Clustering and Mean Reversion in Hawkes Microstructure Models

José Da Fonseca *   Riadh Zaatour †

August 31, 2013

Abstract

This paper provides explicit formulas for the first and second moments and the autocorrelation function of the number of jumps over a given interval for the multivariate Hawkes process. These computations are possible thanks to the affine property of this process. We unify the stock price models of Bacry et al. (2013a) and Da Fonseca and Zaatour (2013) both of them based on the Hawkes process, the first one having a mean reverting behaviour whilst the second one has a clustering behaviour, and build a model having these two properties. We compute various statistics as well as the diffusive limit for the stock price that determines the connection between the parameters driving the high frequency activity to the daily volatility. Lastly, the impulse function giving the impact on the stock price of a buy/sell trade is explicitly computed.
In a model for buy and sell orders based on the Hawkes process we compute the market impact of a trade as well as a collection of trades. We obtain a function which is similar to those obtained in previous empirical studies.

JEL Classification: C13, C32, C58.
Keywords: Hawkes process, Clustering and Mean reversion, Volatility, Diffusive limit, Market Impact.

*Corresponding author: 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

†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 both clustering and mean reverting behaviours. 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 these properties. It explains the large number of works on trading activity and more generally high-frequency econometrics based on this process as a modelling framework.

The purpose of this work is to develop closed-form solutions for some functionals of the Hawkes process. The main mathematical tool is the Dynkin formula that we use to derive the ordinary differential equations solved by the first and second moments. Thanks to these equations we compute the autocovariance function of the number of jumps over a given time interval for the Hawkes process. These computations are possible thanks to the affine property of the Hawkes process.

We use these analytical results to understand the statistical properties of a stock whose dynamics is expressed as a function of a Hawkes process. The simplest model proposed in the literature for a stock based on this process appears in Bacry et al. (2013a). Although simple this model enables to capture the mean reverting behaviour of a stock price but not the clustering behaviour. Within this framework in Da Fonseca and Zaatour (2013) a model with clustering behaviour was developed. Our objective is to unify the models of Bacry et al. (2013a) and Da Fonseca and Zaatour (2013) and build a toy model based on the multivariate Hawkes process that can exhibit both clustering and mean reverting behaviours. Thanks to the analytical results obtained for the first and second moments as well as the autocovariance function of multivariate Hawkes process the statistical properties of the price dynamic is explicitly given in terms of the model parameters. Depending on the parameters values the stock dynamic has a mean reverting or a clustering behaviour. Also, the results allow us to compute the signature plot which is the function relating the volatility computed using the stock price to the sampling frequency. Here also, the closed-form solutions permit to understand how the mean reverting and the clustering behaviours affect the computed volatility. Lastly, we compute the diffusive limit associated with the stock dynamic. Therefore, we 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).

Using a model for buy and sell orders based on the Hawkes process we compute the market impact of a trade. We then extend the results to set of orders and compute explicitly the market impact. We recover the usual shape for meta order as presented in the literature.

The structure of the paper is as follows. In the first section, we show empirical evidences of mean reverting and clustering behaviours on high frequency data. In the second 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. In the third section, we develop two applications. The first application is a toy model for the stock price based on a multivariate Hawkes process, we compute different statistical properties and the diffusive limit for the stock. The second application is a toy model for buy and sell orders and we compute explicitly the market impact of a collection of trades. Finally, we conclude and provide some technical results, tables and figures which are gathered in the appendix.

1 Empirical Evidences

We define the clustering as the strong autocorrelation in the occurrence frequency of a certain event. The mean reversion is the strong autocorrelation in the occurrence of two opposite events, such as buy and sell order arrivals or up and down price movements. Both clustering and mean reversion are important characteristics of many aspects of the trading activity. The random walk nature of the price process observed on a coarse time scale is the result of their subtle interaction on a microscopic time scale. Borrowing the language of (Bouchaud et al., 2004), we can say that clustering (*persistence*), leads to *super diffusion* whereas mean reversion (*anti-persistence*) leads to *sub-diffusion*, the overall effect being the diffusive aspect of the price process.

For example, time series of the number of buy orders occurring during consecutive time intervals of length τ are highly autocorrelated. This correlation persists for several lags as is clear in Figure 1. The same is true for sell orders. The same phenomena is observed daily and for any choice of τ.

[Insert Figure 1 here]

Additionally, if one considers time occurrences of up and down mid price jumps, the same phenomenon is observed as is clear in Figure 2.

[Insert Figure 2 here]

Conversely, cross correlation functions between the number of occurrences of opposite phenomena during a time interval also present the same property. Buy orders seem then to trigger sell orders, which in turn trigger other buy orders, as is clear in Figure 3. Up and down jumps of the mid price present the same property, and their cross correlation functions are plotted in Figure 4.

[Insert Figure 3 here]

[Insert Figure 4 here]

Therefore, the clustering effect is compensated by a mean reversion phenomenon as illustrated in the cross correlation functions. This rich interaction in the micro structure level results macroscopically in the price trajectory for which the efficient market hypothesis as well as the Brownian diffusion approximation seem to be reasonable.

These kind of figures would not have been obtained if the studied events were realizations of independent Poisson processes, and the estimated autocorrelations would have been insignificant in that case. Hawkes processes present a versatile mathematical framework allowing to deal with such phenomena.

2 Mathematical Framework

2.1 The multivariate Hawkes process

Let $X_t = (\lambda_t, N_t)$ is a Markov process in the state space $D = \mathbb{R}_+^n \times \mathbb{N}^n$, which satisfies the dynamic:

$$d\lambda_t = \beta(\lambda_\infty - \lambda_t)dt + \alpha dN_t \tag{1}$$

with β, α two $n \times n$ real matrices and λ_∞ a vector of $\mathbb{R}^n$. Ito's lemma to $e^{\beta t}\lambda_t$ yields:

$$\lambda_t = e^{-\beta t}\left(\lambda_0 - \lambda_\infty\right) + \lambda_\infty + \int_0^t e^{-\beta(t-v)}\alpha dN_v. \tag{2}$$

From (2) 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 λ_0, the initial value for the intensity, vanishes leaving us with:

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

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

$$\lambda_t = \lambda_\infty + \int_{-\infty}^t e^{-\beta(t-v)}\alpha dN_v. \tag{3}$$

The equation (3) leads to a stochastic differential equation similar to (1), 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 λ_0 in equation (2) 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 infinitesimal generator of the diffusion is given by

$$\mathcal{L}f = (\beta(\lambda_\infty - \lambda))^\top \nabla^\top f + \lambda^\top \begin{pmatrix} f(\lambda + \alpha e_1, N_t + e_1) - f \\ \ldots \\ \ldots \\ f(\lambda + \alpha e_n, N_t + e_n) - f \end{pmatrix} \tag{4}$$

for $f : D \to \mathbb{R}$, where $\nabla f = (\partial_{\lambda_1} f, ..., \partial_{\lambda_n} f)$ is a $1 \times n$ vector and $(e_i)_{i=1..n}$ is the canonical basis of $\mathbb{R}^n$.

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

$$M_t = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_v)\,dv$$

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

$$\mathbb{E}\left[f(X_s) - \int_0^s \mathcal{L}f(X_v)\,dv|\mathcal{F}_t\right] = f(X_t) + \int_t^s \mathcal{L}f(v, X_v)\,dv$$

by the martingale property, so finally:

$$\mathbb{E}\left[f(X_s)|\mathcal{F}_t\right] = f(X_t) + \mathbb{E}\left[\int_t^s \mathcal{L}f(X_v)\,dv|\mathcal{F}_t\right]. \tag{5}$$

This gives a very convenient way to calculate conditional expectations of functional of the Markov process $X_t = (N_t, \lambda_t)$ when the expectation of the right hand side of the preceding equation can be easily computed.

Notice also that further taking the expectation of the above formula allows to get rid of the conditioning and gives unconditional expectations of quantities of interest depending on the process (N_t, λ_t).

The infinitesimal generator of the diffusion leads, thanks to Feynman-Kac formula, to the computation of the moment-generating function. Denote by $\phi(t, z, u) = \mathbb{E}\left[e^{z^\top \lambda_t + u^\top N_t}\right]$ this function, it solves the partial differential equation with initial condition:

$$\begin{cases} \partial_t \phi = \mathcal{L}\phi \\ \phi(0, z, u) = e^{z^\top \lambda_0 + u^\top N_0}. \end{cases}$$

The model being affine we look for a solution of the form $e^{a_t + b_t^\top \lambda + u^\top N}$ that leads to a set of ordinary differential equations:

$$\begin{cases} \partial_t a = b^\top \beta \lambda_\infty \\ \partial_t b = -\beta^\top b + h - \mathbf{1} \end{cases}$$

with initial conditions $a_0 = 0$, $b_0 = z$, and $\mathbf{1} = (1, \ldots, 1)^\top$ whilst the function h is defined as:

$$h = \begin{pmatrix} e^{b^\top \alpha e_1 + u^\top e_1} \\ \ldots \\ e^{b^\top \alpha e_n + u^\top e_n}. \end{pmatrix}.$$

To compute the moments we still need to derive the function which leads to some tedious calculus. It is easier to compute by integrating the ordinary differential equations associated with the moments, these equations can be obtained by repeatedly applying the Dynkin formula (5) as shown in the following subsection.

2.2 Computing the moments and autocorrelation 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 τ . To achieve this we rely on the infinitesimal generator of the process given by (4) and Dynkin's formula (5). 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* (1) *then the expected number of jumps $\mathbb{E}[N_t]$ and the expected intensity $\mathbb{E}[\lambda_t]$ satisfy the set of ODE:*

$$d\mathbb{E}[\lambda_t] = \beta(\lambda_\infty - \mathbb{E}[\lambda_t])dt + \alpha\mathbb{E}[\lambda_t]dt \tag{6}$$

$$d\mathbb{E}[N_t] = \mathbb{E}[\lambda_t]dt. \tag{7}$$

These equations can be integrated explicitly as we have:

$$\mathbb{E}[\lambda_t] = (\alpha-\beta)^{-1}\left(e^{(\alpha-\beta)t} - I\right)\beta\lambda_\infty + e^{(\alpha-\beta)t}\lambda_0 \tag{8}$$

$$= c_0(t)\lambda_0 + c_1(t) \tag{9}$$

and

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

$$= N_0 + c_2(t)\lambda_0 + c_3(t). \tag{10}$$

We will need the following asymptotic expectations

Lemma 2 *Given a Hawkes process with dynamic given by* (1) *then long term expected intensity is given by:*

$$\lim_{t \to +\infty} \mathbb{E}\left[\lambda_t\right] = \bar{\lambda}_\infty = -(\alpha - \beta)^{-1} \beta \lambda_\infty. \tag{11}$$

whilst the long term expected number of jumps over an interval τ is:

$$\lim_{t \to +\infty} \mathbb{E}\left[N_{t+\tau} - N_t\right] = -(\alpha - \beta)^{-1} \beta \lambda_\infty \tau \tag{12}$$

$$= \bar{\lambda}_\infty \tau. \tag{13}$$

Lemma 3 *Given a Hawkes process $X_t = (\lambda_t, N_t)$ with dynamic given by (1) then the functions $\mathbb{E}\left[N_t N_t^\top\right]$, $\mathbb{E}[\lambda_t N_t^\top]$ and $\mathbb{E}\left[\lambda_t \lambda_t^\top\right]$ solve the set of ODE:*

$$\frac{d}{dt}\mathbb{E}\left[N_t N_t^\top\right] = \mathbb{E}\left[\lambda_t N_t^\top\right] + \mathbb{E}[N_t \lambda_t^\top] + diag(\mathbb{E}[\lambda_t]) \tag{14}$$

$$\frac{d}{dt}\mathbb{E}[\lambda_t N_t^\top] = \beta \lambda_\infty \mathbb{E}[N_t^\top] + (\alpha - \beta)\mathbb{E}[\lambda_t N_t^\top] + \mathbb{E}[\lambda_t \lambda_t^\top] + \alpha\, diag(\mathbb{E}[\lambda_t]) \tag{15}$$

$$\frac{d}{dt}\mathbb{E}\left[\lambda_t \lambda_t^\top\right] = \beta \lambda_\infty \mathbb{E}\left[\lambda_t^\top\right] + \mathbb{E}[\lambda_t]\lambda_\infty^\top \beta^\top + (\alpha - \beta)\mathbb{E}\left[\lambda_t \lambda_t^\top\right] + \mathbb{E}\left[\lambda_t \lambda_t^\top\right](\alpha - \beta)^\top + \alpha\, diag(\mathbb{E}[\lambda_t])\alpha^\top \tag{16}$$

and the long term covariance matrix for the intensity $\Lambda_\infty = \lim_{t \to +\infty} \mathbb{E}[\lambda_t \lambda_t^\top]$ solves the algebraic matrix equation

$$(\alpha - \beta)\bar{\Lambda}_\infty + \bar{\Lambda}_\infty(\alpha - \beta)^\top + \alpha\, diag(\bar{\lambda}_\infty)\alpha^\top = 0 \tag{17}$$

with $\bar{\Lambda}_\infty = \Lambda_\infty - \bar{\lambda}_\infty \bar{\lambda}_\infty^\top$ where $\bar{\lambda}_\infty$ is given by (11)*.*

Using the previous lemmas we can compute second order moment as well as the auto-covariance function. We have the following lemma

Lemma 4 *The long term second order moment of the Hawkes over a given interval $\tau > 0$ is*

$$\begin{aligned} \mathrm{Cov}(\tau) &= \lim_{t \to +\infty} \mathbb{E}\left[(N_{t+\tau} - N_t)(N_{t+\tau} - N_t)^\top\right] - \mathbb{E}\left[(N_{t+\tau} - N_t)\right]\mathbb{E}\left[(N_{t+\tau} - N_t)^\top\right] \\ &= J_1 + J_1^\top + \tau\, diag(\bar{\lambda}_\infty) \end{aligned} \tag{18}$$

with $J_1 = c_5(\tau)(\bar{\Lambda}_\infty + \alpha\, diag(\bar{\lambda}))$ and

$$c_5(\tau) = -(\alpha - \beta)^{-1}\tau + (\alpha - \beta)^{-2}(e^{(\alpha - \beta)\tau} - I)$$

As we are interested by the autocorrelation structure of the process the following quantity proved to be essential

Lemma 5 *Given $t_1 < t_2 \leq t_3 < t_4$ with $t_2 - t_1 = \tau_1$, $t_4 - t_3 = \tau_2$ and $t_3 - t_2 = \delta$ we have*

$$\begin{aligned}
\textsc{Cov}_1(\tau_1,\tau_2,\delta) &= \lim_{t_1\to+\infty} \mathbb{E}\left[(N_{t_4}-N_{t_3})(N_{t_2}-N_{t_1})^\top\right] - \mathbb{E}\left[(N_{t_4}-N_{t_3})\right]\mathbb{E}\left[(N_{t_2}-N_{t_1})^\top\right] \\
&= c_2(\tau_2)c_0(\delta)c_2(\tau_1)\left(\bar{\Lambda}_\infty + \alpha\, diag(\bar{\lambda}_\infty)\right) \qquad (19)
\end{aligned}$$

with $\bar{\Lambda}_\infty$ given by (17) *and $\bar{\lambda}_\infty$ by* (11)*.*

It is possible to relax the assumption of overlapping intervals made in the previous lemma. We have

Lemma 6 *Given $t_1 < t_3 \leq t_2 < t_4$ with $t_2 - t_1 = \tau_1$, $t_4 - t_3 = \tau_2$ and $t_3 - t_1 = \delta$ (note the difference with the previous lemma), we have*

$$\begin{aligned}
\textsc{Cov}_2(\tau_1,\tau_2,\delta) &= \lim_{t_1\to+\infty} \mathbb{E}\left[(N_{t_4}-N_{t_3})(N_{t_2}-N_{t_1})^\top\right] - \mathbb{E}\left[N_{t_4}-N_{t_3}\right]\mathbb{E}\left[(N_{t_2}-N_{t_1})^\top\right] \qquad (20) \\
&= \textsc{Cov}_1(\tau_1,\tau_2-(\tau_1-\delta),0) + \textsc{Cov}(\tau_1-\delta) + \textsc{Cov}_1(\delta,\tau_1-\delta,0).
\end{aligned}$$

Having the expression for the autocovariance functions we can define the autocorrelation function that will be used in the calibration procedure.

$$\textsc{Corr}(\tau,\delta) = \lim_{t\to+\infty} \frac{\mathbb{E}\left[(N_{t+\tau}-N_t)(N_{t+\tau+\delta}-N_{t+\delta})^\top\right] - \mathbb{E}\left[N_{t+\tau}-N_t\right]\mathbb{E}\left[(N_{t+\tau+\delta}-N_{t+\delta})^\top\right]}{\sqrt{\textsc{Cov}(\tau)}\sqrt{\textsc{Cov}(\tau+\delta)}} \qquad (21)$$

where depending whether $\delta \geq \tau$ or not we use either (19) or (20).
Lemma 5 leads to the following useful result whose proof is straightforward:

Lemma 7 *Given the covariance matrix $\textsc{Cov}_1(\tau_1,\tau_2,\delta)$ defined by* (19) *then we have*

$$\begin{aligned}
\bar{\Sigma} &= \sum_{j=0}^{\infty} \textsc{Cov}_1(1,1,j) \\
&= -(\alpha-\beta)^{-2}(I - e^{\alpha-\beta})(\bar{\Lambda}_\infty + \alpha\, diag(\bar{\lambda}_\infty)) \qquad (22)
\end{aligned}$$

2.3 Data Description and Estimation Algorithm

We rely on tick-by-tick data from TRTH (Thomson Reuters Tick History). We deal with futures on indices such as Dax and Eurostoxx, as well as some other commodity, interest rates and forex futures. The data covers the period between 2010/01/01 to 2011/12/31. It consists of quotes files recording quote changes (bid/ask prices and quantities) timestamped up to the millisecond, as well as trade files recording the transactions (prices and quantities) timestamped to the millisecond.

For some of our empirical investigations, we needed to infer the sign of the trades and to this purpose we rely on the Lee and Ready algorithm as introduced in Lee and Ready (1991).

The estimation algorithm relies on maximum likelihood method. From Proposition 7.2.III of Daley and Jones (2002), the log-likelihood of a point process $(N_t)_{t\geq 0}$ writes up to an additive constant:

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

So that, the log-likelihood of the multidimensional Hawkes process is the sum of the log-likelihoods generated by the observation of each coordinate process:

$$L = \sum_{i=1}^{n} L_m.$$

This takes a particularly simple form if we consider a diagonal structure for the β matrix, that is $\beta = diag(\bar{\beta}_1, ..., \bar{\beta}_n)$, we obtain thanks to (Ogata, 1981):

$$L_m = -\lambda_\infty^m T - \sum_{i=1}^{n} \sum_{j=1}^{N} \frac{\alpha_{mi}}{\bar{\beta}_i} \left(1 - e^{\bar{\beta}_i (T - t_j)}\right) + \sum_{i=1}^{N_T} \ln\left[\lambda_\infty^m + \sum_{j=1}^{N_T} \alpha_{mi} R^{mi}(j)\right],$$

where:

$$\begin{aligned} R^{mi}(1) &= 0 \\ R^{mm}(j) &= e^{-\bar{\beta}_m (t_j^m - t_{j-1}^m)} (1 + R^{mm}(j-1)) \\ R^{mi}(j) &= e^{-\bar{\beta}_i (t_i^m - t_{j-1}^m)} R_m^u(j-1) + \sum_{k: t_{j-1}^m \leq t_k^i < t_j^m} e^{-\bar{\beta}(t_k^m - t_j^i)} \text{ for } i \neq m \end{aligned}$$

This recursive element enables efficient calculation of the function.

3 Applications

3.1 Generalized Bacry-Delattre-Hoffmann-Muzy model

3.1.1 Model specification and its estimation

In this section, we closely follow the spirit of (Bacry et al., 2013a) and generalize their results. Our purpose is to model the evolution of the *mid* price of a traded asset. The *mid* price is the mean of the best ask price in the order book and the best bid price. As these move up or down by one tick at a time, the mid price moves up (down) by *half* a tick every time that the best ask price or best bid price moves up (down) by one tick. The model writes :

$$S_t = S_0 + \left(N_t^u - N_t^d\right)\frac{\nu}{2}, \tag{23}$$

where ν is the tick value. The N_t^u and N_t^d are Hawkes processes capturing the *up* ad *down* jumps of the mid price. We consider that both processes are self exciting as well as mutually exciting. We are then capturing both characteristics of microstructure price formation: clustering and mean reversion. In the stationary regime, their intensities write:

$$\lambda_t^u = \lambda_\infty + \int_0^t \alpha_s e^{-\bar{\beta}(t-v)} dN_v^u + \int_0^t \alpha_m e^{-\bar{\beta}(t-v)} dN_v^d \tag{24}$$

$$\lambda_t^d = \lambda_\infty + \int_0^t \alpha_m e^{-\bar{\beta}(t-v)} dN_v^u + \int_0^t \alpha_s e^{-\bar{\beta}(t-v)} dN_v^d, \tag{25}$$

where with our original notations this translates to:

$$\alpha = \begin{pmatrix} \alpha_s & \alpha_m \\ \alpha_m & \alpha_s \end{pmatrix}; \beta = \begin{pmatrix} \bar{\beta} & 0 \\ 0 & \bar{\beta} \end{pmatrix}$$

where α_s stands for the self excitation parameter, α_m stands for the mutual excitation parameter, the two intensities lead to a two-dimensional vector $\lambda_t = (\lambda_t^u, \lambda_t^d)^\top$ whilst the two jumps processes give the vector $N_t = (N_t^u, N_t^d)^\top$.

Notice that α is component-wise positive, that is to say $\alpha_s > 0$ and $\alpha_m > 0$. As a result, we exclude any inhibitory effects (that is a jump in one component of the process decreasing the probability of jump for the other component).

A stability condition for the model is that the matrix $\alpha - \beta$ has eigenvalues with negative real parts. The eigenvalues of the matrix $\alpha - \beta$ are:

$$x_1 = \alpha_s - \bar{\beta} - \alpha_m \tag{26}$$

$$x_2 = \alpha_s - \bar{\beta} + \alpha_m. \tag{27}$$

So that the stability condition translates to:

$$x_1 < 0 \Leftrightarrow \bar{\beta} - \alpha_s + \alpha_m > 0 \tag{28}$$

$$x_2 < 0 \Leftrightarrow \bar{\beta} - \alpha_s - \alpha_m > 0. \tag{29}$$

We choose to consider perfectly symmetric processes so that the resultant price to be balanced (i.e. to have a martingale property). This property will be necessary in order to derive the asymptotic volatility. Note that this asymptotic result can be obtained under more general matrices α and β so long that some constraints are satisfied. For example, if we consider a non symmetric setting we can still ensure the martingale property of the price. For instance, considering

$$\alpha = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}; \beta = \begin{pmatrix} \bar{\beta}_1 & 0 \\ 0 & \bar{\beta}_2 \end{pmatrix}$$

the formulas for the moment allow us to show that for the price to be a martingale, that is in order for the up jumps to completely offset the down jumps on average, the above variables have to satisfy:

$$\frac{\alpha_{11} + \alpha_{12}}{\bar{\beta}_1} = \frac{\alpha_{21} + \alpha_{22}}{\bar{\beta}_2}$$

We nevertheless stick with our simpler setting as the aim here is to highlight clustering and mean reversion interactions. Considering a non symmetric setting enables to uncover asymmetries in such interactions and we let that interesting problem for a future research. Lastly, to the extent that we are not interest in asymptotic results and that some constraints ensuring the stability of the SDE are satisfied the matrices can be rather general.

We rely on the previously cited quote files giving best bid and ask changes timestamped up to the millisecond. This allow us to produce time records of up and down mid price jumps corresponding to a bivariate Hawkes process N in our framework. Parameter estimation is then conducted by maximum likelihood. The likelihood function is computed thanks to Ozaki (1979) and the optimisation is conducted by a Nelder-Mead algorithm[1]. We conducted the parameter estimation daily on our 2-year dataset. After the first estimation, we use the found parameters as a first guess for the consecutive day calibration. Results are reported in Table 1, where they seem very stable.

[Insert Table 1 here]

Additionally, the model is generalized version of the pioneering work of Bacry et al. (2013a) who consider only mean reversion. It also generalizes the work in Da Fonseca and Zaatour (2013) who consider only the clustering component. It is then interesting, as the general model encompasses both of the previously cited, to see if both of the phenomena is really needed. To that purpose, we conduct a likelihood ratio test. Results are reported in Table 1 where one can see that both phenomena are generally needed, but for some days for Eurostoxx, and the German bonds (Bobl, Bund and Schatz), where mean reversion alone can be a best data fit.

3.1.2 Statistical properties

Having specified the dynamics for the intensities we can analyse the statistical properties of the asset in this toy model. Thanks to the computation carried out in the analytical section many these properties can be explicitly expressed in terms of the parameters driving the Hawkes processes.

Within this simple model we can compute the autocovariance function of the price increments. To this end we consider the expected covariance of price increments over two non-overlapping time intervals of length τ and with lag δ, it is defined as:

$$\begin{aligned} CovStock(\tau,\delta) &= \mathbb{E}\left[\left(S_{t+\tau}-S_t\right)\left(S_{t+2\tau+\delta}-S_{t+\tau+\delta}\right)\right] \\ &= \mathbb{E}\left[\left((N^u_{t+\tau}-N^u_t)-(N^d_{t+\tau}-N^d_t)\right)\left((N^u_{t+2\tau+\delta}-N^u_{t+\tau+\delta})-(N^d_{t+2\tau+\delta}-N^d_{t+\tau+\delta})\right)\right]\frac{\nu^2}{4}. \end{aligned}$$

For the model considered this quantity can be computed explicitly as we have:

Proposition 8 *The autocovariance function $CovStock(\tau,\delta)$ is given by*

$$CovStock(\tau,\delta) = \left(M_{11}+M_{22}-M_{12}-M_{21}\right)\frac{\nu^2}{4} \tag{30}$$

[1]We used the open source library NL-opt, see `http://ab-initio.mit.edu/wiki/index.php/NLopt`

where M is given by 5.

For intensities following the dynamics (24) *and* (25)*. The above equation leads to the expression for the autocorrelation function of price increments:*

$$CorrStock(\tau,\delta) = -\frac{e^{-(\delta+2\tau)(\beta+\alpha_m-\alpha_s)}\left(e^{\tau(\beta+\alpha_m-\alpha_s)}-1\right)^2(\alpha_m-\alpha_s)(2\beta+\alpha_m-\alpha_s)}{2\beta^2(\beta+\alpha_m-\alpha_s)}. \tag{31}$$

First, if $\alpha_m = 0$ then we conclude of the positiveness of the above correlation. An up jump increases the mid price and the up intensity (without affecting the down intensity) which increases the likelihood of another up jump. As a result, there is a positive autocorrelation of the stock price and a clustering of jumps in the same direction. On the contrary, if $\alpha_s = 0$ the above correlation is negative. An up jump increases the mid price and the down intensity which increases the likelihood of a down jump and a decrease of the mid price, it leads to a mean reverting behaviour of the mid price. Whenever $\alpha_s = \alpha_m$ the above autocorrelation is equal to zero because the two opposite effects offset each other. Moreover, the decay of the autocorrelation as a function of the lag increases α_m whilst α_s has the opposite effect.

Having a better understanding of the clustering and mean reverting behaviour in this model we can analyse its impact on the signature plot. The use of high-frequency data leads to the estimation of the volatility from returns sampled at possibly different frequencies. The dependency of the resulting volatility on the sampling frequency is called the signature plot and is of tremendous importance in practice. Within the toy model this effect can be explicitly analysed. Indeed, the realized variance over a period T calculated by sampling the data by time intervals of length τ can be written as:

$$\begin{aligned}
\hat{C}(\tau) &= \frac{1}{T}\sum_{n=0}^{T/\tau-1}\left(S_{(n+1)\tau}-S_{n\tau}\right)^2 = \frac{1}{T}\sum_{n=0}^{T/\tau-1}\left(\left(N^u_{(n+1)\tau}-N^u_{n\tau}\right)-\left(N^d_{(n+1)\tau}-N^d_{n\tau}\right)\right)^2\frac{\nu^2}{4}\\
&= \frac{1}{T}\sum_{n=0}^{T/\tau-1}\left(N^u_{(n+1)\tau}-N^u_{n\tau}\right)^2\frac{\nu^2}{4}+\frac{1}{T}\sum_{n=0}^{T/\tau-1}\left(N^d_{(n+1)\tau}-N^d_{n\tau}\right)^2\frac{\nu^2}{4}\\
&-2\frac{1}{T}\sum_{n=0}^{T/\tau-1}\left(N^u_{(n+1)\tau}-N^u_{n\tau}\right)\left(N^d_{(n+1)\tau}-N^d_{n\tau}\right)\frac{\nu^2}{4}.
\end{aligned}$$

The mean signature plot, or more simply signature plot, is the expectation of the above quantity and is explicitly given by:

Proposition 9 *The signature plot $C(\tau) = \mathbb{E}[\hat{C}(\tau)]$ is:*

$$C(\tau) = \frac{\nu^2}{4}\left(M_{11} + M_{22} - M_{12} - M_{21}\right) \tag{32}$$

$$= \frac{\nu^2}{2}\Lambda\left(\kappa^2 + \left(1-\kappa^2\right)\frac{\left(1-e^{-\tau\gamma}\right)}{\gamma\tau}\right) \tag{33}$$

where M in (32) *is the second moment matrix calculated in 4 whilst* (33) *is the expression when the intensities follow the dynamics* (24) *and* (25) *and the parameters are:*

$$\Lambda = \frac{\bar{\beta}\lambda_\infty}{\bar{\beta} - \alpha_s - \alpha_m}\ ,\ \kappa = \frac{\bar{\beta}}{\bar{\beta} + \alpha_m - \alpha_s}\ \text{ and } \gamma = \bar{\beta} + \alpha_m - \alpha_s.$$

Notice that $\alpha_s = 0$ enables to retrieve (Bacry et al., 2013a) results, that is to say the case with intensities which are only mutually excited. Whenever $\alpha_m = 0$ leads to the results of (Da Fonseca and Zaatour, 2013), the case of self-excitation only. Also, whenever the clustering and mean reversion effects are equal (i.e. $\alpha_m = \alpha_s$) the signature plot is flat and the volatility estimated at every time resolution is equal to $\frac{\nu^2\Lambda}{2}$, there is no bias due to mean reversion or clustering because these two effects perfectly offset each other. Lastly, taking the limit τ to infinity leads the the asymptotic volatility as it will be proven in the following propositions.

We focus on the computation of the diffusive limit associated with the stock dynamic. It allows the determination of the connection between 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.

To fulfil this objective a limit theorem is needed. In Bacry et al. (2013b), the authors rely on 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. In our case, as the kernel is exponential the process $X_t = \left(N_t^u, \lambda_t^u, N_t^d, \lambda_t^d\right)$ is a Markov process. Its infinitesimal generator writes:

$$\begin{aligned}\mathcal{L}f(x) &= \bar{\beta}\left(\lambda_\infty - \lambda_t^u\right)\frac{\partial f}{\partial \lambda^u}(x) + \bar{\beta}\left(\lambda_\infty - \lambda_t^d\right)\frac{\partial f}{\partial \lambda^d}(x) \\ &+ \lambda_t^u\Big[f\left(N_t^u + 1, \lambda_t^u + \alpha_s, N_t^d, \lambda_t^d + \alpha_m\right) - f(x)\Big] \\ &+ \lambda_t^d\Big[f\left(N_t^u, \lambda_t^u + \alpha_m, N_t^d + 1, \lambda_t^d + \alpha_s\right) - f(x)\Big].\end{aligned}$$

Thanks to the explicit expression for the infinitesimal generator we rely on a simple method to establish some stability results. For instance, ergodicity of the process X_t, that is its convergence to a stationary regime can be easily established thanks to the Foster-Lyapounov test function criterion. In our case, define the function $V(x) = \frac{\lambda^u + \lambda^d}{2\lambda_\infty}$, then a simple calculation yields the *geometric drift condition*:

$$\mathcal{L}V(x) \leq -(\bar{\beta} - \alpha_m - \alpha_s)V(x) + \bar{\beta}, \tag{34}$$

which grants, thanks to Theorems 6.1 and 7.1 in (Meyn and Tweedie, 2009) (and especially (CD3)), the V-uniform ergodicity of the process X_t.

Let us then write unit-time price increments:

$$\eta_i = \Big[(N_i^u - N_{i-1}^u) - (N_i^d - N_{i-1}^d)\Big] \times \frac{\nu}{2},$$

and consider the random sums

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

where $\{\eta_i; i = 1 \ldots n\}$ denotes a set of price increments (note that $\mathbb{E}[\eta_i] = 0$). We are interested in the asymptotic behaviour of the price process

$$\bar{S_t^n} = \frac{S_{\lfloor nt \rfloor}}{\sqrt{n}}.$$

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.

$$\bar{S_t^n} \Rightarrow \sigma W_t.$$

with the volatility is given by

$$\sigma^2 = \lim_{n \to \infty} \frac{\mathrm{Var}(S_n)}{n}.$$

The closed-form formulas obtained in the analytical part enable to compute explicitly this volatility as we have the following proposition:

Proposition 10 *The volatility σ^2 is given:*

$$\sigma^2 = \frac{\nu^2}{4}\left(M_{11} + M_{11} - M_{12} - M_{21}\right) \tag{35}$$

with

$$M = \mathrm{Cov}(1) + 2\bar{\Sigma} \tag{36}$$

where the two matrices in (36) *are given by* (18) *and* (22)*, respectively.*

This expression is valid for a general dynamic but for the particular choice made in this work the volatility turns out to have a very simple expression, as the next proposition shows:

Proposition 11 *If the dynamic intensities are given by* (24) (25) *the volatility* (35) *has the expression:*

$$\begin{aligned}\sigma^2 &= \frac{\nu^2}{2}\Lambda\kappa^2\\ &= \frac{\nu^2}{2}\frac{\bar{\beta}^3\lambda_\infty}{(\bar{\beta} - \alpha_s - \alpha_m)(\bar{\beta} + \alpha_m - \alpha_s)^2}\end{aligned}$$

with:

$$\Lambda = \frac{\bar{\beta}\lambda_\infty}{\bar{\beta} - \alpha_s - \alpha_m}, \quad \kappa = \frac{\bar{\beta}}{\bar{\beta} + \alpha_m - \alpha_s}.$$

As can be expected, increasing the self excitation parameter α_s, and hence the clustering effect, increases the volatility as this leads to positive autocorrelation of the returns. On the contrary, increasing the mutual excitation parameter α_m decreases the overall volatility because of the negative autocorrelation of the returns. Obviously, these parameters have to remain in the region $\bar{\beta} - \alpha_s - \alpha_m > 0$ in order for the stability condition to be satisfied. Also, taking the limit $\tau \to +\infty$ in (33) leads to an equality between the signature plot and the asymptotic volatility.

In Bacry et al. (2013a) the authors consider a mean reversion effect only, and obtain the same formula as above with $\alpha_s = 0$, whereas in Da Fonseca and Zaatour (2013) the authors consider clustering alone and obtain the same formula as above with $\alpha_m = 0$. In the latter article, calibration is done on real data of both models (clustering only or mean reversion only), and empirical results show that considering only one effect results in a systematically underestimated or systematically overestimated realized volatility.

We conducted here the same realized variance estimations as in Da Fonseca and Zaatour (2013), based on the calibration of this more complete model. Results are reported in Table 2. They clearly show that calibrating the clustering as well as mean reversion effects results in better fits to realized volatility.

[Insert Table 2 here]

Within this simple model it is possible to explicitly analyse the price impact of a trade. This is a key quantity that appears in many, if not all, papers on microstructure models, see for example Dufour and Engle (2000). We illustrate the impact of a "up" move but similar results apply to a down move.

Proposition 12 *Given τ and δ the impact of a "up" move on the stock returns is*

$$\begin{aligned}
\textsc{imp}(\tau,\delta) &= \lim_{t\to+\infty} \mathbb{E}[S_{t+\tau+\delta} - S_{t+\delta} | dN_t^u = 1] && (37)\\
&= \frac{\nu}{2\bar{\lambda}_\infty^u}(M_{11} - M_{12}) && (38)\\
&= -\nu \frac{e^{-\tau(\beta+\alpha_m-\alpha_s)} e^{-\delta(\beta+\alpha_m-\alpha_s)} \left(e^{\tau(\beta+\alpha_m-\alpha_s)} - 1\right)(\alpha_m - \alpha_s)(2\beta + \alpha_m - \alpha_s)}{4\,(\beta+\alpha_m-\alpha_s)^2} && (39)
\end{aligned}$$

with $M = c_2(\tau)c_0(\delta)(\bar{\Lambda}_\infty + \alpha\, diag(\bar{\lambda}_\infty))$.

It might be convenient to consider the stock evolution over an infinitesimal interval. From the previous proposition it is straightforward to deduce

Proposition 13 *Given the price impact function 40 then*

$$\begin{aligned}
\textsc{imp}(\delta) &= \lim_{\tau\to 0} \frac{\textsc{imp}(\tau,\delta)}{\tau} && (40)\\
&= \frac{\nu}{2\bar{\lambda}_\infty^u}(M_{11} - M_{12}) && (41)\\
&= -\nu \frac{e^{-\delta(\beta+\alpha_m-\alpha_s)}(\alpha_m - \alpha_s)(2\beta + \alpha_m - \alpha_s)}{4\,(\beta+\alpha_m-\alpha_s)} && (42)
\end{aligned}$$

with $M = c_0(\delta)(\bar{\Lambda}_\infty + \alpha\, diag(\bar{\lambda}_\infty))$.

It is often meaningful to consider the *cumulative* impact of trade on the stock return, it leads to integrate with respect to δ the above quantity. Simple computations give

Proposition 14 *Given the function* $\text{IMP}(\delta)$ *of proposition 13 then the cumulative price impact of trades up to a given time t is*

$$\text{CIMP}(t) = \int_0^t \text{IMP}(\delta)d\delta \tag{43}$$
$$= \frac{\nu}{2\bar{\lambda}_\infty^u}(M_{11}(t) - M_{12}(t)) \tag{44}$$
$$= -\nu\frac{\left(1 - e^{-t(\beta+\alpha_m-\alpha_s)}\right)(\alpha_m - \alpha_s)(2\beta + \alpha_m - \alpha_s)}{4(\beta + \alpha_m - \alpha_s)^2} \tag{45}$$

with $M(t) = \int_0^t c_0(\delta)d\delta(\bar{\Lambda}_\infty + \alpha diag(\bar{\lambda}_\infty))$.

Note that if $\alpha_s > \alpha_m$ then the function is increasing and concave and in that case the shape is similar to the one obtain in Figure 1 of Dufour and Engle (2000).

3.2 Market Impact in a Buy-Sell Toy Model

As stated before, another aspect of the trading activity presenting both clustering and mean reversion effects is the sell and buy order arrivals. Time arrivals of buy and sell orders can then be modelled by a bivariate Hawkes process. Let N_t^{Buy} and N_t^{Sell} be the cumulated number of Buy and sell orders respectively, and let their respective intensities obey:

$$\lambda_t^{Buy} = \lambda_\infty + \int_0^t \alpha_s e^{-\bar{\beta}(t-u)}dN_u^{Buy} + \int_0^t \alpha_m e^{-\bar{\beta}(t-u)}dN_u^{Sell}$$
$$\lambda_t^{Sell} = \lambda_\infty + \int_0^t \alpha_m e^{-\bar{\beta}(t-u)}dN_u^{Buy} + \int_0^t \alpha_s e^{-\bar{\beta}(t-u)}dN_u^{Sell}.$$

The model is mathematically the same as before in the sense the the matrices α and β in (1) have very particular forms. Namely, they are given by:

$$\alpha = \begin{pmatrix} \alpha_s & \alpha_m \\ \alpha_m & \alpha_s \end{pmatrix}; \beta = \begin{pmatrix} \bar{\beta} & 0 \\ 0 & \bar{\beta} \end{pmatrix}$$

where α_s stands for the self excitation parameter, α_m stands for the mutual excitation parameter. The two intensities lead to a two-dimensional vector $\lambda_t = (\lambda_t^{Buy}, \lambda_t^{Sell})^\top$ whilst the two jumps processes give the vector $N_t = (N_t^{Buy}, N_t^{Sell})^\top$.

Although the model is similar to the previous one we do not focus the same phenomena. We work are at a more *microscopic* level than the mid price up and down jumps of the previous subsection because we are now interested in market orders (that might not lead to a mid-price move).

Notice that we choose symmetric processes in order for the price buy and sell pressures to be equal. As stated previously, this balance can be achieved even with some asymmetry in the parameters but we prefer to remain in the simplest possible setting.

In order to estimate the parameters of our model with real data, we need to preprocess TRTH files to make records of Buy and Sell transaction time arrivals. In fact, transaction files of TRTH do not contain any flag specifying the sign of the transaction. We then take trade and quotes files of our data as well as transaction file and apply Lee and Ready algorithm (Lee and Ready, 1991) to sign trades, obtaining records of buy and sell order arrival times corresponding to our bivariate Hawkes process $N_t = (N_t^{Buy}, N_t^{Sell})$. Calibration is done by MLE as before. We conduct the calibration daily on a 2-year data sample, where we have the futures on Eurostoxx and Dax, and two stocks: BNPP and Sanofi.

Results are reported in Table 3, where they seem to be stable and where likelihood ratio tests clearly reject the hypothesis of only clustering or only mean reversion.

[Insert Table 3 here]

As the primary interest of order book events modelling is the price dynamics, let us consider a model of price formation based on buy and sell order arrival times. To remain simple, we suppose that every trade has a unit quantity. Following (Kyle, 1985), the long run price return is assumed to depend linearly on the imbalance of buy and sell trades. More precisely, denoting by $F_t = N_t^{Buy} - N_t^{Sell}$ the cumulated buy and sell orders imbalance at time t, the price at t writes:

$$\mathbb{E}[P_\infty] = P_0 + \lambda_{\text{K}} \mathbb{E}[F_\infty],$$

where we noted λ_{K} the coefficient of linear dependence of price return on buy-sell order imbalance, this coefficient is known as the *Kyle's Lambda*. In (Hewlett, 2006), the author uses the same reasoning in the case of two independent univariate Hawkes processes, and obtain a simple market impact function.

In our case, we have:

$$\begin{aligned} P_t &= \mathbb{E}_t[P_\infty] \\ &= P_0 + \lambda_{\mathrm{K}} \mathbb{E}_t[F_\infty]. \end{aligned}$$

Moreover, following (Bouchaud et al., 2004), we consider that the price at time t results from the superposition of market impacts of all the trades that took place prior to t. The market impact of a trade is attenuated as time passes. Formally, we have:

$$P_t = P_0 + \int_0^t I(t-u) dF_u. \tag{46}$$

Equating with these two equalities we obtain:

$$\lambda_{\mathrm{K}} F_t + \lambda_{\mathrm{K}} \mathbb{E}_t[F_\infty - F_t] = \int_0^t I(t-u) dF_u. \tag{47}$$

Additionally, we have:

$$\mathbb{E}_t[F_\infty - F_t] = \int_t^\infty \mathbb{E}_t[\lambda_u^{Buy} - \lambda_u^{Sell}] du.$$

From (9) the expected intensity is equal to:

$$\mathbb{E}_t[\lambda_u] = (\alpha - \beta)^{-1} (e^{(\alpha-\beta)(u-t)} - I) \beta \lambda_\infty + e^{(\alpha-\beta)(u-t)} \lambda_t.$$

where λ_t is the vector of buy and sell intensities (i.e. $\lambda_t = (\lambda_t^{Buy}, \lambda_t^{Sell})^\top$).

Using this equality as well as the closed form formulas for the matrix exponential and matrix inverse for the model considered we arrive to:

$$\begin{aligned} \mathbb{E}_t[F_\infty - F_t] &= \int_t^\infty e^{-(\bar{\beta} + \alpha_m - \alpha_s)(u-t)} (\lambda_t^{Buy} - \lambda_t^{Sell}) du \\ &= \frac{\lambda_t^{Buy} - \lambda_t^{Sell}}{\bar{\beta} + \alpha_m - \alpha_s}. \end{aligned}$$

Then finally the impact writes:

$$\begin{aligned} \int_0^t I(t-u) dF_u &= \lambda_{\mathrm{K}} F_t + \lambda_{\mathrm{K}} \mathbb{E}_t[F_\infty - F_t] \\ &= \lambda_{\mathrm{K}} F_t + \lambda_{\mathrm{K}} \frac{\lambda_t^{Buy} - \lambda_t^{Sell}}{\bar{\beta} + \alpha_m - \alpha_s} \\ &= \lambda_{\mathrm{K}} (N_t^{Buy} - N_t^{Sell}) + \lambda_{\mathrm{K}} \frac{\lambda_t^{Buy} - \lambda_t^{Sell}}{\bar{\beta} + \alpha_m - \alpha_s}. \end{aligned}$$

Then, knowing that:

$$\lambda_t^{Buy} = \lambda_\infty + \int_0^t \alpha_s e^{-\bar{\beta}(t-u)} dN_u^{Buy} + \int_0^t \alpha_m e^{-\bar{\beta}(t-u)} dN_u^{Sell}$$

and

$$N_t^{Buy} = \int_0^t dN_u^{Buy}.$$

Similar computations for the sell-side give:

$$\begin{aligned}\int_0^t I(t-u) dF_u &= \lambda_{\mathrm{K}}(\int_0^t 1 + \frac{(\alpha_s - \alpha_m)e^{-\bar{\beta}(t-u)}}{\bar{\beta} + \alpha_m - \alpha_s}(dN_u^{Buy} - dN_u^{Sell})) \\ &= \lambda_{\mathrm{K}}(\int_0^t 1 + \frac{(\alpha_s - \alpha_m)e^{-\bar{\beta}(t-u)}}{\bar{\beta} + \alpha_m - \alpha_s} dF_u).\end{aligned}$$

So that, the impact of a single trade executed at time 0 as seen at time t is:

$$I(t) = \lambda_{\mathrm{K}} \left(1 + \frac{(\alpha_s - \alpha_m)e^{-\bar{\beta}t}}{\bar{\beta} + \alpha_m - \alpha_s}\right), \tag{48}$$

where one can see that this impact decomposes in a permanent impact and a time attenuating one. This result enables the study the impact of a large order splitting. We have the following result whose proof is omitted:

Proposition 15 *Denote by $\mathcal{I}(T)$ the price impact at time T of a continuum of orders executed over the interval $[0;t]$ then this function is given by:*

$$\mathcal{I}(T) = \int_0^t I(T-s) ds \tag{49}$$

where the function I is (48) and for the model considered the integral leads to:

$$\mathcal{I}(T) = \begin{cases} \lambda_{\mathrm{K}} \left(T + \frac{(\alpha_s - \alpha_m)}{(\bar{\beta} + \alpha_m - \alpha_s)\bar{\beta}}(1 - e^{-\bar{\beta}T})\right) & T \le t \\ \lambda_{\mathrm{K}} \left(t + \frac{(\alpha_s - \alpha_m)}{(\bar{\beta} + \alpha_m - \alpha_s)\bar{\beta}}(e^{-\bar{\beta}(T-t)} - e^{-\bar{\beta}T})\right) & T \ge t. \end{cases}$$

We illustrate graphically this function, Figure 5 presents a plot of the market impact of an order to buy 50 shares of stock, executed during 50 seconds at the rate of 1 stock per second. One clearly sees that the impact increased during the execution time of the order, and then decreases steadily to attain the permanent impact.

[Insert Figure 5 here]

In the previous computation we suppose that one unit was bought, this hypothesis can be relaxed by introducing the notion of *execution profile* as defined for example in (Moro et al., 2009).

We define the *execution profile* as a function of time f with suport on the invertal $[0;t]$ with $\int_0^t f(s)ds = Q$ which is the total quantity to be executed. With this definition, the impact of the meta order writes in integral notation:

$$\mathcal{I}(T) = \int_0^T I(T-s)f(s)ds \tag{50}$$

which is just the convolution of the execution profile and market impact.

4 Conclusion

In this paper we compute explicitly the first and second moments and the autocovariance function of the number of jumps over a given time interval of a multivariate Hawkes process. These computations are possible thanks to the affine property of the process and the use of the Dynkin formula.

Using these quantities we compute several statistical properties for a stock dynamic based on the Hawkes process. It allows us to unify the models Bacry et al. (2013a) and Da Fonseca and Zaatour (2013). The first one possesses a mean reverting behaviour whilst the second one has a clustering behaviour. We compute explicitly the the signature plot and analyse the impact of the parameters on the shape of this function. Furthermore, we compute the diffusive limit which enables to connect the parameters driving the price at high-frequency with the parameter driving the stock price a low frequency (i.e. the daily Black-Scholes volatility). Lastly, we compute explicitly the impulse response function which quantifies the impact of a trade of the stock price. For all these results we can analyse the impact of the mean reverting and clustering parameters.

We further exploit the results by analysing the price impact of a trade in a Buy and Sell toy model based on the Hawkes process. We provide an explicit expression and extend the results by computing the impact of a collection of trades. It allows us to recover the typical pattern of execution of meta order.

Although some of the formulas are derived for a specific choice of the matrix parameters most of the results are valid for general matrices (to the extent that some stability conditions are satisfied). Within this framework it is possible to consider other important market microstructure problems. As an example, the lead-lag relationship between two (or more) stocks can be analysed within that framework. The interaction between different order types (cancellation, amend, etc..) can be easily quantified and the impact on the daily volatility can be evaluated. Lastly, extending these models by adding a process for the volume might be feasible but challenging. These interesting problems are left for future investigation.

Tables

Symbol	Measure	λ_∞	α_s	α_m	β	Clustering	Mean Reversion
Eurostoxx	Mean	0.0175	0.0062	0.0698	0.0958	1364.29	-9.12
	Std. dev.	0.0065	0.0082	0.0242	0.0252	415.88	93.49
Dax	Mean	0.0661	0.0347	0.0500	0.1019	673.07	404.00
	Std. dev.	0.0286	0.0102	0.0158	0.0221	319.35	299.83
BNPP	Mean	0.0495	0.2685	0.1792	0.4660	3394.05	12359.29
	Std. dev.	0.0220	0.1241	0.0841	0.1919	2612.05	13150.40
Sanofi	Mean	0.0345	0.1941	0.1572	0.3691	2969.66	7086.07
	Std. dev.	0.0188	0.1182	0.0902	0.1968	2123.99	6553.23
Bund	Mean	0.0250	0.0189	0.0792	0.1180	1179.34	68.45
	Std. dev.	0.0123	0.0115	0.0226	0.0275	245.24	82.22
Bobl	Mean	0.0172	0.0138	0.0673	0.1009	1259.34	5.82
	Std. dev.	0.0073	0.0108	0.0226	0.0215	277.01	81.47
Schatz	Mean	0.0153	0.0196	0.0662	0.1056	1480.55	27.34
	Std. dev.	0.0047	0.0124	0.0210	0.0224	385.31	79.29
JPY	Mean	0.0449	0.2516	0.2420	0.5089	4410.78	4780.38
	Std. dev.	0.0174	0.1029	0.0941	0.1800	2085.32	2892.44
EURO	Mean	0.0622	0.3733	0.3550	0.7360	8956.69	10563.86
	Std. dev.	0.0240	0.1255	0.1145	0.1922	4041.59	5433.36
GOLD	Mean	0.0902	0.4080	0.3109	0.7300	3486.06	9290.28
	Std. dev.	0.0266	0.1481	0.1248	0.2288	2418.24	7088.23
Crude Oil Brent	Mean	0.0865	0.4506	0.2411	0.7068	7224.19	13651.24
	Std. dev.	0.0265	0.1407	0.0755	0.1838	5580.90	8525.66
Natural GAS	Mean	0.0614	0.3503	0.3044	0.6683	3766.40	6345.64
	Std. dev.	0.0194	0.1120	0.1121	0.1943	1563.27	2623.62
Sugar	Mean	0.0373	0.1998	0.1766	0.3932	1440.07	1868.66
	Std. dev.	0.0118	0.0877	0.0861	0.1625	1353.06	1546.63
CORN	Mean	0.0337	0.1879	0.1803	0.3861	1521.93	1704.87
	Std. dev.	0.0130	0.0609	0.0618	0.1166	762.05	781.92
WHEAT	Mean	0.0473	0.2370	0.2104	0.4659	1698.37	2221.75
	Std. dev.	0.0168	0.0952	0.0843	0.1644	1220.72	1486.22

Table 1: Calibration results for two years of data. We calibrate daily a bivariate Hawkes process to the up and down mid price jump times for each symbol. The column Clustering presents the likelihood ratio statistic when we consider only clustering. The column Mean reversion presents the likelihood ratio statistic when we consider only mean reversion.

Symbol	λ_∞	α_s	α_m	β	Empirical σ	Toy model σ
Eurostoxx	0.0119	0.0474	0.0401	0.1074	24.64%	19.08%
Dax	0.0581	0.0341	0.0491	0.1050	25.75%	23.45%
BNPP	0.0739	0.5287	0.2311	0.7798	56.43%	55.31%
Sanofi	0.0658	0.3503	0.3543	0.7246	31.42%	31.47%
Bund	0.0136	0.0471	0.0410	0.1081	5.35%	4.62%
Bobl	0.0125	0.0465	0.0396	0.1062	4.45%	4.44%
Schatz	0.0108	0.0360	0.0484	0.1044	2.01%	1.86%
JPY	0.0520	0.3320	0.3295	0.6814	10.74%	10.67%
EURO	0.0169	0.1007	0.1017	0.2224	5.49%	5.65%
GOLD	0.0677	0.3168	0.3561	0.6929	14.28%	14.35%
Crude Oil Brent	0.0359	0.1394	0.1394	0.2989	14.04%	14.09%
Natural GAS	0.0495	0.2752	0.2801	0.5744	48.84%	48.70%
Sugar	0.0315	0.1702	0.1642	0.3544	61.24%	61.21%
CORN	0.0591	0.3160	0.2855	0.6215	58.16%	58.08%
WHEAT	0.0642	0.2819	0.2884	0.5903	59.49%	59.46%

Table 2: Median values of asymptotic volatilities as calculated by our toy model, compared to the realized volatilities of the day. We also put median values of Hawkes model parameters.

Symbol	Measure	λ_∞	α_s	α_m	β	Clustering	Mean Reversion
Eurostoxx	Mean	0.0907	0.5674	0.1395	0.7233	3367.51	52672.29
	Std. dev.	0.0229	0.1433	0.0598	0.1738	2047.06	22999.03
Dax	Mean	0.0938	0.5929	0.1588	0.7663	3496.19	47232.60
	Std. dev.	0.0235	0.1408	0.0613	0.1708	3054.18	20459.51
BNPP	Mean	0.0427	0.3383	0.0707	0.4263	2006.10	34381.42
	Std. dev.	0.0172	0.1358	0.0456	0.1672	1826.07	19204.96
Sanofi	Mean	0.0264	0.2063	0.0484	0.2738	1318.24	19197.74
	Std. dev.	0.0096	0.0814	0.0368	0.1099	1133.54	8615.12

Table 3: Calibration results for two years of data. We calibrate daily a bivariate Hawkes process to the time arrivals of buy and sell orders. The column Clustering presents the likelihood ratio statistic when we consider only clustering. The column Mean reversion presents the likelihood ratio statistic when we consider only mean reversion.

Figures

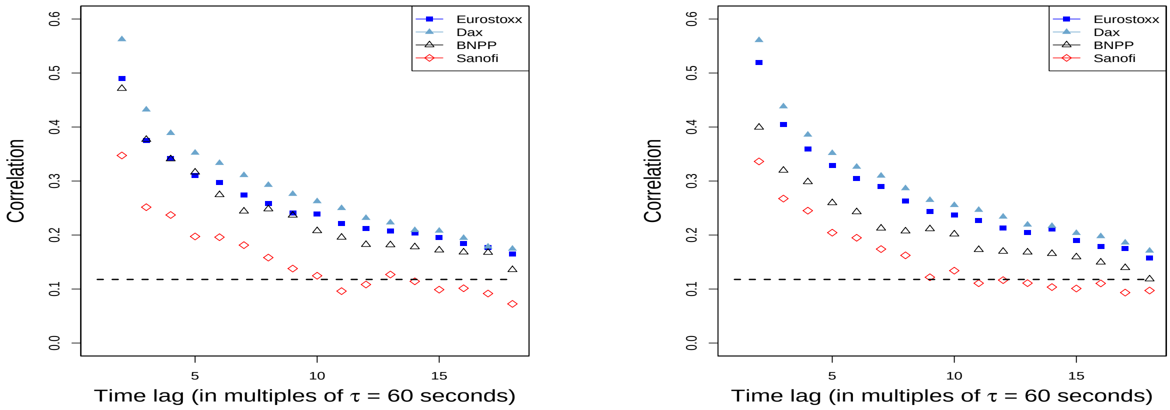

Figure 1: Autocorrelation of the number of buy trades (left) and sell trades (right) occurring on consecutive, non overlapping time interval of length τ=1 minute. We take average values of September 2011. The same plots can be observed for other values of τ. The dashed line presents the significance threshold of the estimated correlations at a 99% confidence level.

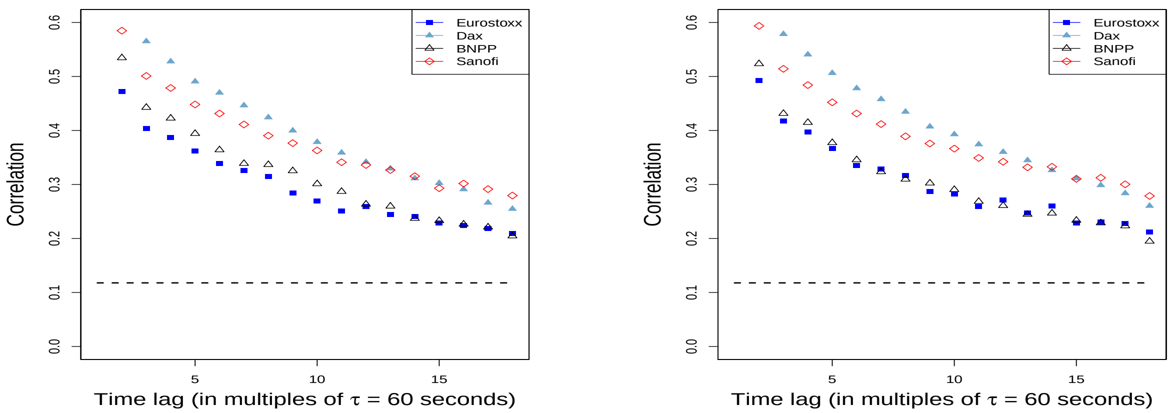

Figure 2: Autocorrelation of the up (left) and down (right) mid price jumps occurring on consecutive, non overlapping time interval of length τ=1 minute. We take average values of September 2011. The same plots can be observed for other values of τ. The dashed line presents the significance threshold of the estimated correlations at a 99% confidence level.

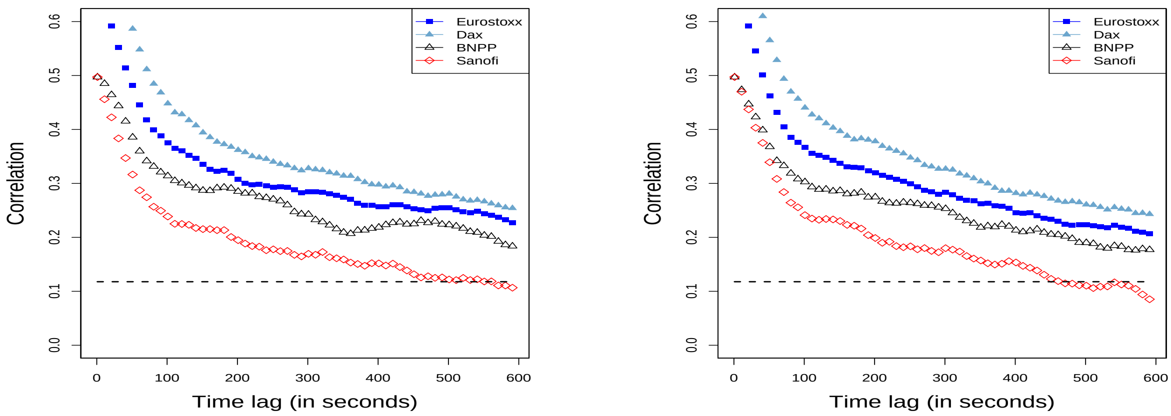

Figure 3: Cross correlation of the number of buy and sell trades, where in the left figure sell trades are lagged, whereas in the right figure buy trades are lagged. Lags are measured in seconds and the intervals can be overlapping. Time interval length is τ=1 minute. We take average values of September 2011.

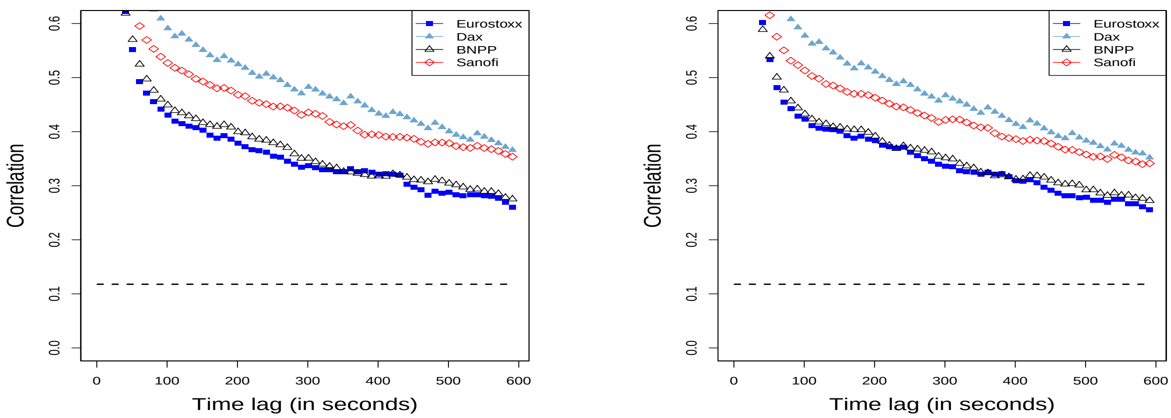

Figure 4: Cross correlation of up and down jumps of the mid price, where in the left figure down jumps are lagged, whereas in the right figure up jumps are lagged. Lags are measured in seconds and the intervals can be overlapping. Time interval length is τ=1 minute. We take average values of September 2011.

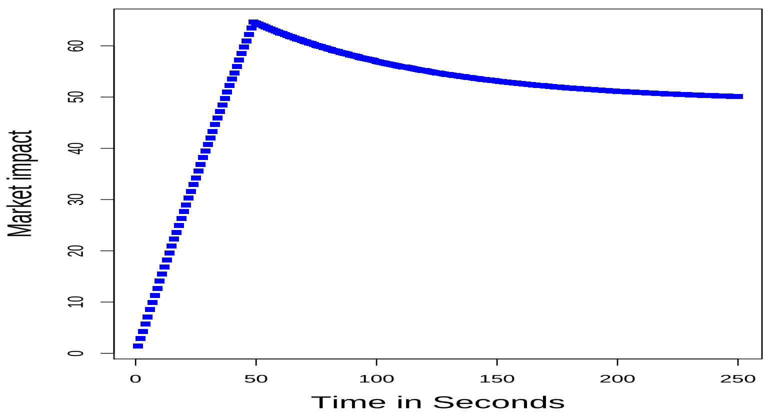

Figure 5: Market impact of an order to buy 50 shares of stock, executed during 50 seconds at the rate of 1 stock per second. We took the scaling parameter $\lambda_{\mathrm{K}} = 1$.

References

F. Abergel and A. Jedidi. A Mathematical Approach to Order Book Modeling. *International Journal of Theoretical and Applied Finance*, 16(5), 2013.

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.

J.-P. Bouchaud, Y. Gefen, M. Potters, and M. Wyart. Fluctuations and response in financial markets: the subtle nature of 'random' price changes. *Quantitative Finance*, 10:176–190, April 2004.

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.

J. Da Fonseca and R. Zaatour. Hawkes process: Fast calibration, application to trade clustering and diffusive limit. *Working Paper SSRN-2294112*, 2013.

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.

A. Dufour and R. F. Engle. Time and the Price Impact of a Trade. *The Journal of Finance*, 55(6), 2000.

E. Errais, K. Giesecke, and L. R. Goldberg. Affine point processes and portfolio credit risk. *SIAM J. Financial Math.*, 1(1):642–665, 2010.

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.

A. S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335, 1985.

C. M. C. Lee and M. J. Ready. Inferring Trade Direction from Intraday Data. *The Journal of Finance*, 46(2), June 1991.

S. Meyn and R. L. Tweedie. *Markov Chains and Stochastic Stability.* Cambridge University Press, 2nd edition, 2009.

E. Moro, J. Vicente, L. G. Moyano, A. Gerig, J. D. Farmer, G. Vaglica, F. Lillo, and R. N. Mantegna. Market impact and trading profile of hidden orders in stock markets. *Physical Revue E*, 80(6), Dec. 2009.

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.

D. Revuz and M. Yor. *Continuous martingales and Brownian motion.* 3rd ed, Springer-Verlag, 1999.

A Appendix

Proof. of Lemma 1.
We apply Dynkin's formula (5) 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:

$$\mathbb{E}[N_t] = N_0 + \int_0^t \mathbb{E}[\lambda_s] ds. \tag{51}$$

Differentiating this integral equation gives (7). 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 (6) we rely again on Dynkin's formula. Following Errais et al. (2010), let $f \equiv \lambda_t$ in (4) 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 (6). ■

Proof. of Lemma 4
We start with

$$\begin{aligned}I_1 &= \mathbb{E}\left[(N_{t+\tau} - N_t)(N_{t+\tau} - N_t)^\top\right] \\ &= \mathbb{E}\left[N_{t+\tau}N_{t+\tau}^\top\right] - \mathbb{E}\left[N_{t+\tau}N_t^\top\right] - \mathbb{E}\left[N_t N_{t+\tau}^\top\right] + \mathbb{E}\left[N_t N_t^\top\right] \qquad (52) \\ &= 2\mathbb{E}\left[N_t N_t^\top\right] + \int_t^{t+\tau} \mathbb{E}\left[\lambda_s N_s^\top\right] + \mathbb{E}\left[N_s\lambda_s^\top\right] + \mathrm{diag}(\mathbb{E}[\lambda_s])ds - \mathbb{E}\left[N_{t+\tau}N_t^\top\right] - \mathbb{E}\left[N_t N_{t+\tau}^\top\right] \qquad (53)\end{aligned}$$

where from (52) to (53) we used (14). Moreover, we have

$$
\begin{aligned}
I_2 &= \int_t^{t+\tau} \mathbb{E}\left[\lambda_s N_s^\top\right] ds \\
&= \int_t^{t+\tau} e^{(\alpha-\beta)(s-t)} \mathbb{E}\left[\lambda_t N_t^\top\right] ds + \int_t^{t+\tau} \int_t^s e^{(\alpha-\beta)(s-u)} \left\{\beta\lambda_\infty \mathbb{E}\left[N_u^\top\right] + \mathbb{E}\left[\lambda_u \lambda_u^\top\right] + \alpha \mathrm{diag}(\mathbb{E}\left[\lambda_u\right])\right\} duds \\
&= \int_t^{t+\tau} e^{(\alpha-\beta)(s-t)} \mathbb{E}\left[\lambda_t N_t^\top\right] ds + \int_t^{t+\tau} \int_t^s e^{(\alpha-\beta)(s-u)} \left\{\beta\lambda_\infty \mathbb{E}\left[N_t^\top\right] + \beta\lambda_\infty \int_t^u \mathbb{E}\left[\lambda_r^\top\right] dr\right\} duds \qquad (54) \\
&+ \int_t^{t+\tau} \int_t^s e^{(\alpha-\beta)(s-u)} \left\{\mathbb{E}\left[\lambda_u \lambda_u^\top\right] + \alpha \mathrm{diag}(\mathbb{E}\left[\lambda_u\right])\right\} duds \qquad (55)
\end{aligned}
$$

where we used successively (15) and (10). The fifth term of (53) is, after using the ODE for $\mathbb{E}[N_t]$ and conveniently conditioning, equal to

$$
\begin{aligned}
I_3 &= \mathbb{E}\left[N_{t+\tau} N_t^\top\right] \\
&= \mathbb{E}\left[\left(N_t + \int_t^{t+\tau} (c_0(s-t)\lambda_t + c_1(s-t)) ds\right) N_t^\top\right]. \qquad (56)
\end{aligned}
$$

The first term of (56) will cancel with the first term of (53), the second term of (56) with will cancel with the first term of (54) whilst the last term of (56) with the second term of (53) when $t \to +\infty$. Therefore, for t large we have

$$
I_1 = \underbrace{\int_t^{t+\tau} \int_t^s e^{(\alpha-\beta)(s-u)} \left\{\beta\lambda_\infty \int_t^u \mathbb{E}\left[\lambda_r^\top\right] dr + \mathbb{E}\left[\lambda_u \lambda_u^\top\right] + \alpha \mathrm{diag}(\mathbb{E}\left[\lambda_u\right])\right\} duds}_{K_1} \qquad (57)
$$

$$
+ K_1^\top + \int_t^{t+\tau} \mathrm{diag}(\mathbb{E}\left[\lambda_s\right]) ds \qquad (58)
$$

Replacing in K_1 the expectations involving λ_t by their long term values we obtain

$$
\begin{aligned}
K_1 &= c_4(\tau)\beta\lambda_\infty \bar{\lambda}_\infty^\top + c_5(\tau)(\Lambda_\infty + \alpha \mathrm{diag}(\bar{\lambda}_\infty)) \qquad (59) \\
c_4(\tau) &= -(\alpha-\beta)^{-1}\frac{\tau^2}{2} - (\alpha-\beta)^{-2}\tau + (\alpha-\beta)^{-3}\left(e^{(\alpha-\beta)\tau} - I\right) \qquad (60) \\
c_5(\tau) &= -(\alpha-\beta)^{-1}\tau + (\alpha-\beta)^{-2}(e^{(\alpha-\beta)\tau} - I) \qquad (61)
\end{aligned}
$$

As $c_4(t)$ is related to $c_5(\tau)$ through

$$
c_4(\tau) = \left(\frac{\tau^2}{2} - c_5(\tau)\right)(-(\alpha-\beta)^{-1}) \qquad (62)
$$

K_1 can be rewritten as

$$
K_1 = \frac{\tau^2}{2}\bar{\lambda}_\infty \bar{\lambda}_\infty^\top + c_5(\tau)(\bar{\Lambda}_\infty + \alpha \mathrm{diag}(\bar{\lambda}_\infty)).
$$

Taking into account

$$\lim_{t\to\infty} \mathbb{E}\left[N_{t+\tau} - N_t\right] \mathbb{E}\left[(N_{t+\tau} - N_t)^\top\right] = \tau^2 \bar{\lambda}_\infty \bar{\lambda}_\infty^\top$$

we deduce the result. ■

Proof. of Lemma 5
We need to determine

$$\begin{aligned}
I_4 &= \mathbb{E}\left[(N_{t_4} - N_{t_3})(N_{t_2} - N_{t_1})^\top\right] \\
&= \mathbb{E}\left[\mathbb{E}_{t_3}\left[(N_{t_4} - N_{t_3})\right](N_{t_2} - N_{t_1})^\top\right] && (63) \\
&= \mathbb{E}\left[(c_2(\tau_2)\lambda_{t_3} + c_3(\tau_2))(N_{t_2} - N_{t_1})^\top\right] && (64) \\
&= c_2(\tau_2)c_0(\delta)\mathbb{E}\left[\lambda_{t_2}(N_{t_2} - N_{t_1})^\top\right] + c_2(\tau_2)c_1(\delta)\mathbb{E}\left[(N_{t_2} - N_{t_1})^\top\right] + c_3(\tau_2)\tau_1\bar{\lambda}_\infty^\top && (65)
\end{aligned}$$

where from (63) to (64) we used (10), and from (64) to (65) we used (9) as well as (13). Taking into account that

$$\begin{aligned}
\mathbb{E}\left[\lambda_{t_2}(N_{t_2} - N_{t_1})^\top\right] &= \mathbb{E}\left[\lambda_{t_2} N_{t_2}^\top\right] - \mathbb{E}\left[\lambda_{t_2} N_{t_1}^\top\right] \\
&= e^{(\alpha-\beta)\tau_1}\mathbb{E}\left[\lambda_{t_1} N_{t_1}^\top\right] + \int_{t_1}^{t_2} e^{(\alpha-\beta)(t_2-s)}\left\{\beta\lambda_\infty \mathbb{E}\left[N_s^\top\right] + \mathbb{E}\left[\lambda_s\lambda_s^\top\right] + \alpha\mathrm{diag}(\mathbb{E}\left[\lambda_s\right])\right\} ds \\
&- \left(c_0(\tau_2)\mathbb{E}\left[\lambda_{t_1} N_{t_1}^\top\right] + c_1(\tau_2)\mathbb{E}\left[N_{t_1}^\top\right]\right).
\end{aligned}$$

The first term of the last equation simplifies with last-but-one term. Replacing $\mathbb{E}\left[N_s^\top\right]$ by its integral given by (10) allows us to simply the last term of the equation and we are left with

$$\mathbb{E}\left[\lambda_{t_2}(N_{t_2} - N_{t_1})^\top\right] = \int_{t_1}^{t_2} e^{(\alpha-\beta)(t_2-s)}\left\{\beta\lambda_\infty \int_{t_1}^{s} \mathbb{E}\left[\lambda_u^\top\right] du + \mathbb{E}\left[\lambda_s\lambda_s^\top\right] + \alpha\mathrm{diag}(\mathbb{E}\left[\lambda_s\right])\right\} ds$$

Taking the long term values for the expectations (involving only the process λ_t) we get

$$\mathbb{E}\left[\lambda_{t_2}(N_{t_2} - N_{t_1})^\top\right] = c_5(\tau_1)\beta\lambda_\infty\bar{\lambda}_\infty^\top + c_2(\tau_1)\left(\Lambda_\infty + \alpha\mathrm{diag}(\bar{\lambda}_\infty)\right). \quad (66)$$

As $c_5(\tau)$ is related to $c_2(\tau)$ through

$$c_5(\tau) = (\tau - c_2(\tau))\left(-(\alpha - \beta)^{-1}\right) \quad (67)$$

when used in conjunction with (66) in (65) leads to, after taking into account (13) for the second term and the definition for $\bar{\lambda}_\infty$ given by (11), it gives

$$\begin{aligned}
I_4 &= c_2(\tau_2)c_0(\delta)\left\{\tau_1\bar{\lambda}_\infty\bar{\lambda}_\infty^\top + c_2(\tau_1)\left(\bar{\Lambda}_\infty + \alpha\mathrm{diag}(\bar{\lambda}_\infty)\right)\right\} \\
&+ c_2(\tau_2)c_1(\delta)\bar{\lambda}_\infty^\top\tau_1 + c_3(\tau_2)\tau_1\bar{\lambda}_\infty^\top.
\end{aligned}$$

Taking into account the equalities $c_3(\tau) = (\tau - c_2(\tau))\bar{\lambda}_\infty$ and $c_1(\delta) = (I - c_0(\delta))\bar{\lambda}_\infty$ then if we subtract to I_4 the following quantity

$$\lim_{t_1 \to \infty} \mathbb{E}\left[N_{t_4} - N_{t_3}\right] \mathbb{E}\left[(N_{t_2} - N_{t_1})^\top\right] = \tau_2 \tau_1 \bar{\lambda}_\infty \bar{\lambda}_\infty^\top$$

we obtain the result. ■

Proof. of Lemma 6

Under the hypothesis of the lemma we can decompose the expectation as

$$\begin{aligned}
I_5 &= \mathbb{E}[(N_{t_4} - N_{t_3})(N_{t_2} - N_{t_1})^\top] \\
&= \mathbb{E}[((N_{t_4} - N_{t_2}) + (N_{t_2} - N_{t_3}))\,(N_{t_2} - N_{t_1})^\top] \\
&= \mathbb{E}[(N_{t_4} - N_{t_2})(N_{t_2} - N_{t_1})^\top] + \mathbb{E}[(N_{t_2} - N_{t_3})(N_{t_2} - N_{t_1})^\top] \\
&= \mathbb{E}[(N_{t_4} - N_{t_2})(N_{t_2} - N_{t_1})^\top] + \mathbb{E}[(N_{t_2} - N_{t_3})(N_{t_2} - N_{t_3})^\top] \\
&+ \mathbb{E}[(N_{t_2} - N_{t_3})(N_{t_3} - N_{t_1})^\top].
\end{aligned}$$

Similarly, if we decompose the product $\mathbb{E}[N_{t_4} - N_{t_3}]\mathbb{E}[(N_{t_2} - N_{t_1})^\top]$ then using Lemma 5 we obtain the announced result. ■

Proof. of Lemma 10

The volatility is given by

$$\begin{aligned}
\sigma^2 &= \lim_{n \to \infty} \frac{\mathrm{Var}(S_n)}{n} && (68)\\
&= \frac{\nu^2}{4}\mathbb{E}[\Big((N_1^u - N_0^u) - \Big(N_1^d - N_0^d\Big)\Big)^2] && (69)\\
&+ 2\frac{\nu^2}{4}\sum_{n=1}^{\infty}\mathbb{E}[\Big((N_1^u - N_0^u) - \Big(N_1^d - N_0^d\Big)\Big)\Big((N_{1+n}^u - N_n^u) - \Big(N_{1+n}^d - N_n^d\Big)\Big)] && (70)\\
&= \frac{\nu^2}{4}\left(\mathrm{Cov}(1)_{11} + \mathrm{Cov}(1)_{22} - \mathrm{Cov}(1)_{12} - \mathrm{Cov}(1)_{21}\right) && (71)\\
&+ 2\frac{\nu^2}{4}\left(\bar{\Sigma}_{11} + \bar{\Sigma}_{22} - \bar{\Sigma}_{12} - \bar{\Sigma}_{21}\right). && (72)
\end{aligned}$$

Define $M = \mathrm{Cov}(1) + 2\bar{\Sigma}$ then we obtain the result. ■

Proof. of Proposition

Given τ_0, τ and δ the quantity

$$\begin{aligned}
\mathbb{E}[(S_{t+\tau+\delta} - S_{t+\delta})(N_{t+\tau_0}^u - N_t^u)] &= \frac{\nu}{2}\mathbb{E}[(N_{t+\tau+\delta}^u - N_{t+\delta}^u)(N_{t+\tau_0}^u - N_t^u)] && (73)\\
&- \frac{\nu}{2}\mathbb{E}[(N_{t+\tau+\delta}^d - N_{t+\delta}^d)(N_{t+\tau_0}^u - N_t^u)]. && (74)
\end{aligned}$$

As we have:

$$\begin{aligned}
I_1 &= \mathbb{E}[(N_{t+\tau+\delta}^u - N_{t+\delta}^u)(N_{t+\tau_0}^u - N_t^u)] \\
&= \sum_{i=0}^{\infty}\mathbb{E}[N_{t+\tau+\delta}^u - N_{t+\delta}^u | N_{t+\tau_0}^u - N_t^u = i] \times P[N_{t+\tau_0}^u - N_t^u = i] \times i \\
&\sim \mathbb{E}[N_{t+\tau+\delta}^u - N_{t+\delta}^u | N_{t+\tau_0}^u - N_t^u = 1] \times \lambda_t^u \tau_0
\end{aligned}$$

where the last equation is for τ_0 small enough. Then, taking the limit $\tau_0 \to 0$, we get for t large:

$$\lim_{\tau_0 \to 0} \frac{I_1}{\tau_0} \sim \mathbb{E}[N^u_{t+\tau+\delta} - N^u_{t+\delta} | dN^u_t = 1] \bar{\lambda}^u_\infty.$$

Moreover, we have:

$$\mathbb{E}[(N^u_{t+\tau+\delta} - N^u_{t+\delta})(N^u_{t+\tau_0} - N^u_t)] = (\mathrm{Cov}_1(\tau_0, \tau, \delta))_{11},$$

we deduce:

$$\mathbb{E}[N^u_{t+\tau+\delta} - N^u_{t+\delta} | dN^u_t = 1] = \frac{1}{\bar{\lambda}^u_\infty}(c_2(\tau)c_0(\delta)(\bar{\Lambda}_\infty + \alpha \mathrm{diag}(\bar{\lambda}_\infty)))_{11}.$$

Similar computations can be carried out for (74) and the announced result is obtained. ■

Proof. of Proposition 9
The computations are analytically tractable as we have closed formula for the inverse and exponential of a 2x2 symmetric matrix:

$$M = \begin{pmatrix} a & b \\ b & a \end{pmatrix} \tag{75}$$

$$e^M = \frac{e^{a-b}}{2} \begin{pmatrix} 1 + e^{2b} & -1 + e^{2b} \\ -1 + e^{2b} & 1 + e^{2b} \end{pmatrix} \tag{76}$$

and the inverse of M writes:

$$M^{-1} = \frac{1}{a^2 - b^2} \begin{pmatrix} a & -b \\ -b & a \end{pmatrix} \tag{77}$$

yielding:

$$M_{11} + M_{22} - M_{12} - M_{21} = 2\Lambda \left(\kappa^2 \tau + \left(1 - \kappa^2\right) \frac{(1 - e^{-\tau\gamma})}{\gamma} \right)$$

so that the expected signature plot is given by:

$$\mathbb{E}[\hat{C}(\tau)] = \frac{\nu^2}{4\tau}(M_{11} + M_{22} - M_{12} - M_{21})$$

which leads to the announced result. ■