha < \beta$ , albeit no constraints were used during the optimization procedure. Lastly, let us underline the fact that  $\alpha - \beta$  controls the decreasing shape of the autocorrelation function as it is clear from equation (23) and quantifies to which extent the data generating process departs from a Poisson process or equivalently to which extent the clustering behaviour of the Hawkes process is fruitful.

### 2.3 Branching structure of trading activity

An interesting property of the Hawkes process is its branching structure. Indeed, the occurrence of a jump increases the intensity of the process, thereby the probability to observe another jump. As pointed out by Hewlett (2006), this results in a direct and indirect impulse response of the process intensity to a jump event. Denoting the expected increase of the process intensity at time  $t$  as a response to a jump occurring at time 0 by  $f(t)$ , we have the following decomposition:

- Direct response: an increase of the intensity by  $\alpha$  that will decay exponentially as time passes, thus leading to an increase of the intensity at time  $s$  of  $\alpha e^{-\beta s}$ .
- Indirect response: at any time  $s$  between 0 and  $t$ , the direct increase of the intensity by  $\alpha e^{-\beta s}$  leads to an indirect increase of the expected number of jumps at time  $t$  which equals to  $\alpha e^{-\beta s} ds f(t - s)$ ; we then need to integrate over the range  $[0, t]$  to obtain the total indirect effect.

As a consequence, the expected direct and indirect increase of the intensity at time  $t$  caused by a jump at time 0 writes as:

$$f(t) = \alpha e^{-\beta t} + \int_0^t \alpha e^{-\beta s} f(t - s) ds.$$

The solution of this integral equation is given by:

$$f(t) = \alpha e^{-(\beta - \alpha)t}.$$

Therefore, the  $N_{response}$  which is the expected number of jumps triggered by one jump occurring at time 0 if the process is observed indefinitely (the impulse response) is:

$$N_{response} = \int_0^\infty f(s)ds = \frac{\alpha}{\beta - \alpha}. \quad (30)$$

We use our daily calibrations on real data of the Hawkes process to measure an average  $N_{response}$  for the studied assets and report the results in Table 4.

[ Insert Table 4 here ]

We clearly observe a difference between futures, given by the symbols Dax and Eurostoxx, and the stocks represented by BNPP and Sanofi. This points towards considering this number as an indicator of liquidity and trading activity. Formula (30) suggests the ratio  $\frac{\beta}{\alpha}$  as the key quantity to evaluate the impulse response value and the numbers are consistent with the fact that futures are more actively traded than the stocks due to a stronger branching structure (controlled by  $\alpha$  and  $\beta$ ).

A robustness check was also performed for other assets in Table 3. According to this measure a trade on the Bund triggers more other trades than does a trade on the Bobl. The Euro currency, and to a lesser extent the JPY, seem to be more reactive markets than the others. Among the commodities, the Crude Oil Brent dominates the Natural Gas, Sugar, Corn and the Wheat.

## 2.4 Diffusive limit and signature plot

In all the preceding sections we dealt with the trading process from a microscopic point of view, that is to say at transaction level. In the classical High-Frequency literature, mainly developed by econometricians, most of the studies are carried out at this time scale and aim at explaining the price formation process. Many models encompass the subtle interactions of the many components of the trading process such as order flow, order signs, volumes and other quantities to achieve a microscopic foundation of the price process. For instance, the ACD models, as proposed in Engle and Russell (1998), fall in this trend of research. For an overview of models and techniques we refer to Hautsch (2012) and references therein.

In our paper, we consider a simpler framework focusing only on the order flow. Obviously, this simplification comes at the cost of neglecting important aspects of the price formation process, such as

the volume for example, but it allows us to address the important question of connecting the microscopic price formation process observed at transaction level to its macroscopic properties at a coarser time scale. In other words, we connect the stochastic differential equations used to model an asset price evolution at a daily frequency, such as in the Black-Scholes model which relies mainly on the continuous Brownian motion, to the discontinuous point process describing individual transactions. The Hawkes process, thanks to its strong analytical tractability, enables us to relate these two time scales.

In recent years, many authors developed this bottom-up point of view in price modelling, establishing connections between order-book level price formation mechanisms and statistical macroscopic price properties. Among other references let us mention Abergel and Jedidi (2013) whose model for the order book is based on a multidimensional Markov chain with independent Poissonian order arrival times. Within their framework they prove the convergence of the price process to a Brownian motion, thus bridging the gap between high frequency and low frequency quantities. In Cont and De Larrard (2011) and Cont and De Larrard (2012), the order book is described as a Markovian queuing system for which the authors establish a diffusive limit and calculate some quantities of interest such as volatility. In Kirilenko et al. (2013), the authors use this same idea of different time scales and relate microscopic causes to macroscopic effects, they study the influence of high frequency traders on asset volatility. Bacry et al. (2013a) introduces a model for microstructure price evolution based on mutually exciting Hawkes processes. They connect the signature plot of volatility and Epps effect of asset correlations to the model parameters driving the price process. Also of great interest is Bacry et al. (2013b) where the authors establish diffusive limits for such kind of models.

In this subsection, we consider the modelling framework proposed by Bacry et al. (2013a)<sup>6</sup> and develop a toy model for the movements of the *mid* price of a traded asset using the Hawkes processes presented in the analytical part. The *mid* price is the mean of the best ask price and the best bid price in the order book. As the best ask price or the best bid price moves up (down) by one tick the mid price will move (down) by *half* a tick. Despite its simplicity, the model captures the essential features of the price process.<sup>7</sup> The model writes:

---

<sup>6</sup>We might use the symbol “BDHM” which stands for Bacry, Delattre, Hoffmann, and Muzy.

<sup>7</sup>For instance, the so-called trade-throughs Pomponio and Abergel (2013) (i.e. trades consuming many successive limits and then moving the best quote by more than one tick) can be regarded in this model as successive one-tick movements occurring very closely in time.

$$S_t = S_0 + \left( N_t^{up} - N_t^{down} \right) \frac{\delta}{2} \quad (31)$$

where  $\delta$  is the tick value. The  $N_t^{up}$  and  $N_t^{down}$  are Hawkes processes capturing the up and down jumps of the mid price. Both of them follow a dynamic of the form (4). We consider them independent and with the same parameters in order to avoid price explosion. In the stationary regime their intensities are given by:

$$\lambda_t^{up} = \lambda_\infty + \int_0^t \alpha e^{-\beta(t-s)} dN_s^{up} \quad (32)$$

$$\lambda_t^{down} = \lambda_\infty + \int_0^t \alpha e^{-\beta(t-s)} dN_s^{down}. \quad (33)$$

Later, we provide the specification of Bacry et al. (2013a)'s model and explain the differences. In order to relate this high frequency description for the price to its low frequency description behaviour we need a limit theorem. In Bacry et al. (2013b), the authors rely on the martingale theory and limit theorems for semi-martingales to prove stability and convergence results for a general model with mutually exciting processes and a generic kernel<sup>8</sup>. In our case, as the kernel is exponential the process  $X_t = (S_t, N_t^{up}, \lambda_t^{up}, N_t^{down}, \lambda_t^{down})$  is a Markov process. Its infinitesimal generator writes:

$$\begin{aligned} \mathcal{L}f(x) &= \beta (\lambda_\infty - \lambda_t^{up}) \frac{\partial f}{\partial \lambda^{up}}(x) + \beta (\lambda_\infty - \lambda_t^{down}) \frac{\partial f}{\partial \lambda^{down}}(x) \\ &\quad + \lambda_t^{up} \left[ f \left( S_t + \frac{\delta}{2}, N_t^{up} + 1, \lambda_t^{up} + \alpha, N_t^{down}, \lambda_t^{down} \right) - f(x) \right] \\ &\quad + \lambda_t^{down} \left[ f \left( S_t - \frac{\delta}{2}, N_t^{up}, \lambda_t^{up}, N_t^{down} + 1, \lambda_t^{down} + \alpha \right) - f(x) \right]. \end{aligned}$$

The explicit form of the infinitesimal generator allows us to apply Foster-Lyapounov techniques in order to establish stability results. We refer to Meyn and Tweedie (2009) for a detailed exposition on such techniques as well as stochastic stability concepts. For instance, ergodicity of the process  $X_t$ , that is to say its convergence to a stationary regime, can be easily established thanks to the Foster-Lyapounov test function criterion. In our case, we define the function  $V(x) = \frac{\lambda^{up} + \lambda^{down}}{2\lambda_\infty}$ , then a simple calculation yields the *geometric drift condition*:

$$\mathcal{L}V(x) \leq (\alpha - \beta)V(x) + \beta \quad (34)$$

---

<sup>8</sup>The function  $g(t) = \alpha e^{-\beta t}$  is called the kernel of the Hawkes process. Other forms are possible but this choice leads to the most tractable Hawkes process.

which grants, thanks to Theorems 6.1 and 7.1 in Meyn and Tweedie (1993)(and especially (CD3)), the V-uniform ergodicity of the process  $X_t$ .

Let us then write unit-time price increments:

$$\eta_i = \left[ (N_i^{up} - N_{i-1}^{up}) - (N_i^{down} - N_{i-1}^{down}) \right] \times \frac{\delta}{2}$$

and consider the random sums:

$$S_n = \sum_{i=1}^n \eta_i$$

with  $\{\eta_i; i = 1 \dots n\}$  being the price increments (note that  $\mathbb{E}[\eta_i] = 0$ ). We focus on the asymptotic behaviour of the rescaled price process:

$$\bar{S}_t^n = \frac{S_{[nt]}}{\sqrt{n}}.$$

The V-uniform ergodicity and Theorem 16.1.5 in Meyn and Tweedie (2009) allows us to conclude that the increments are geometrically mixing and Theorem 19.3 of Billingsley (1999) proves that  $\bar{S}_t^n$  converges to a Brownian motion in the sense of Skorokhod topology:

$$\bar{S}_t^n \Rightarrow \sigma W_t.$$

Moreover, calculations done before for the moments of the Hawkes process increments lead to a very simple expression for the volatility. In fact, we have:

$$\begin{aligned} \sigma^2 &= \lim_{n \rightarrow \infty} \frac{\text{Var}(S_n)}{n} \\ &= 2 \frac{\delta^2}{4} (\mathbb{E}[(N_1^{up} - N_0^{up})^2] - \mathbb{E}[N_1^{up} - N_0^{up}]^2) \\ &\quad + 4 \frac{\delta^2}{4} \sum_{n=1}^{\infty} \mathbb{E}[(N_1^{up} - N_0^{up})(N_{1+n}^{up} - N_n^{up})] - \mathbb{E}[N_1^{up} - N_0^{up}] \mathbb{E}[N_{1+n}^{up} - N_n^{up}] \\ &= \frac{\delta^2}{2} \left( V(1) + 2 \sum_{n=0}^{\infty} \text{Cov}(1, n) \right) \end{aligned}$$

where thanks to (22)  $\text{Cov}(1, n)$  writes:

$$\text{Cov}(1, n) = \frac{\lambda_{\infty} \beta \alpha (2\beta - \alpha) (e^{(\alpha - \beta)} - 1)^2}{2(\alpha - \beta)^4} e^{(\alpha - \beta)n}.$$

Then, summing up with the expression of the variance and after some simplifications we obtain:

$$\sigma^2 = \frac{\delta^2}{2} \frac{\lambda_{\infty}}{\left(1 - \frac{\alpha}{\beta}\right)^3}. \quad (35)$$

Note the dependence of the volatility on the ratio  $\frac{\alpha}{\beta}$ . The larger this ratio is, the larger is the volatility is (under the hypothesis of stability  $\alpha < \beta$ ). An upward (downward) chock is likely to trigger another upward (downward) chock if this ratio is large and therefore it induces a positive autocorrelation for the mid price and a more persistent path with the effect of increasing asset's volatility.

Besides giving a framework that allows the connection of the microscopic price formation mechanism to its macroscopic behaviour, as shown above, the Hawkes process can reproduce some stylized facts across time scales. Among these stylized facts is the volatility signature plot which depends on the realized variance over a period  $T$  calculated by sampling the data by time intervals of length  $\tau$ . Within the toy model (31) we have:

$$\begin{aligned}\hat{C}(\tau) &= \frac{1}{T} \sum_{n=0}^{T/\tau-1} (S_{(n+1)\tau} - S_{n\tau})^2 \\ &= \frac{1}{T} \sum_{n=0}^{T/\tau-1} \left( (N_{(n+1)\tau}^{up} - N_{n\tau}^{up}) - (N_{(n+1)\tau}^{down} - N_{n\tau}^{down}) \right)^2 \frac{\delta^2}{4} \\ &= \frac{1}{T} \sum_{n=0}^{T/\tau-1} (N_{(n+1)\tau}^{up} - N_{n\tau}^{up})^2 \frac{\delta^2}{4} + \frac{1}{T} \sum_{n=0}^{T/\tau-1} (N_{(n+1)\tau}^{down} - N_{n\tau}^{down})^2 \frac{\delta^2}{4} \\ &\quad - 2 \frac{1}{T} \sum_{n=0}^{T/\tau-1} (N_{(n+1)\tau}^{up} - N_{n\tau}^{up}) (N_{(n+1)\tau}^{down} - N_{n\tau}^{down}) \frac{\delta^2}{4}.\end{aligned}$$

By definition the mean signature plot is the expectation of the above quantity and can be computed explicitly as we have (we omit the proof as it is straightforward):

**Proposition 5** *The mean signature plot, defined by  $C(\tau) = \mathbb{E}[\hat{C}(\tau)]$ , is given by:*

$$\begin{aligned}C(\tau) &= \mathbb{E}[\hat{C}(\tau)] \\ &= \frac{\delta^2}{2\tau} V(\tau) \\ &= \frac{\delta^2}{2} \Lambda \left( \kappa_-^2 + (1 - \kappa_-^2) \frac{(1 - e^{-\tau\gamma_-})}{\tau\gamma_-} \right)\end{aligned}$$

where:

$$\Lambda = \frac{\lambda_\infty}{1 - \alpha/\beta}, \quad \kappa_- = \frac{1}{1 - \alpha/\beta} \quad \text{and} \quad \gamma_- = \beta - \alpha.$$

We eventually refer to the signature plot instead of the mean signature plot. Notice that when  $\tau$  becomes larger the above expression converges to the asymptotic diffusive variance of the model calculated in (35). The mean signature plot is an increasing function with respect to  $\tau$  (or equivalently the signature plot is decreasing with the sampling frequency) and this is due to the positive serial autocorrelation of the returns. This captures situations as the one observed in Figure 7.

[ Insert Figure 7 here ]

Lastly, within our simple toy model we can determine the autocorrelation function of the price increments computed over intervals of size  $\tau$  and lagged by  $\delta$ , it is given by:

$$CorrStock(\tau, \delta) = \frac{\mathbb{E}[(S_{t+\tau} - S_t)(S_{t+2\tau+\delta} - S_{t+\tau+\delta})]}{\sqrt{\mathbb{E}[(S_{t+\tau} - S_t)] \mathbb{E}[(S_{t+2\tau+\delta} - S_{t+\tau+\delta})]}} \quad (36)$$

which thanks to the closed-form formula obtained in the analytical section leads to:

$$\frac{e^{-2\beta\tau+\delta(\alpha-\beta)} (e^{\tau\beta} - e^{\tau\alpha})^2 \alpha (2\beta - \alpha)}{2\beta^2 (\beta - \alpha)}. \quad (37)$$

The stability condition (i.e.  $\alpha < \beta$ ) ensures the positiveness of this serial autocorrelation and it confirms the intuition developed above. Moreover, the decay of the autocorrelation as a function of the lag depends on the parameter  $\alpha$  which tends to reduce the decaying behaviour of this function.

In order to compare our results to Bacry et al. (2013a) let us briefly recapitulate their main findings. In Bacry et al. (2013a), the authors propose a model similar to (31) but with Hawkes processes that are *mutually-excited* and not *self-excited* as in our case. To be more precise the dynamics for the intensities are given by:

$$\lambda_t^{up} = \lambda_\infty + \int_0^t \alpha e^{-\beta(t-s)} dN_s^{down} \quad (38)$$

$$\lambda_t^{down} = \lambda_\infty + \int_0^t \alpha e^{-\beta(t-s)} dN_s^{up}. \quad (39)$$

Notice that an up jump increases the down intensity that increases the probability of a down jump and if this one occurs it will increase the up jump intensity. The process is purely mutually excited and possesses a mean reversion behaviour. These dynamics are significantly different from those driving

the toy model. The diffusive limit for this model is:

$$\sigma_{\text{BDHM}}^2 = \frac{\delta^2}{2} \frac{\lambda_\infty}{\left(1 - \frac{\alpha}{\beta}\right) \left(1 + \frac{\alpha}{\beta}\right)^2} \quad (40)$$

and the mean signature plot is:

$$C(\tau) = \frac{\delta^2}{2} \Lambda \left( \kappa_+^2 + (1 - \kappa_+^2) \frac{1 - e^{-\tau\gamma_+}}{\tau\gamma_+} \right)$$

with:

$$\Lambda = \frac{\lambda_\infty}{1 - \alpha/\beta}, \quad \kappa_+ = \frac{1}{1 + \alpha/\beta} \quad \text{and} \quad \gamma_+ = \alpha + \beta.$$

In Bacry et al. (2013a)'s model an upward chock will increase the down intensity and trigger a downward chock on the mid price, thereby leading to a mean reverting behaviour for the mid price. Therefore, as a function of the sampling period the signature plot is decreasing with respect to  $\tau$  (or equivalently the signature plot is increasing with the sampling frequency) because of this negative serial autocorrelation of the returns. In Andersen et al. (1999) the authors mention "the patterns of bias injected in realized volatility as underlying returns are sampled progressively more frequently". This bias can lead to a decreasing or increasing volatility as a function of the sampling period. The decreasing pattern is the most frequent but inverted pattern is also possible and can be found in data, even for liquid stocks, as illustrated in Figure 7. Bacry et al. model is compatible with a decreasing pattern whilst the toy model is compatible with an increasing pattern<sup>9</sup>. Due to the positive (negative) autocorrelation of the returns in the toy (Bacry et al.) model we have, for a given pair  $(\alpha, \beta)$ , the inequality  $\sigma > \sigma_{\text{BDHM}}$ .

Within the specification (38) and (39) the autocorrelation function (36) can be computed and is given by:

$$- \frac{e^{-(\delta+2\tau)(\beta+\alpha)} (e^{\tau(\beta+\alpha)} - 1)^2 \alpha (2\beta + \alpha)}{2\beta^2 (\beta + \alpha)}. \quad (41)$$

This quantity is negative and this confirms the intuition. Also, the decay of the autocorrelation as a function of the lag depends on parameter  $\alpha$ , now controlling the mutual excitation property of the

---

<sup>9</sup>Let us stress that these figures are based on mid price sampling so that any noise due to bid-ask bounce is eliminated.



process, and this parameter increases the decaying behaviour of the function.

Notice that both models produce different asymptotic volatility formulas. In order to assess their plausibility, we calibrate a Hawkes process to the mid price up-jumps and calculate the asymptotic volatilities for the two models; the toy model given by (35) and Bacry et al. (2013a)’s specification (40). For the toy model the parameter estimation is performed as in the previous subsection. We use only the autocorrelation function with  $\tau = 60$  seconds and  $\delta$  ranging from 0 to 600 seconds by step of 60 seconds and  $\lambda_\infty$  is deduced thanks to (20) with  $\tau = 60$  seconds (it corresponds to the “Fast calibration II”). For Bacry et al. (2013a)’s model we use the MLE algorithm<sup>10</sup>.

Results are reported in Table 5 (apart from the asymptotic volatilities we only report the estimated parameters for the toy model). We rescaled the obtained volatilities by the spot value in order to obtain the more usual Black-Scholes volatility corresponding to a lognormal model. One can clearly see that our model systematically overestimates volatility whereas Bacry et al. (2013a)’s model systematically underestimates it. Qualitatively, this is because our model magnifies up and down movements, thus inflating the realized volatility, whereas Bacry et al. (2013a)’s model moderates up and down movements thanks to its mean reversion mechanism, thus underestimating the volatility.

[ Insert Table 5 here ]

Our results suggest a more general specification allowing for both self and mutual excitations. The particularly simple approach we adopted in our calculations based on the infinitesimal generator and Dynkin’s formula may be generalized to a multidimensional setting and may lead to tractable results. It would allow for both effects, that were underlined in the empirical literature (see Hautsch (2012)), on the price dynamic and provide a deeper understanding of how the volatility measured at a macroscopic level (at a daily or low frequency) depends on the trading activity observed at a microscopic level (at high frequency).

---

<sup>10</sup>For this model we only report the volatility value  $\sigma_{\text{BDHM}}$ , the parameter values are available upon request.

### 3 Conclusion

In this paper we explicitly compute the moments and the autocorrelation function of the number of jumps over an interval for the Hawkes process. Using these quantities we develop a method of moments estimation strategy which is extremely fast compared with the usual maximum likelihood estimation strategy. This aspect is essential as we are interested in the trade clustering activity observed in high frequency data or if we wish to apply the model in real time. We use our estimation framework to calibrate the Hawkes process on trades for four stocks over a two-year sample. The Hawkes process can cope with the trade clustering effect thanks to its autocorrelation structure. As our calibration is fast we roll the daily estimation over two years to analyse the parameters stability, and they are found to be reasonably stable. We perform a robustness check on other assets and obtain similar results.

Thanks to the analytical tractability of the Hawkes process we explicitly compute the impulse response associated with the process which determines the market impact of a trade. We use this as a measure of liquidity and the estimated parameters lead to reasonable conclusions.

Lastly, within a simple model based on the Hawkes process we explicitly compute the diffusive limit for the price process. This allows us to connect the microscopic dynamic, that is to say the high frequency dynamic, to the macroscopic dynamic, the volatility computed at a daily frequency (with the Black-Scholes volatility being the most well-known quantity).

Our work points towards several extensions. First, we computed the diffusive limit under the restrictive hypothesis that the Hawkes processes are only self excited whereas in Bacry et al. (2013a), on which we heavily rely, the Hawkes processes are only mutually excited. The reality should lie between the two and requires Hawkes processes that are both self and mutually excited. To this end we would need to perform the computations in the multidimensional case. We suspect that they can be performed for that case.

Another aspect of interest is the diffusive limit concept. In this work we connect the dynamic driving the trade process, using a Hawkes process, to the daily volatility. It would be of interest to go further at the microscopic level by modelling, for example, the level I quotes. The Hawkes process provides a natural modelling framework and would extend the interesting existing models based on the Poisson

process. To compute the diffusive limit for a model based on the Hawkes process the moments as well as the autocorrelation are needed and they can be obtained using the computation strategy developed in this work. These interesting problems are left for future investigation.

## References

- F. Abergel and A. Jedidi. A Mathematical Approach to Order Book Modeling. *International Journal of Theoretical and Applied Finance*, 16(5), 2013.
- Y. Aït-Sahalia, J. Cacho-Diaz, and R. J.A. Laeven. Modeling financial contagion using mutually exciting jump processes. Working paper, National Bureau of Economic Research, 2010.
- T. G. Andersen, T. Bollerslev, F. X. Diebold, and P. Labys. (understanding, optimizing, using and forecasting) realized volatility and correlation. New York University, Leonard N. Stern School Finance Department Working Paper Seires 99-061, New York University, Leonard N. Stern School of Business-, Oct 1999.
- E. Bacry, S. Delattre, M. Hoffmann, and J.-F. Muzy. Modelling microstructure noise with mutually exciting point processes. *Quantitative Finance*, 13:65–77, Jan. 2013a.
- E. Bacry, S. Delattre, M. Hoffmann, and J. F. Muzy. Scaling limits for hawkes processes and application to financial statistics. *Stochastic Processes and Applications*, 123:2475–2499, 2013b.
- P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. 1999.
- T. Bollerslev and H. Zhou. Estimating stochastic volatility diffusion using conditional moments of integrated volatility. *Journal of Econometrics*, 109:33–65, 2002.
- T. Bollerslev and H. Zhou. Corrigendum to "Estimating stochastic volatility diffusion using conditional moments of integrated volatility". *Journal of Econometrics*, 119:221–222, 2004.
- C. G. Bowsher. Modelling security market events in continuous time: Intensity based, multivariate point process models. *Journal of Econometrics*, 141:876–912, 2007.
- P. Brémaud. *Point Processes and Queues, Martingale Dynamics*. Springer-Verlag, 1981.
- P. Brémaud and L. Massoulié. Imbedded construction of stationary sequences and point processes with a random memory. *Queueing Systems*, 17:213–234, 1994.
- R. Cont and A. De Larrard. Price Dynamics in a Markovian Limit Order Market. *Social Science Research Network Working Paper Series*, Jan. 2011.
- R. Cont and A. De Larrard. Order Book Dynamics in Liquid Markets: Limit Theorems and Diffusion Approximations. *Social Science Research Network Working Paper Series*, Feb. 2012.
- R. Cont, S. Stoikov, and R. Talreja. A stochastic model for order book dynamics. *Operations Research*, 58(3):549–563, 2010.
- C. Cuchiero, M. Keller-Ressel, and J. Teichmann. Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16:711–740, 2012.
- D. J. Daley and D. V. Jones. *An introduction to the theory of point processes, Vol 1, Elementary theory and methods*. Springer, 2. ed. edition, 2002.
- D. J. Daley and D. V. Jones. *An introduction to the theory of point processes, Vol 2, General theory and structure*. Springer, 2. ed. edition, 2008.
- D. Duffie and R. Kan. A Yield-Factor Model of Interest Rates. *Mathematical Finance*, 6(4):379–406, 1996.
- R. F. Engle and J. R. Russell. Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data. *Econometrica*, 66(5):1127–1162, Sept. 1998.
- E. Errais, K. Giesecke, and L. R. Goldberg. Affine point processes and portfolio credit risk. *SIAM J. Financial Math.*, 1(1):642–665, 2010.
- D. Filipović, E. Mayerhofer, and P. Schneider. Density approximations for multivariate affine jump-diffusion processes. *forthcoming Journal of Econometrics*, 2013.
- G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins Studies in Mathematical Sciences. The Johns Hopkins University Press, 1996.
- A. R. Hall. *Generalized Method of Moments*. Advanced Texts in Econometrics. OUP Oxford, 2004. ISBN 9780198775201.
- N. Hautsch. *Econometrics of Financial High-frequency Data*. Springer Berlin Heidelberg, 2012.

- A. G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971.
- P. Hewlett. Clustering of order arrivals, price impact and trade path optimisation. 2006.
- A. Kirilenko, R. B. Sowers, and X. Meng. A multiscale model of high-frequency trading. *Algorithmic Finance*, 2(1):59–98, Feb. 2013.
- J. Large. Measuring the resiliency of an electronic limit order book. *Journal of Financial Markets*, 10(1):1–25, February 2007.
- E. Lewis and G. Mohler. A nonparametric em algorithm for multiscale hawkes processes, 2011.
- M. Lourakis. levmar: Levenberg-marquardt nonlinear least squares algorithms in C/C++. [web page] <http://www.ics.forth.gr/~lourakis/levmar/>, Jul. 2004. [Accessed on 31 Jan. 2005].
- N. Meddahi. Arma representation of integrated and realized variances. *Econometrics Journal*, 6(2):335–356, 2003.
- S. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, 2nd edition, 2009.
- S. P. Meyn and R. L. Tweedie. Stability of Markovian Processes III: Foster-Lyapunov Criteria for Continuous-Time Processes. *Advances in Applied Probability*, 25(3), 1993.
- G. O. Mohler, M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita. Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493):100–108, 2011.
- I. Muni Toke and F. Pomponio. Modelling trades-through in a limited order book using hawkes processes. Economics Discussion Paper 2011-32, Kiel Institute for the World Economy, 2011. URL <http://www.economics-ejournal.org/economics/discussionpapers/2011-32>.
- Y. Ogata. On lewis simulation method for point processes. *IEEE Transactions on Information Theory*, 27(1):23–31, 1981.
- T. Ozaki. Maximum likelihood estimation of hawkes’ self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31(1):145–155, 1979.
- F. Pomponio and F. Abergel. Multiple-limit trades: empirical facts and application to lead-lag measures. *Quantitative Finance*, 13(5):783–793, 2013.
- D. Revuz and M. Yor. *Continuous martingales and Brownian motion*. 3rd ed, Springer-Verlag, 1999.
- A. Sarkar and R. A. Schwartz. Two-sided markets and intertemporal trade clustering: insights into trading motives. Staff Reports 246, Federal Reserve Bank of New York, 2006.
- A. Veen and F. P. Schoenberg. Estimation of spacetime branching process models in seismology using an emtype algorithm. *Journal of the American Statistical Association*, 103:614–624, June 2008.
- D. Vere-Jones. Stochastic models for earthquake occurrence. *Journal of the Royal Statistical Society Series B*, 32(1):1–62, 1970.

## Appendix

**Computing the first moment of  $N_T$  by derivation of the moment-generating function:**  
 Choosing  $u_1 = 0$  in the system of equations (11), (12) and (13) yields the moment-generating function for  $N_t$  which writes if we take  $u_2 = u$ :

$$\mathbb{E}_t^x [e^{uN_T}] = e^{a(t)+b(t)\lambda_t+uN_t}.$$

Expressed in terms of  $\tau = T - t$  the expectation and the system of ordinary differential equations are given by:

$$\mathbb{E}_t^x [e^{u(N_{t+\tau}-N_t)}] = e^{a(\tau)+\lambda_t b(\tau)} \quad (42)$$

with:

$$\frac{\partial a}{\partial \tau} = \beta \lambda_\infty b(\tau), \quad (43)$$

$$\frac{\partial b}{\partial \tau} = -\beta b(\tau) - 1 + e^{\alpha b(\tau)+u} \quad (44)$$

and  $a(0) = b(0) = 0$ .

The above system of ODE fully characterizes the moment-generating function and the Laplace transform of the process which completely determines its distribution. However, an explicit solution for the equation (44) is most often not available. From the moment-generating function (42) we can retrieve the moments of the process *after* differentiating with respect to  $u$  and evaluating for  $u = 0$  the function. This leads to differentiate with respect to  $u$  the above system. For instance, for the expected number of jumps this yields to a system of ODEs for  $b_u = \frac{db}{du}$  and  $a_u = \frac{da}{du}$  of the form:

$$\frac{\partial a_u}{\partial \tau} = \beta \lambda_\infty b_u(\tau) \quad (45)$$

$$\frac{\partial b_u}{\partial \tau} = -\beta b_u(\tau) - 1 + \alpha b_u(\tau) e^{\alpha b(\tau)+u} \quad (46)$$

with  $a_u(0) = b_u(0) = 0$  that can be explicitly solved in conjunction with equations (43) and (44) when  $u = 0$ . Therefore, we can compute the first moment. Iterating this procedure for higher moments leads to tedious computations and whenever the moments of  $\lambda_t N_t$  are needed the computations get out of hand. Needless to say that difficulties significantly increase when the dimension of the process increases.

**Proof.** of Lemma 1.

We apply Dynkin's formula (8) to  $f \equiv N_t$  and taking into account the fact that:

$$\mathcal{L}f(X_t) = \lambda_t$$

we obtain:

$$\mathbb{E}[N_t] = N_0 + \mathbb{E} \left[ \int_0^t \lambda_s ds \right].$$

Using Fubini-Tonnelli's theorem we have:

$$\mathbb{E}[N_t] = N_0 + \int_0^t \mathbb{E}[\lambda_s] ds. \quad (47)$$

Differentiating this integral equation gives (15). This equation could have been obtained by recalling that  $N_t - \int_0^t \lambda_s ds$  is a martingale, by definition of the intensity of a point process, as explained in Brémaud (1981). We nevertheless quote the Dynkin formula method as the same reasoning will prove useful for other functions as well.

To obtain the ODE (16) we rely again on Dynkin's formula. Following Errais et al. (2010), let  $f \equiv \lambda_t$  in (7) then as we have:

$$\mathcal{L}f(X_t) = \beta(\lambda_\infty - \lambda_t) + \alpha\lambda_t.$$

Dynkin's formula leads to:

$$\begin{aligned} \mathbb{E}[\lambda_t] &= \lambda_0 + \mathbb{E}\left[\int_0^t (\beta(\lambda_\infty - \lambda_s) + \alpha\lambda_s) ds\right] \\ &= \lambda_0 + \beta\lambda_\infty t + (\alpha - \beta) \int_0^t \mathbb{E}[\lambda_s] ds \end{aligned}$$

where, as before, we used Fubini-Tonnelli's theorem to swap the integration and expectation operators. Taking the differential with respect to  $t$  yields the ordinary differential equation satisfied by the expected intensity (16). ■

**Proof.** of Lemma 2.

For function  $f \equiv N^2$  the equation (7) gives:

$$\mathcal{L}f(X_t) = 2\lambda_t N_t + \lambda_t$$

and Dynkin's formula (8) results in:

$$\mathbb{E}[N_t^2] = N_0^2 + 2 \int_0^t \mathbb{E}[\lambda_u N_u] du + \int_0^t \mathbb{E}[\lambda_u] du \quad (48)$$

and differentiating this equation with respect to  $t$  leads to (17).

Following the same procedure for  $f \equiv \lambda N$  and  $f \equiv \lambda^2$  give the ODE (18) and (19), respectively. ■

**Proof.** of Proposition 3.

Taking into account the initial condition  $\mathbb{E}[\lambda_0] = \lambda_0$  the solution is found to be:

$$\mathbb{E}[\lambda_t] = \lambda_\infty \beta \frac{e^{(\alpha-\beta)t} - 1}{\alpha - \beta} + e^{(\alpha-\beta)t} \lambda_0. \quad (49)$$

From the above equation a stability condition is given by  $\frac{\alpha}{\beta} < 1$ . Using the above result in equation (47) yields the expression for the mean number of jumps:

$$\mathbb{E}[N_t] = N_0 + \frac{\lambda_\infty \beta (-1 + e^{(\alpha-\beta)t} - (\alpha - \beta)t)}{(\alpha - \beta)^2} + \frac{(-1 + e^{(\alpha-\beta)t})}{\alpha - \beta} \lambda_0.$$

We are interested in the expected number of jumps during an interval of length  $\tau$ . Using the previous computation we conclude that it is given by:

$$\mathbb{E}[N_{t+\tau} - N_t] = \frac{-\lambda_\infty \beta \tau}{\alpha - \beta} + e^{t(\alpha-\beta)} \frac{(-\lambda_\infty \beta + e^{(\alpha-\beta)\tau} \lambda_\infty \beta - \alpha \lambda_0 + e^{(\alpha-\beta)\tau} \alpha \lambda_0 + \beta \lambda_0 - e^{(\alpha-\beta)\tau} \beta \lambda_0)}{(\alpha - \beta)^2}. \quad (50)$$

The equation (50) depends on  $\lambda_0$  the initial value for the intensity which is unknown. To eliminate this value we take the limit  $t \rightarrow \infty$  and under the stability condition  $\frac{\alpha}{\beta} < 1$  we obtain (20).

We need to compute the second moment of the number of jumps during a given interval, namely:

$$I = \mathbb{E}[(N_{t_2} - N_{t_1})^2] = \mathbb{E}[\mathbb{E}_{t_1}[N_{t_2}^2] - 2N_{t_1}\mathbb{E}_{t_1}[N_{t_2}] + N_{t_1}^2]. \quad (51)$$

Using the ODE (17) it results that:

$$\mathbb{E}_{t_1}[N_{t_2}^2] = N_{t_1}^2 + 2 \int_{t_1}^{t_2} \mathbb{E}_{t_1}[\lambda_u N_u] du + \int_{t_1}^{t_2} \mathbb{E}_{t_1}[\lambda_u] du \quad (52)$$

and when inserted in the previous equation leads to:

$$I = 2 \int_{t_1}^{t_2} \mathbb{E}[\lambda_u N_u] du + \int_{t_1}^{t_2} \mathbb{E}[\lambda_u] du - 2\mathbb{E}\left[N_{t_1} \int_{t_1}^{t_2} \mathbb{E}_{t_1}[\lambda_u] du\right]. \quad (53)$$

The first integral of (53) can be computed thanks to (18) and it gives:

$$\begin{aligned} I_1 &= \int_{t_1}^{t_2} \mathbb{E}[\lambda_u N_u] du \\ &= \int_{t_1}^{t_2} e^{(\alpha-\beta)(u-t_1)} \mathbb{E}[\lambda_{t_1} N_{t_1}] du + \int_{t_1}^{t_2} \int_{t_1}^u e^{(\alpha-\beta)(u-s)} \{\beta \lambda_\infty \mathbb{E}[N_s] + \mathbb{E}[\lambda_s^2] + \alpha \mathbb{E}[\lambda_s]\} ds du \end{aligned}$$

whilst for the third term of (53) is:

$$\begin{aligned} I_2 &= \mathbb{E}\left[N_{t_1} \int_{t_1}^{t_2} \mathbb{E}_{t_1}[\lambda_u] du\right] \\ &= \mathbb{E}\left[N_{t_1} \left(\int_{t_1}^{t_2} e^{(\alpha-\beta)(u-t_1)} \lambda_{t_1} du + \int_{t_1}^{t_2} \int_{t_1}^u e^{(\alpha-\beta)(u-r)} \beta \lambda_\infty dr du\right)\right] \\ &= \int_{t_1}^{t_2} e^{(\alpha-\beta)(u-t_1)} du \mathbb{E}[N_{t_1} \lambda_{t_1}] + \int_{t_1}^{t_2} \int_{t_1}^u e^{(\alpha-\beta)(u-r)} dr du \beta \lambda_\infty \mathbb{E}[N_{t_1}]. \end{aligned}$$

As we have  $\mathbb{E}[N_s] = \mathbb{E}[N_{t_1}] + \int_{t_1}^s \mathbb{E}[\lambda_r] dr$  we arrive after substitution and simplification to:

$$I = \int_{t_1}^{t_2} \mathbb{E}[\lambda_u] du + 2 \int_{t_1}^{t_2} \int_{t_1}^u e^{(\alpha-\beta)(u-s)} \left\{ \beta \lambda_\infty \int_{t_1}^s \mathbb{E}[\lambda_r] dr + \mathbb{E}[\lambda_s^2] + \alpha \mathbb{E}[\lambda_s] \right\} ds du.$$

We can therefore calculate the time  $t$  expected second moment of the number of jumps occurring during an interval of length  $\tau$ , by first conditioning on  $\mathcal{F}_t$ , obtaining  $\mathbb{E}_t[(N_{t+\tau} - N_t)^2]$ , which is an expression depending only on the expectations  $\mathbb{E}[\lambda_t]$  and  $\mathbb{E}[\lambda_t^2]$ . These last two terms depend on  $\lambda_0$  but by letting  $t \rightarrow \infty$  we obtain its stationary regime value and get an expression independent of the



initial intensity. As a result we have the second moment of the number of jumps over a time interval of length  $\tau$ .

To specify further the result, we note  $\lim_{t \rightarrow \infty} \mathbb{E}[\lambda_t] = \Lambda$ ,  $\lim_{t \rightarrow \infty} \mathbb{E}[\lambda_t^2] = \Lambda_2$  and taking the limit on the expression for  $I$  we reach:

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{E}[(N_{t+\tau} - N_t)^2] &= \lim_{t \rightarrow \infty} \tau \Lambda + 2 \int_t^{t+\tau} \int_t^u e^{(\alpha-\beta)(u-s)} \int_t^s dr ds du \beta \lambda_\infty \Lambda \\ &\quad + 2 \int_t^{t+\tau} \int_t^u e^{(\alpha-\beta)(u-s)} ds du \{\Lambda_2 + \alpha \Lambda\} \end{aligned}$$

where the integral expressions are given by:

$$\begin{aligned} \int_t^{t+\tau} \int_t^u e^{(\alpha-\beta)(u-s)} \int_t^s dr ds du &= -(\alpha - \beta)^{-1} \frac{\tau^2}{2} - (\alpha - \beta)^{-2} \tau + (\alpha - \beta)^{-3} (e^{(\alpha-\beta)\tau} - 1) \\ \int_t^{t+\tau} \int_t^u e^{(\alpha-\beta)(u-s)} ds du &= -(\alpha - \beta)^{-1} \tau + (\alpha - \beta)^{-2} (e^{(\alpha-\beta)\tau} - 1). \end{aligned}$$

The equation (21) is deduced using the above equation and (20).

To compute the autocovariance function of the number of jumps during different time intervals we need to determine  $\mathbb{E}_t[(N_{t_1} - N_t)(N_{t_3} - N_{t_2})]$ , where  $t < t_1 < t_2 < t_3$ . In order to simplify notations we consider the variables  $\Delta_1 = t_1 - t$ ,  $\Delta_2 = t_3 - t_2$  and  $\delta = t_2 - t_1$ . By performing successive conditionings we get:

$$\mathbb{E}[(N_{t_1} - N_t)(N_{t_3} - N_{t_2})] = \mathbb{E}[\mathbb{E}_t[\mathbb{E}_{t_1}[\mathbb{E}_{t_2}[(N_{t_1} - N_t)(N_{t_3} - N_{t_2})]]]].$$

The innermost conditional expectation is:

$$\mathbb{E}_{t_2}[(N_{t_1} - N_t)(N_{t_3} - N_{t_2})] = (N_{t_1} - N_t) \times \left[ \frac{\lambda_\infty \beta (-1 + e^{(\alpha-\beta)\Delta_2} - (\alpha - \beta)\Delta_2)}{(\alpha - \beta)^2} + \frac{(-1 + e^{(\alpha-\beta)\Delta_2})}{\alpha - \beta} \lambda_{t_2} \right]$$

thanks to calculations done for the first moment. Then, conditioning down by  $\mathcal{F}_{t_1}$ , one has to compute:

$$\mathbb{E}_{t_1}[\lambda_{t_2}] = \lambda_\infty \beta \frac{e^{(\alpha-\beta)\delta} - 1}{\alpha - \beta} + e^{(\alpha-\beta)\delta} \lambda_{t_1}. \quad (54)$$

This results in an expression depending on  $N_{t_1} \lambda_{t_1}$  and  $N_t \lambda_{t_1}$ . Further conditioning down with respect to  $\mathcal{F}_t$ , one has to calculate :

$$\mathbb{E}_t[\lambda_{t_1}] = \lambda_\infty \beta \frac{e^{(\alpha-\beta)\Delta_1} - 1}{\alpha - \beta} + e^{(\alpha-\beta)\Delta_1} \lambda_t.$$

Lastly, the quantity  $\mathbb{E}_t[\lambda_{t_1} N_{t_1}]$  is already known from the previous computations. Collecting all results together we determine the autocovariance of the process and to simplify the final expression we suppose  $\Delta_1 = \Delta_2 = \tau$ , so that we obtain:

$$\lim_{t \rightarrow \infty} \mathbb{E}[(N_{t+\tau} - N_t)(N_{t+2\tau+\delta} - N_{t+\tau+\delta})] = \frac{\lambda_\infty \beta \alpha (2\beta - \alpha) (e^{(\alpha-\beta)\tau} - 1)^2}{2(\alpha - \beta)^4} e^{(\alpha-\beta)\delta} + \frac{\lambda_\infty^2 \beta^2}{(\alpha - \beta)^2} \tau^2. \quad (55)$$

If we subtract the mean value we obtain the expression (22). ■

**Expression for the skewness:** by application of the infinitesimal generator operator to adequate functions, one has the following ordinary differential equations:

$$\begin{aligned}
d\mathbb{E}[N^3] &= \mathbb{E}[\lambda_t]dt + 3\mathbb{E}[\lambda_t N_t]dt + 3\mathbb{E}[\lambda_t N_t^2]dt \\
d\mathbb{E}[\lambda_t N_t^2] &= \mathbb{E}[\lambda_t^2]dt + 2\mathbb{E}[\lambda_t^2 N_t]dt + \alpha\mathbb{E}[\lambda_t]dt + 2\alpha\mathbb{E}[\lambda_t N_t]dt + (\alpha - \beta)\mathbb{E}[\lambda_t N_t^2]dt + \beta\lambda_\infty\mathbb{E}[N_t^2]dt \\
d\mathbb{E}[\lambda_t^2 N_t] &= \mathbb{E}[\lambda_t^3]dt + 2\alpha\mathbb{E}[\lambda_t^2]dt + 2(\alpha - \beta)\mathbb{E}[\lambda_t^2 N_t]dt + \alpha^2\mathbb{E}[\lambda_t]dt + (\alpha^2 + 2\lambda_\infty\beta)\mathbb{E}[\lambda_t N_t]dt \\
d\mathbb{E}[\lambda_t^3] &= 3(\alpha - \beta)\mathbb{E}[\lambda_t^3]dt + 3(\alpha^2 + \lambda_\infty\beta)\mathbb{E}[\lambda_t^2]dt + \alpha^3\mathbb{E}[\lambda_t]dt.
\end{aligned}$$

The stationary regime third moment then writes:

$$\begin{aligned}
\lim_{t \rightarrow \infty} \mathbb{E}[(N_{t+\tau} - N_t)^3] &= \frac{1}{2(\alpha - \beta)^6} \lambda_\infty \beta \\
&\left[ -e^{2(\alpha - \beta)\tau} \alpha^2 (2\alpha - 3\beta)(\alpha - \beta) \right. \\
&\quad + 2e^{(\alpha - \beta)\tau} \alpha (\alpha^3 - 4\alpha^2\beta + 3\alpha\beta^2 + 6\beta^3 + 3(\lambda_\infty + \alpha)(\alpha - 2\beta)(\alpha - \beta)\beta\tau) \\
&\quad + \beta \left( 3\alpha (\alpha^2 - \alpha\beta - 4\beta^2) \right. \\
&\quad + 2(-\alpha + \beta) (3\lambda_\infty\alpha(\alpha - 2\beta) + \beta^2(2\alpha + \beta)) \tau \\
&\quad \left. \left. + 6\lambda_\infty(\alpha - \beta)^2\beta^2\tau^2 + 2\lambda_\infty^2\beta(-\alpha + \beta)^3\tau^3 \right) \right].
\end{aligned}$$

**Expression for the kurtosis:** similarly to the preceding paragraph, we have the following ordinary differential equations:

$$\begin{aligned}
d\mathbb{E}[N^4] &= \mathbb{E}[\lambda_t]dt + 4\mathbb{E}[\lambda_t N_t]dt + 6\mathbb{E}[\lambda_t N_t^2]dt + 4\mathbb{E}[\lambda_t N_t^3]dt \\
d\mathbb{E}[\lambda_t N_t^3] &= \mathbb{E}[\lambda_t^2]dt + 3\mathbb{E}[\lambda_t^2 N_t]dt + 3\mathbb{E}[\lambda_t^2 N_t^2]dt + \alpha\mathbb{E}[\lambda_t]dt + 3\alpha\mathbb{E}[\lambda_t N_t]dt + 3\alpha\mathbb{E}[\lambda_t N_t^2]dt \\
&\quad + (\alpha - \beta)\mathbb{E}[\lambda_t N_t^3]dt + \lambda_\infty\beta\mathbb{E}[N_t^3]dt \\
d\mathbb{E}[\lambda_t^2 N_t^2] &= \mathbb{E}[\lambda_t^3]dt + 2\mathbb{E}[\lambda_t^3 N_t]dt + 2\alpha\mathbb{E}[\lambda_t^2]dt + 4\alpha\mathbb{E}[\lambda_t^2 N_t]dt + 2\alpha\mathbb{E}[\lambda_t^2 N_t^2]dt + \alpha^2\mathbb{E}[\lambda_t]dt \\
&\quad + 2\alpha^2\mathbb{E}[\lambda_t N_t]dt + (\alpha^2 + 2\lambda_\infty\beta)\mathbb{E}[\lambda_t N_t^2]dt - 2\beta\mathbb{E}[\lambda_t^2 N_t^2]dt \\
d\mathbb{E}[\lambda_t^3 N_t] &= \mathbb{E}[\lambda_t^4]dt + 3\alpha\mathbb{E}[\lambda_t^3]dt + 3(\alpha - \beta)\mathbb{E}[\lambda_t^3 N_t]dt + 3\alpha^2\mathbb{E}[\lambda_t^2]dt + 3(\alpha^2 + \lambda_\infty\beta)\mathbb{E}[\lambda_t^2 N_t]dt \\
&\quad + \alpha^3\mathbb{E}[\lambda_t]dt + \alpha^3\mathbb{E}[\lambda_t N_t]dt \\
d\mathbb{E}[\lambda_t^4] &= (4\alpha - 4\beta)\mathbb{E}[\lambda_t^4]dt + (6\alpha^2 + 4\lambda_\infty\beta)\mathbb{E}[\lambda_t^3]dt + 4\alpha^3\mathbb{E}[\lambda_t^2]dt + \alpha^4\mathbb{E}[\lambda_t]dt.
\end{aligned}$$

The stationary regime fourth moment of the process writes:

$$\begin{aligned}
\lim_{t \rightarrow \infty} \mathbb{E} [(N_{t+\tau} - N_t)^4] &= \frac{1}{6(\alpha - \beta)^8} \times \\
&\lambda_\infty \beta \left[ -2e^{3(\alpha-\beta)\tau} \alpha^3 (3\alpha - 4\beta)(\alpha - \beta)(2\alpha - \beta) \right. \\
&\quad + 3e^{2(\alpha-\beta)\tau} \alpha^2 \left( 6\alpha^4 + 6(\lambda_\infty - 3\alpha)\alpha^2\beta + 3\alpha(-8\lambda_\infty + \alpha)\beta^2 \right. \\
&\quad \left. + 6(4\lambda_\infty + 5\alpha)\beta^3 - 18\beta^4 + 4(\lambda_\infty + 2\alpha)(2\alpha - 3\beta)(\alpha - \beta)^2\beta\tau \right) \\
&\quad - 6e^{(\alpha-\beta)\tau} \alpha \left( \alpha^5 + 6\lambda_\infty\alpha^3\beta - 3\alpha^4\beta - 24\lambda_\infty\alpha^2\beta^2 - \alpha^3\beta^2 \right. \\
&\quad + 24\lambda_\infty\alpha\beta^3 + 20\alpha^2\beta^3 - 45\alpha\beta^4 - 14\beta^5 \\
&\quad + 2(\lambda_\infty + \alpha)(\alpha - \beta)\beta (2\alpha^3 - 8\alpha^2\beta + 3\alpha\beta^2 + 18\beta^3) \tau \\
&\quad \left. + 6(\lambda_\infty + \alpha)^2(\alpha - 2\beta)(\alpha - \beta)^2\beta^2\tau^2 \right) \\
&\quad + \beta \left( \alpha (2\alpha^4 + 18\lambda_\infty\alpha(\alpha - 2\beta)^2 + 15\alpha^3\beta + 22\alpha^2\beta^2 - 216\alpha\beta^3 - 84\beta^4) \right. \\
&\quad + 6\beta(-\alpha + \beta) (6\lambda_\infty\alpha^3 + 6\alpha(-6\lambda_\infty + \alpha)\beta^2 + 8\alpha\beta^3 + \beta^4) \tau \\
&\quad \left. + 6\lambda_\infty(\alpha - \beta)^2\beta (6\lambda_\infty\alpha(\alpha - 2\beta) + \beta^2(8\alpha + 7\beta)) \tau^2 + 36\lambda_\infty^2\beta^3(-\alpha + \beta)^3\tau^3 + 6\lambda_\infty^3(\alpha - \beta)^4\beta^2\tau^4 \right) \Big].
\end{aligned}$$

## Tables

|           | MLE              |          |         | Fast Calibration I |          |         | Fast Calibration II |          |         |
|-----------|------------------|----------|---------|--------------------|----------|---------|---------------------|----------|---------|
|           | $\lambda_\infty$ | $\alpha$ | $\beta$ | $\lambda_\infty$   | $\alpha$ | $\beta$ | $\lambda_\infty$    | $\alpha$ | $\beta$ |
| mean      | 0.33             | 0.16     | 0.23    | 0.68               | -0.26    | -0.23   | 0.18                | 0.07     | 0.25    |
| Std. dev. | 5.74             | 7.13     | 6.69    | 8.39               | 10.66    | 11.04   | 4.41                | 7.47     | 7.69    |

Table 1: A Monte Carlo experiment comparing the MLE and fast calibration estimation strategies. Fast calibration I is an LS estimation based on the first, second moments and the autocorrelation function computed for several lags. Fast calibration II is a LS estimation based on the autocorrelation function computed for several lags ( $\lambda_\infty$  is then deduced using (20)). For each method and each parameters we report the mean relative error value and the corresponding standard deviation (both expressed in percent).

| <b>Symbol</b> | Measure   | Poisson- $\lambda$ | $\lambda_\infty$ | $\alpha$ | $\beta$ |
|---------------|-----------|--------------------|------------------|----------|---------|
| Eurostoxx     | Mean      | 1.4480             | 0.0625           | 0.0869   | 0.0911  |
|               | Std. dev. | 0.6283             | 0.0209           | 0.0229   | 0.0237  |
|               | Median    | 1.3343             | 0.0593           | 0.0843   | 0.0882  |
| Dax           | Mean      | 1.7814             | 0.0664           | 0.0993   | 0.1034  |
|               | Std. dev. | 0.7322             | 0.0249           | 0.0218   | 0.0226  |
|               | Median    | 1.6214             | 0.0609           | 0.0988   | 0.1028  |
| BNPP          | Mean      | 0.8627             | 0.0556           | 0.0760   | 0.0819  |
|               | Std. dev. | 0.3923             | 0.0231           | 0.0192   | 0.0219  |
|               | Median    | 0.7438             | 0.0508           | 0.0724   | 0.0772  |
| Sanofi        | Mean      | 0.6704             | 0.0453           | 0.0747   | 0.0806  |
|               | Std. dev. | 0.1873             | 0.0213           | 0.0212   | 0.0240  |
|               | Median    | 0.6087             | 0.0414           | 0.0700   | 0.0758  |

Table 2: Calibration results for two years of data. We calibrate daily a Hawkes process to the trade arrival times for each symbol. For comparison, we put the Poisson equivalent  $\lambda$ , defined as the mean number of trades per second for every day. For every measure, we put mean, standard deviation and median values for the calibrated parameters.

| <b>Symbol</b>   | Measure   | Poisson $\lambda$ | $\lambda_\infty$ | $\alpha$ | $\beta$ | $N_{response}$ |
|-----------------|-----------|-------------------|------------------|----------|---------|----------------|
| Bund            | Mean      | 1.2742            | 0.0671           | 0.0956   | 0.1013  | 19.683         |
|                 | Std. dev. | 0.4434            | 0.0244           | 0.0218   | 0.0232  | 8.3592         |
|                 | Median    | 1.2063            | 0.0645           | 0.0934   | 0.0983  | 18.256         |
| Bobl            | Mean      | 0.6699            | 0.0546           | 0.0816   | 0.0894  | 13.236         |
|                 | Std. dev. | 0.1705            | 0.0216           | 0.0225   | 0.0257  | 6.4551         |
|                 | Median    | 0.6173            | 0.0498           | 0.0790   | 0.0858  | 11.329         |
| Schatz          | Mean      | 0.6245            | 0.0473           | 0.0877   | 0.0952  | 14.896         |
|                 | Std. dev. | 0.1397            | 0.0209           | 0.0211   | 0.0236  | 7.4515         |
|                 | Median    | 0.5712            | 0.0440           | 0.0835   | 0.0903  | 13.556         |
| JPY             | Mean      | 1.6023            | 0.0536           | 0.1130   | 0.1173  | 29.737         |
|                 | Std. dev. | 0.7144            | 0.0172           | 0.0207   | 0.0215  | 11.123         |
|                 | Median    | 1.5510            | 0.0518           | 0.1133   | 0.1178  | 29.139         |
| EURO            | Mean      | 4.1955            | 0.0788           | 0.1220   | 0.1245  | 53.526         |
|                 | Std. dev. | 1.7419            | 0.0271           | 0.0192   | 0.0194  | 18.999         |
|                 | Median    | 4.1956            | 0.0770           | 0.1258   | 0.1282  | 52.045         |
| GOLD            | Mean      | 2.3191            | 0.0852           | 0.1104   | 0.1149  | 27.716         |
|                 | Std. dev. | 0.8824            | 0.0290           | 0.0237   | 0.0244  | 10.472         |
|                 | Median    | 2.1555            | 0.0815           | 0.1175   | 0.1212  | 25.743         |
| Crude Oil Brent | Mean      | 2.0453            | 0.0550           | 0.1243   | 0.1279  | 37.528         |
|                 | Std. dev. | 0.7018            | 0.0154           | 0.0143   | 0.0147  | 12.877         |
|                 | Median    | 1.9787            | 0.0535           | 0.1255   | 0.1302  | 36.86          |
| Natural GAS     | Mean      | 1.4524            | 0.0688           | 0.1177   | 0.1241  | 21.17          |
|                 | Std. dev. | 0.4362            | 0.0181           | 0.0156   | 0.0166  | 8.4653         |
|                 | Median    | 1.3653            | 0.0680           | 0.1168   | 0.1246  | 18.532         |
| Sugar           | Mean      | 0.8082            | 0.0434           | 0.1213   | 0.1289  | 19.964         |
|                 | Std. dev. | 0.3214            | 0.0196           | 0.0174   | 0.0190  | 11.93          |
|                 | Median    | 0.6869            | 0.0382           | 0.1272   | 0.1351  | 18.539         |
| CORN            | Mean      | 1.0338            | 0.0626           | 0.1069   | 0.1146  | 17.348         |
|                 | Std. dev. | 0.4563            | 0.0213           | 0.0332   | 0.0334  | 9.3997         |
|                 | Median    | 0.9563            | 0.0537           | 0.1226   | 0.1312  | 17.333         |
| WHEAT           | Mean      | 1.1562            | 0.0639           | 0.1119   | 0.1193  | 18.807         |
|                 | Std. dev. | 0.4334            | 0.0242           | 0.0215   | 0.0227  | 8.9511         |
|                 | Median    | 1.0926            | 0.0594           | 0.1182   | 0.1244  | 17.372         |

Table 3: Calibration results for two years of data (2010 and 2011) and for different asset classes (interest rates, foreign exchange, metal commodities) to assess the robustness of the model. Schatz, Bobl and Bund are respectively the 2-year, 5-year and 10-year futures on German government bonds. For each asset class we take for daily data the front maturing futures to perform the calibration. The last column  $N_{response}$  will be defined in subsection 2.3

| <b>Symbol</b> | Average $N_{response}$ |
|---------------|------------------------|
| Dax           | 26                     |
| Eurostoxx     | 22                     |
| BNPP          | 14                     |
| Sanofi        | 10                     |

Table 4:  $N_{response}$  as a characteristic of market liquidity. Symbols are ranked from the most liquid to the less liquid.

| <b>Symbol</b>   | $\lambda_\infty$ | $\alpha$ | $\beta$ | $\sigma_{BDHM}$ | Empirical $\sigma$ | Toy model $\sigma$ |
|-----------------|------------------|----------|---------|-----------------|--------------------|--------------------|
| Eurostoxx       | 0.0184           | 0.0160   | 0.0219  | 12.79%          | 40.77%             | 80.64%             |
| Dax             | 0.0429           | 0.0226   | 0.0259  | 4.80%           | 25.75%             | 71.36%             |
| BNPP            | 0.0379           | 0.0481   | 0.0569  | 5.36%           | 41.86%             | 72.24%             |
| Sanofi          | 0.0279           | 0.0488   | 0.0587  | 4.00%           | 30.30%             | 45.37%             |
| Bund            | 0.0267           | 0.0180   | 0.0261  | 2.70%           | 8.25%              | 14.47%             |
| Bobl            | 0.0228           | 0.0187   | 0.0288  | 2.45%           | 6.92%              | 11.19%             |
| Schatz          | 0.0257           | 0.0223   | 0.0372  | 1.38%           | 2.63%              | 5.39%              |
| JPY             | 0.0313           | 0.0659   | 0.0764  | 1.98%           | 9.89%              | 29.88%             |
| EURO            | 0.0474           | 0.0648   | 0.0725  | 4.87%           | 16.76%             | 112.25%            |
| GOLD            | 0.0728           | 0.0775   | 0.0868  | 3.77%           | 24.48%             | 71.64%             |
| Crude Oil Brent | 0.0474           | 0.0472   | 0.0528  | 6.22%           | 41.25%             | 126.29%            |
| Natural GAS     | 0.0548           | 0.0931   | 0.1090  | 11.24%          | 58.65%             | 150.30%            |
| Sugar           | 0.0410           | 0.0556   | 0.0758  | 10.58%          | 51.37%             | 73.30%             |
| CORN            | 0.0419           | 0.0552   | 0.0694  | 9.34%           | 43.73%             | 75.94%             |
| WHEAT           | 0.0451           | 0.0626   | 0.0763  | 10.59%          | 57.17%             | 94.53%             |

Table 5: Values for the asymptotic volatilities as given by our toy model (35) and the Bacry et al. (2013a) model (40) as well as the realized volatility of the day. We also put median estimated values for the Hawkes processes in the toy model (31), (32) and (33). Both models give plausible values.  $\sigma_{BDHM}$  volatility underestimates systematically the realized volatility, whereas our toy model systematically overestimates it.

## Figures

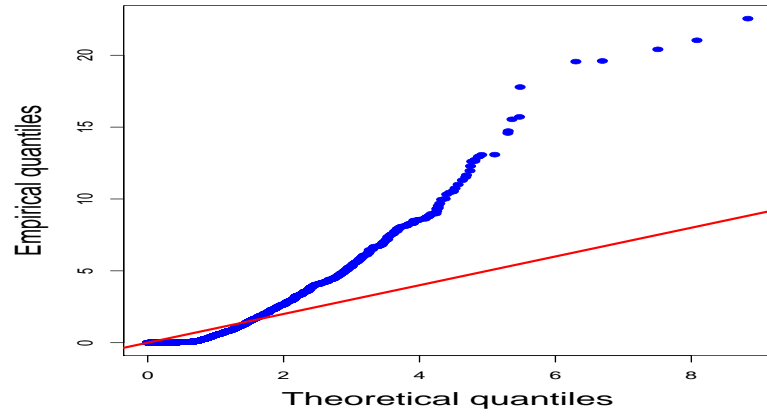


Figure 1: QQplot of inter-trade durations against exponential distribution. Interarrival trade times are clearly not exponential. Graph for Eurostoxx futures trades on 2011/03/03 for the first trading hour.

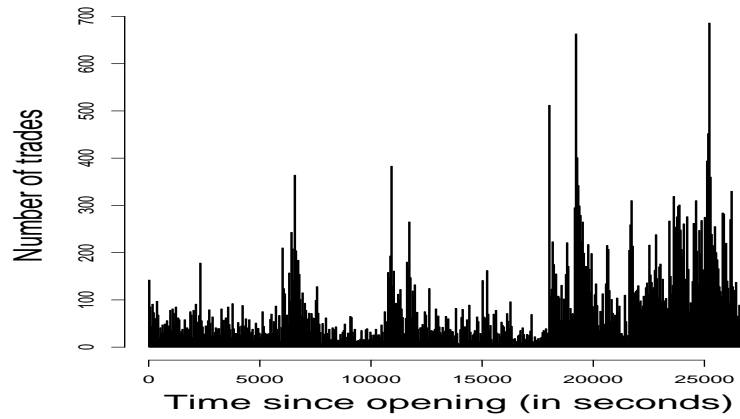


Figure 2: Number of trades by time intervals of 1 minute: presence of clusters is graphically apparent. Graph for Eurostoxx futures trades on 2011/03/03.

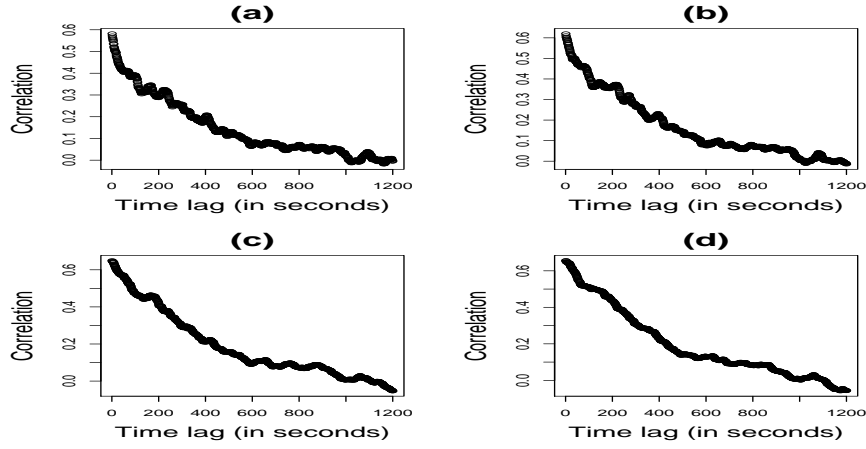


Figure 3: Empirical autocorrelation function  $\delta \rightarrow C(\tau, \delta)$ , given by (29), of the number of trades occurring on consecutive intervals of length  $\tau$  separated by a time lag  $\delta$  as a function of the lag  $\delta$ . In the four plots, we change the length of the time interval  $\tau$  considered. (a) corresponds to  $\tau = 20$  seconds, (b)  $\tau = 30$  seconds, (c)  $\tau = 60$  seconds and (d)  $\tau = 90$  seconds. The shape of the function remains identical, even if the plot is noisier when the time interval length decreases. The data are Eurostoxx futures trades on 2010/01/07.

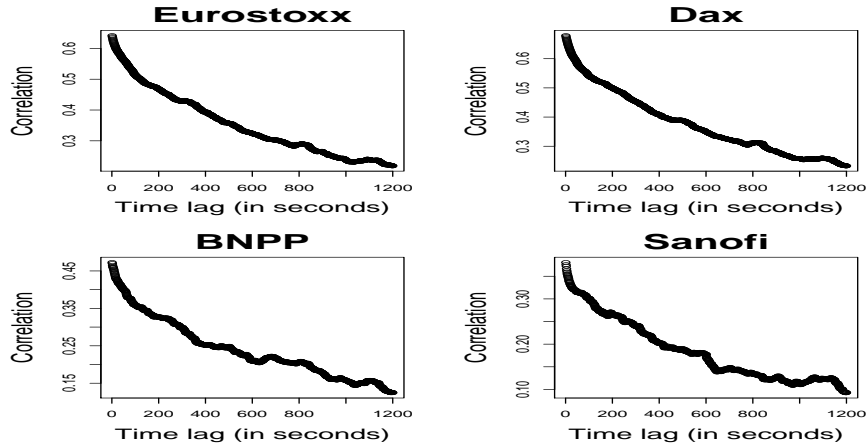


Figure 4: Empirical autocorrelation function  $\delta \rightarrow C(\tau, \delta)$  of the number of trades with  $\tau = 60$  seconds for Eurostoxx, Dax, BNPP and Sanofi, averaged every day for the month January 2010.



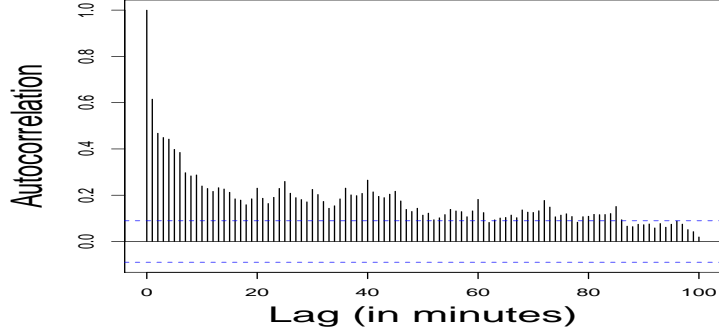


Figure 5: Number of trades empirical autocorrelation. The number of trades are computed for an interval  $\tau = 60$  seconds and the lag  $\delta$  is now measured in *minutes*. The symbol is Eurostoxx on 2010/10/26.

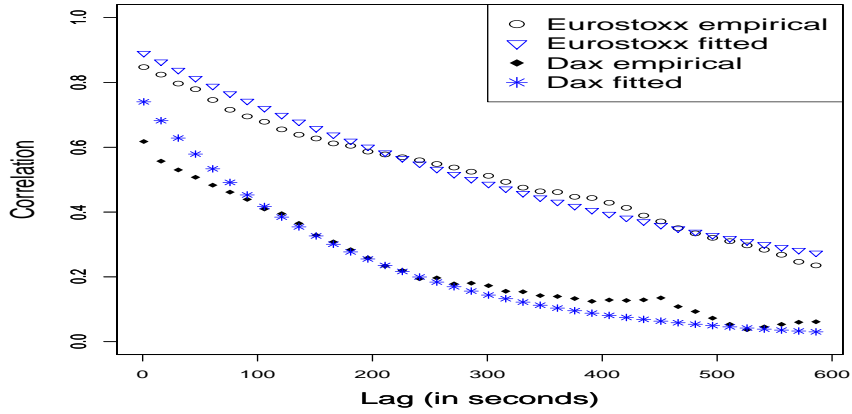


Figure 6: Empirical autocorrelation of the time series of the number of trades occurring during  $\tau = 60$  seconds versus the theoretically fitted one. For the Dax, estimated parameters are :  $\lambda_{\infty} = 0.0326806$ ,  $\alpha = 0.0431643$  and  $\beta = 0.0486235$  and for Eurostoxx, estimated parameters are :  $\lambda_{\infty} = 0.033282$ ,  $\alpha = 0.04259$  and  $\beta = 0.0446049$ .

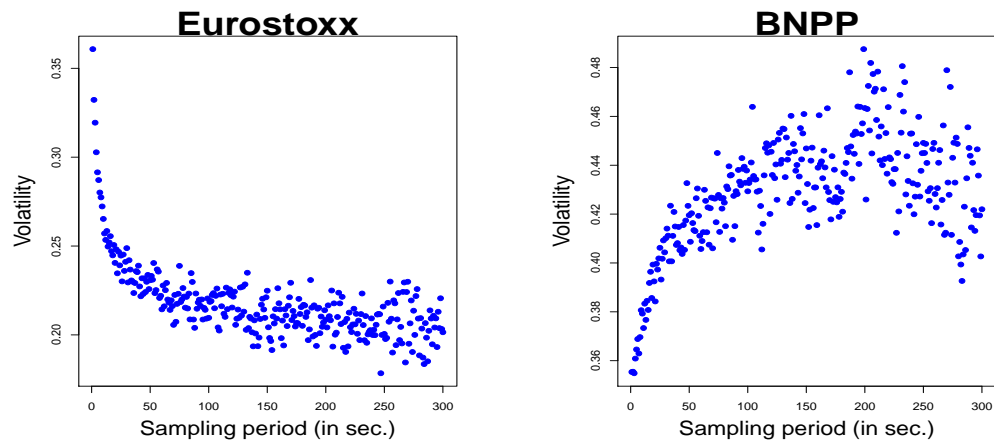


Figure 7: Signature Plot for Eurostoxx futures and BNPP stock on 2011/04/01, computed on mid prices to eliminate bid-ask bounce.