Market impact, price and trades high-frequency dynamics

Mohamed ABD EL WEDOUD and Abdollah RIDA

7 December 2018

Abstract

We start by a quick literature review (based on [3]) and an introduction of the square-root law, before moving to Hawkes' point processes [11] and studying in depth the theory behind self-exciting point processes. We then reintroduce the multivariate Hawkes process that accounts for the dynamics of market prices through the impact of market order arrivals at microstructural level presented in [4]. We try to rebuild the same multivariate Hawkes process and estimate the kernels of the Hawkes' from the empirical conditional mean intensities provided by two different order books using two different techniques. We then provide a computation of the estimated market impact profile from the provided data and briefly discuss extended models.

1 Introduction

An order book is an electronic list of buy and sell orders for a specific security or financial instrument, organized by price level. An order book lists the number of shares being bid or offered at each price point, or market depth. It also identifies the market participants behind the buy and sell orders, although some choose to remain anonymous. An order book is dynamic and constantly updated in real time throughout the day. Exchanges such as Nasdaq refer to it as the "continuous book." Orders that specify execution only at market open or market close are maintained separately. These are known as the "opening (order) book" and "closing (order) book," respectively. These orders can only be placed on a predetermined price grid whose size is called the **tick** size. Therefore an order corresponds to a disclosed buy/sell intention for a fixed volume and at a fixed price. One of the most common ways of ordering orders that fall at the same price level is the *first-in-first-out* time priority queue. The **bid** corresponds to the highest available price for a buy order, and the **ask** to the lowest available price for a sell order.

The order book information helps traders make better-informed trading decisions, since they can see order imbalances, which may provide clues to the stock's direction in the very short term. A massive imbalance of buy orders

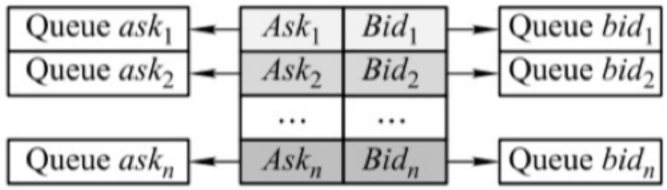

Figure 1: Order Book model

compared to sell orders, for instance, may indicate a move higher in the stock due to buying pressure. This is a form of price impact. Price impact refers to the correlation between an incoming order (to buy or to sell) and the price change resulting from that. As we have seen before, a buy trade imbalance naturally pushes the price up. But how about micro-imbalances? What impact do they have?

2 Literature review

There are several types of orders in an order book: market orders (order to buy or sell a security immediately), stop orders (order to buy or sell a stock once the price of the stock reaches the specified price) and limit orders (order to buy or sell a security at a specific price or better). While limit orders aren't the main focus of this work, they are the basis backbone of most modern financial networks. Historically, the task of supplying liquidity by permanently maintaining limit orders in the book was assumed by designated market makers who, in exchange for this service, kept a spread, namely, they offered to buy at a price lower than their sell price, leading to profit on each transaction. All other actors were liquidity takers, being forced to interact with a market maker. In reality, the idea of market makers profiting on each transaction due to the bid-ask spread is hindered by the challenge of adverse selection: if an informed trader has an accurate prediction about the future evolution of the price (while the market maker is less informed and has not updated his quotes accordingly), he can profit by entering a transaction with them. Due to the nature of their activity, market-makers are not usually affected by long-term trends in price. On the other hand, the second class of market actors are typically in the opposing situation. To understand the difficulties that this second class of actors faces when executing a large order, we need to introduce the concept of liquidity risk.

2.1 Liquidity risk

The concept of Liquidity Risk reflects the extra cost incurred during a buy (resp. sell) order that is due to the scarcity of supply (resp. demand). In the most extreme cases, it can even mean that it is impossible to trade an asset due to the absence of counterpart. One striking example is that during the subprimes crisis, products such as CDOs became practically unsaleable. When a trader seeks to execute a large transaction, the scarcity of instantaneous liquidity means that the order must be executed incrementally. Hence the importance of devising execution strategies that minimize the cost to the investor: given a time horizon and a volume, the investor/ the broker distributes the execution over time so as to obtain the best average execution price.

The concept of price impact is fundamental when it comes to designing execution strategies for large orders. Since the available liquidity at a given time is not sufficient to absorb the entire order, the trader must split his order into several chunks to be executed incrementally. Every time a small order - also called **child order** - is executed, the price is mechanically pushed in its direction, making the average execution price higher than the decision price and leading to the notion of execution shortfall. If the market can guess that the trader intends to buy (resp. sell) large quantities, he can be outrun by informed traders who push the price up (resp. down) in order to benefit from his orders. Without such price pressure, all trading strategies would be infinitely scalable, as the cost of trading would remain unchanged despite the size of the trade increasing. This does not seem plausible and contradicts the fact that at any given time, there are only limited amounts of liquidity available in the real order book.

2.2 The empirical Square-root model

Let $I(Q)$ be the average price variation after executing a volume Q:

$$I(Q) := \mathrm{E}[P_T - P_0 | Q]$$

where P_0 (resp. P_T) is the price of the first (resp. last) child order. Then the typical market impact law reads:

$$I(Q) = \alpha \, \sigma \sqrt{\frac{Q}{V}}$$

Where Q is the total executed volume, V is the (average) daily traded volume, σ the daily volatility, and α a homogenization constant of order 1.

This law has been empirically verified for a large panel of markets and instruments. The square-root law implies that the impact of trading only depends on the traded volume, and not on the duration of execution and the execution path.

This figure, taken from [2], shows that the square-root market impact formula is verified empirically for meta-orders with a range of sizes spanning two to three orders of magnitude! This law should therefore be seen as a (good) first-order approximation and a way to benchmark market impact models. It indicates that markets are inherently fragile: the impact of small volumes is

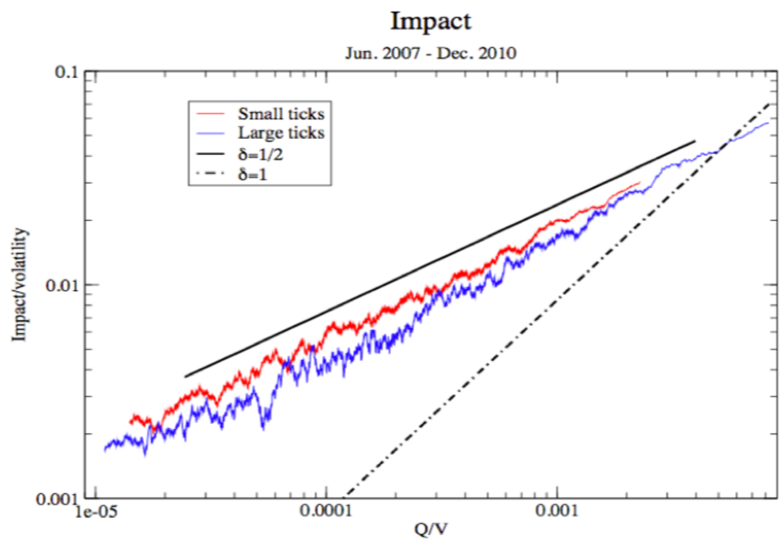

Figure 2: Log-log plot of the volatility-adjusted price impact vs the ratio Q/V taken from [2]

disproportionately high. While this result may seem counter-intuitive, it is perfectly in accordance with the fact that the instantaneous supply of liquidity is limited in real markets. This empirical law also asserts that impact is non-additive but strictly concave: after having traded a $\frac{Q}{2}$ volume, the next $\frac{Q}{2}$ will have less impact on the price.

3 Hawkes' point processes

Trades do not arrive in evenly spaced intervals but usually arrive clustered in time. This should be clear for anyone who has been watching an order book for some time. Similarly, the same trade signs tend to cluster together and result in a sequence of buy or sell orders. Various explanations for this are possible, such as algorithmic traders who split up their orders in smaller blocks or trading systems that react to certain exchange events.

For the sake of demonstration, we will work on dummy data representing 5000 trades between 13:10 and 19:57 on the 20 April 2013. Here is the plot of the trade counts aggregated over a 1 minute window:

The average trade count per minute is 13, however we can make out a couple of instances where it exceeds 50. Usually the higher trade intensity lasts a couple of minutes and then dies down again towards the mean. In particular, the 15 minutes or so after 16:00 we can see very high trading intensity with one instance of over 200 orders per minute, then a slow decrease of intensity over the next

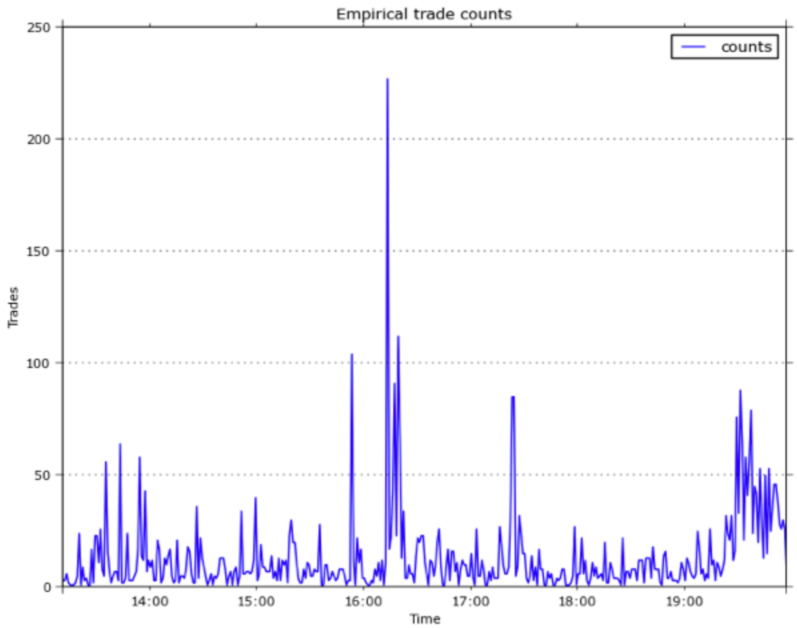

Figure 3: Trade counts aggregated over a 1 minute window

10 minutes.

The most basic way to describe arrival of event counts, such as the time series above, is a Poisson process with one parameters λ. In a Poisson process the expected number of events per unit of time is defined by the one parameter. This method is widely used as it fits well to a lot of data, such as the arrival of telephone calls in a call centre. For our purposes however this is too simple as we need a way to explain the clustering and mean reversion.

Hawkes processes, or also called self-exciting processes, are an extension of the basic Poisson process which aim to explain such clustering. Self-excitable models like this are widely used in various sciences; some examples are seismology (modelling of earthquakes and volcanic eruptions), ecology (wildfire assessment [20]), neuro-science (modelling of brain spike trains which bunch together [18]), even modelling of eruption of violence ([6] on modelling civilian deaths in Iraq, and [10] on crime forecasting), and naturally finance and trading.

3.1 Mathematical notions

3.1.1 Definition and first examples

Definition 1. A general definition of a self-exciting process N reads:

$$\lambda(t) = \lambda_0(t) + \int_{-\infty}^{t} \nu(t-s)\, dN_s$$

where $\lambda_0 : \mathrm{R} \longrightarrow \mathrm{R}_+$ is a deterministic base intensity and $\nu : \mathrm{R}_+ \longrightarrow \mathrm{R}_+$ expresses the positive influence of the past events t_i on the current value of the intensity process.

Definition 2. Hawkes [11] proposes an exponential kernel:

$$\nu(t) = \sum_{j=1}^{P} \alpha_j\, e^{-\beta_j t}\, 1_{\mathrm{R}_+}$$

The intensity becomes:

$$\lambda(t) = \lambda_0(t) + \sum_{t_i<t} \sum_{j=1}^{P} \alpha_j\, e^{-\beta_j (t-t_i)}$$

Here, λ_0 is the base rate the process reverts to, α is the intensity jump right after an event occurrence, and β is the exponential intensity decay. The base rate can also be interpreted as the intensity of exogenous events (to the process) such as news. The other parameters α and β define the clustering properties of the process. It is usually the case that $\alpha < \beta$ which ensures that the intensity decreases quicker than new events increase it – otherwise the process could explode (more on this in the next subsection).

An example realization of a unidimensional Hawkes process is plotted in the next figure for $\lambda_0 = 0.5$, $\alpha = 0.1$ and $\beta = 1$:

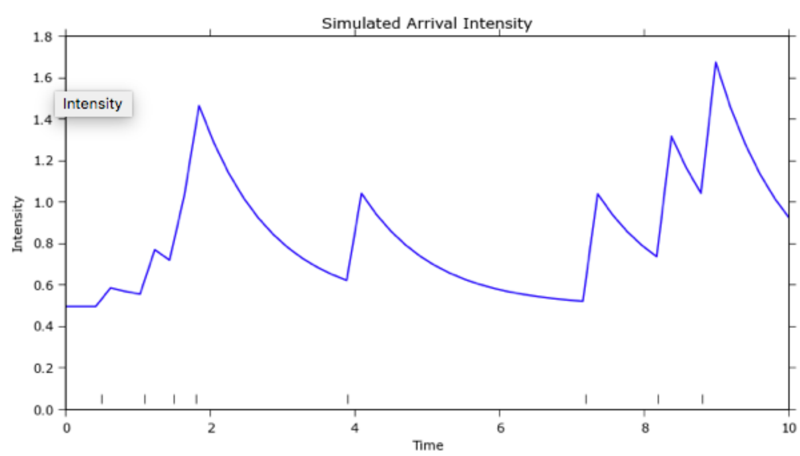

Figure 4: Intensity plot for a Hawkes process with 8 events

3.1.2 Stationarity

Theorem 1. (Stationarity) *Assuming stationarity gives* $\mathrm{E}[\lambda(t)] = \mu$ *constant, we can show that the stationarity condition can be written as:*

$$\sum_{j=1}^{P} \frac{\alpha_j}{\beta_j} < 1$$

Proof.

$$\begin{aligned} \mu &= \mathrm{E}[\lambda(t)] \\ &= \lambda_0 + \mathrm{E}[\int_{-\infty}^{t} \nu(t-s)\, dN_s] \\ &= \lambda_0 + \mathrm{E}[\int_{-\infty}^{t} \nu(t-s)\, \lambda(s)\, ds] \end{aligned}$$

Using Fubini we then find that:

$$\begin{aligned} &= \lambda_0 + \int_{-\infty}^{t} \nu(t-s)\, \mu\, ds \\ &= \lambda_0 + \mu \int_{0}^{\infty} \nu(u)\, du \end{aligned}$$

Which gives:

$$\mu = \frac{\lambda_0}{1 - \int_0^\infty \nu(u)\, du}$$

All we need to do is compute the value of the integral for the exponential kernel to prove the stationarity theorem. Q.E.D. □

The latter proof immediately gives for the one-dimensional Hawkes process.

Lemma 1. *with* $P = 1$ *the unconditional expected value of the intensity process is:*

$$\mathrm{E}[\lambda(t)] = \frac{\lambda_0}{1 - \frac{\alpha}{\beta}}$$

3.1.3 Multi-dimensional Hawkes' Process

Let $M \in \mathbf{N}^*$ and $(t_i^m)_i$ an M-dimensional point process. We will denote by $N_t = (N_t^1, ..., N_t^M)$ the associated counting process.

Definition 3. A multi-dimensional Hawkes' process is defined with intensities λ^m given by:

$$\lambda^m(t) = \lambda_0^m(t) + \sum_{n=0}^{M} \int_0^t \sum_{j=1}^{P} \alpha_j^{mn}\, e^{-\beta_j^{mn}(t-s)}\, dN_s^n$$

In its simplest form (i.e. $P = 1$ and λ_0^m constant) the expression above can be rewritten:

$$\lambda^m(t) = \lambda_0^m + \sum_{n=0}^{M} \sum_{t_i^n < t} \alpha_j^{mn} \, e^{-\beta_j^{mn}(t - t_i^n)}$$

We can rewrite the above equation using vectorial notations:

$$\lambda(t) = \lambda_0 + \int_0^t G(t-s) dN_s$$

where

$$G(t-s) = \left(\alpha_j^{mn} \, e^{-\beta_j^{mn}(t-s)}\right)_{m,n}$$

Assuming stationarity gives $\mathbf{E}[\lambda(t)] = \mu$ a constant vector, and thus stationary intensities must satisfy:

$$\mu = \left(\mathbf{I} - \int_0^\infty G(u) \, du\right)^{-1} \lambda_0$$

Theorem 2. Stationarity of a multivariate Hawkes' Process *A sufficient condition for a multivariate Hawkes' process to be linear is that the spectral radius of the matrix:*

$$\Gamma = \int_0^\infty G(u) \, du$$

be strictly smaller than 1.

3.2 Simulation of a Hawkes process

3.2.1 Lewis' Algorithm

Lewis & Shedler [14] proposes a "thinning procedure" that allows the simulation of a point process with bounded intensity.

Theorem 3. (Basic Thinning Theorem)
Consider a one-dimensional non-homogeneous Poisson process $(N^*(t))_{t \geq 0}$ *with rate function* $\lambda^*(t)$*, so that the number of points* $N^*(T_0)$ *in a fixed interval* $]0, T_0]$ *has a Poisson distribution with parameter* $\mu_0^* = \int_0^{T_0} \lambda^*(s) \, ds$.
Let $t_1^*, \ldots, t_{N^*(T_0)}^*$ *be the points of the process in the interval* $]0, T_0]$. *Suppose that for* $0 \leq t \leq T0, \lambda(t) \leq \lambda^*(t)$.

For $i = 1, 2, \ldots, N^*(T_0)$*, delete the points* t_i *with probability* $1 - \lambda^{(t_i^)}$. *Then the remaining points form a non-homogeneous Poisson process* $(N(t))_{t \geq 0}$ *with rate function* $\lambda(t)$ *in the interval* $]0, T_0]$.

3.2.2 Ogata's Algorithm

Ogata [17] for the simulation of Hawkes processes. Let $\mathcal{U}_{[0,1]}$ denote the uniform distribution on the interval $[0,1]$ and $[0,T]$ the time interval on which the process is to be simulated. We'll assume here that $P = 1$.

We start by initializing:
$\lambda^* \leftarrow \lambda_0(0), n \leftarrow 1$

We then create the first event:
Generate $U \hookrightarrow \mathcal{U}_{[0,1]}$
$s \leftarrow -\frac{1}{\lambda^*} ln(U)$
if $s \leq T$ **then**
$t_1 \leftarrow s$
else
Go to last step
end if

The General routine is then as follows:
$n \leftarrow n + 1$

- Update Maximum intensity
  $\lambda^* \leftarrow \lambda(t_{n-1}) + \alpha$

  λ^* exhibits a jump of size α as an event has just occurred. λ being left-continuous, this jump is not counted in $\lambda(t_{n-1})$, hence the explicit addition.

- New event
  Generate $U \hookrightarrow \mathcal{U}_{[0,1]}$
  $s \leftarrow -\frac{1}{\lambda^*} ln(U)$
  if $s \geq T$ **then**
  Go to last step
  end if

- Rejection test
  Generate $D \hookrightarrow \mathcal{U}_{[0,1]}$
  if $s \leq \frac{\lambda(s)}{\lambda^*}$ **then**
  $t \leftarrow s$
  Go through general routine again
  else
  $\lambda^* \leftarrow \lambda(s)$
  try a new date at step 2 of the general routine.
  end if

Output: Retrieve the simulated process t_n on $[0,T]$.

3.2.3 Examples

Here's an example using the tick [7] package. The parameters are $\lambda_0 = 1.2$, $\alpha = 0.6$ and $\beta = 0.8$.

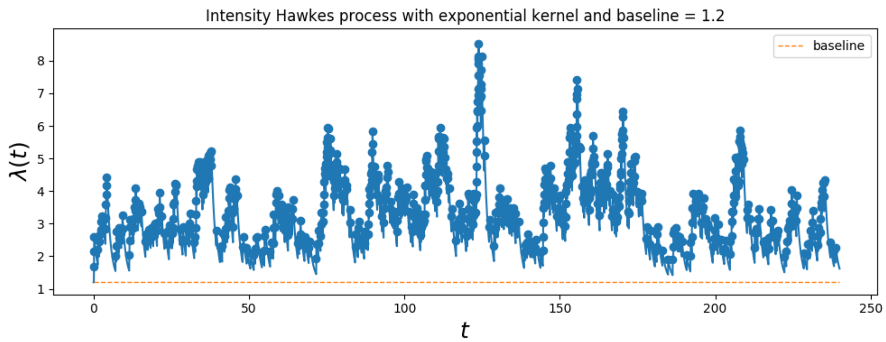

Figure 5: Simulation of a one-dimensional Hawkes process

3.3 Fitting data to a Hawkes process

3.3.1 Likelihood of a Hawkes process

The intensity path is fully defined given a set of ordered trade times $t_1 < ... < t_n$ which, in our case, are just unix timestamps of when the trades were recorded. Given this we can easily apply Maximum Likelihood Estimation to fit the model parameters. The log-likelihood of a simple point process N with intensity λ is written:

$$log\mathcal{L}((N_t)_t) = \int_0^T (1 - \lambda(s))\, ds + \int_0^T log(\lambda(s))\, dN_s$$

Which in the case of a Hawkes process can be explicitely computed as (we define $\Lambda(t,s) := \int_t^s \lambda(s)\, ds$):

$$\begin{aligned} log\mathcal{L}((t_i)_i) &= t_n - \Lambda(0, t_n) + \sum_{i=1}^{n} log(\lambda(t_i)) \\ &= t_n - \Lambda(0, t_n) + \sum_{i=1}^{n} log\left[\lambda_0(t_i) + \sum_{j=1}^{P}\sum_{k=1}^{i-1} \alpha_j\, e^{-\beta_j(t_i - t_k)}\right] \end{aligned}$$

As noted by Ogata [17], this log-likelihood function is easily computed with a recursive formula. We observe that, by denoting $R_j(i) = \sum_{k=1}^{i-1} e^{-\beta_j(t_i - t_k)}$:

$$
\begin{aligned}
R_j(i) &= \sum_{k=1}^{i-1} e^{-\beta_j(t_i - t_k)} \\
&= e^{-\beta_j(t_i - t_{i-1})} \left(1 + R_j(i-1)\right)
\end{aligned}
$$

The log-likelihood can thus be recursively computed with:

$$
log\mathcal{L}((N_t)_t) = t_n - \Lambda(0, t_n) + \sum_{i=1}^{n} log\left[\lambda_0(t_i) + \sum_{j=1}^{P} \alpha_j \, R_j(i)\right]
$$

Direct computation leads then to:

Theorem 4. (Log-likelihood of a 1D-Hawkes process)

$$
\begin{aligned}
log\mathcal{L}((N_t)_t) =& t_n - \int_0^{t_n} \lambda_0(s)\, ds - \sum_{i=1}^{n} \sum_{j=1}^{P} \frac{\alpha_j}{\beta_j}(1 - e^{-\beta_j(t_n - t_i)}) \\
&+ \sum_{i=1}^{n} log\left[\lambda_0(t_i) + \sum_{j=1}^{P} \alpha_j \, R_j(i)\right]
\end{aligned}
$$

We can also write the log-likelihood for a multivariate process. Let's first define a useful function that will be helpful later:

Definition 4. We define the following function:

$$
\begin{aligned}
R_j^{mn}(l) &= \sum_{t_k^n < t_l^n} e^{-\beta_j^{mn}(t_l^m - t_k^n)} \\
&= \begin{cases} e^{-\beta_j^{mn}(t_l^m - t_{l-1}^m)} \, R_j^{mn}(l-1) + \sum_{t_{l-1}^m < t_k^n < t_l^m} e^{-\beta_j^{mn}(t_l^m - t_k^n)} & \text{if} \quad m \neq n \\ e^{-\beta_j^{mn}(t_l^m - t_{l-1}^m)} \left(1 + R_j^{mn}(l-1)\right) & \text{if} \quad m = n \end{cases}
\end{aligned}
$$

The formula for the multivariate log-likelihood is thus:

Theorem 5. (Log-likelihood of a multivariate Hawkes process)

$$
\begin{aligned}
log\mathcal{L}^m(t_i) =& T - \sum_{i=1}^{n} \sum_{n=1}^{M} \sum_{j=1}^{P} \frac{\alpha_j^{mn}}{\beta_j^{mn}} \left(1 - e^{-\beta_j^{mn}(T - t_i)}\right) \\
&+ \sum_{t_l^m} log\left[\lambda_0^m(t_l^m) + \sum_{n=1}^{M} \sum_{j=1}^{P} \alpha_j^{mn} \, R_j^{mn}(l)\right]
\end{aligned}
$$

where R_j^{mn} is defined as above and $R_j^{mn}(0) = 0$.

3.3.2 Properties of the Maximum Likelihood Estimator

Ogata shows that for a stationary one-dimensional Hawkes process with constant λ_0 and $P = 1$, the maximum-likelihood estimator $\hat{\theta}^T = (\hat{\lambda_0}, \hat{\alpha_1}, \hat{\beta_1})$ is:

- Consistent, i.e. converges in probability to the true values $\theta = (\lambda_0, \alpha_1, \beta_1)$ as $T \longrightarrow \infty$.
- Asymptotically normal, i.e.
$$\sqrt{T}(\hat{\theta}^T - \theta) \Longrightarrow \mathcal{N}(0, I^{-1}(\theta))$$
where $I^{-1}(\theta) = \left(\mathbf{E}\left[\frac{1}{\lambda}\frac{\partial\lambda}{\partial\theta_i}\frac{\partial\lambda}{\partial\theta_j}\right]\right)_{ij}$
- Asymptotically efficient, i.e. asymptotically reaches the lower bound of the variance.

3.3.3 Goodness of fit

There are many ways of evaluating the goodness of fit of the estimated Hawkes model. One is by comparing Akaike Information Criterion (AIC) values against a homogeneous Poisson model.We define the AIC by:

Definition 5. Suppose that we have a statistical model of some data. Let k be the number of estimated parameters in the model. Let $\hat{L}$ be the maximum value of the likelihood function for the model. Then the AIC value of the model is the following.

$$\text{AIC} = 2k - 2\ln(\hat{L})$$

Another way to test how well the model fits the data is by evaluating the residuals. Theory says [15] if the model is a good fit, then the residual process should be homogeneous and the Quantile–Quantile plots of the simulated process and the data should be around the 45 line.Another way of assessing the goodness of fit of a model is the Kolmogorov-Smirnov statistical test.

Definition 6. The empirical distribution function F_n for n iid ordered observations X_i is defined as:

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{[-\infty,x]}(X_i)$$

where $I_{[-\infty,x]}$ is the indicator function of $[-\infty, x]$.

The Kolmogorov-Smirnov statistic for a given cumulative distribution function $F(x)$ is:

$$D_n = \sup_x |F_n(x) - F(x)|$$

The test then consists in accepting or rejecting the null hypothesis "the two samples are drawn from the same distribution" by computing the Kolmogorov-Smirnov statistic against a fixed confidence level.

4 First applications

Now that we know the model explains clustering of arrivals well, how can this be applied to trading? The next steps would be to at least consider buy and sell arrivals individually and find a way to make predictions given a fitted Hawkes model. These intensity predictions can then form a part of a market-making or directional strategy. Let us have a look at some basic models first.

4.1 The data

The data used is the is the intraday trade (at best bid/ask) data associated with the most liquid maturity of Eurostoxx (FXSE) and Euro-Bund (FGBL) future contracts. Each time series covers a a period of 800 trading days going from May 2009 to September 2012. Here are some statistics about the database:

Statistic	AskPriceAfter	AskQtyAfter	BidPriceAfter	BidQtyAfter
Mean	7980.09	4.78	7979.16	4.78
std	29.98	4.85	30.004	5.13

We also plotted the mid-price evolution throughout a trading day for the Eurostoxx:

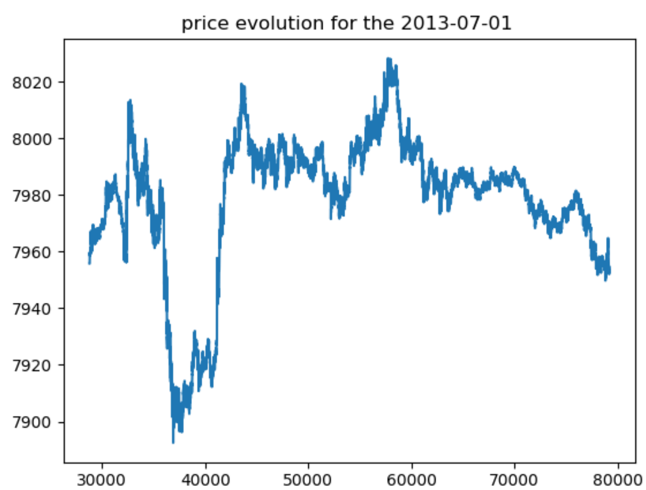

Figure 6: Mid-price evolution through a single day for the Eurostoxx data

4.2 A simple model for buy and sell intensities

Hewlett [12] proposes to model the clustered arrivals of buy and sell trades using Hawkes processes. Using the exponent B for buy variables and S for sell variables, the model is written :

$$\lambda^B(t) = \lambda_0^B + \int_0^t \alpha^{BB} e^{-\beta^{BB}(t-u)} dN_u^B + \int_0^t \alpha^{BS} e^{-\beta^{BS}(t-u)} dN_u^S$$

$$\lambda^S(t) = \lambda_0^S + \int_0^t \alpha^{SB} e^{-\beta^{SB}(t-u)} dN_u^B + \int_0^t \alpha^{SS} e^{-\beta^{SS}(t-u)} dN_u^S$$

Hewlett [12] then imposes some symmetry constraints, stating that mutual excitation and self-excitation should be the same for both processes, which is written:

$$\begin{aligned} \lambda_0^B &= \lambda_0^S = \lambda_0 \\ \alpha^{SB} &= \alpha^{BS} = \alpha^{cross} \\ \beta^{SB} &= \beta^{BS} = \beta^{cross} \\ \alpha^{SS} &= \alpha^{BB} = \alpha^{self} \\ \beta^{SS} &= \beta^{BB} = \beta^{self} \end{aligned}$$

Hewlett fits his model on two-month data of EUR/PLN transactions (no dates given): the Hawkes model is a much better fit of the empirical data than the Poisson model.

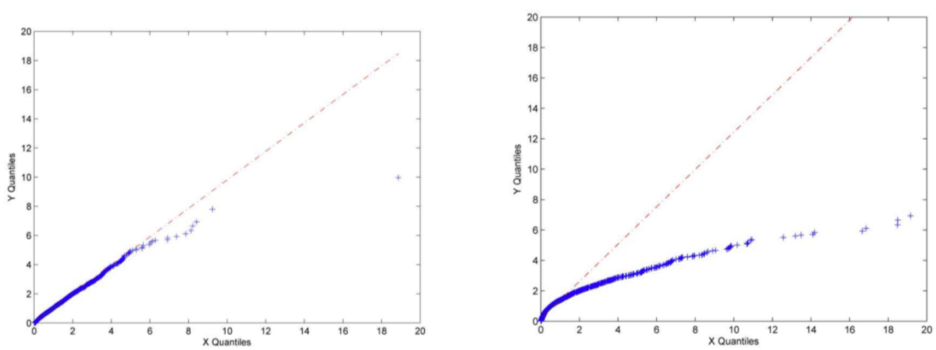

Figure 7: Quantile plots of integrated intensities for the Hawkes model (left) and a Poisson model (right) on EUR/PLN buy and sell data. Reproduced from (Hewlett 2006).

The numerical values obtained are:

$$\lambda_0 = 0.0033\,,\ \alpha^{cross} = 0\,,\ \alpha^{self} = 0.0169\,,\ \beta^{self} = 0.0286$$

In other words:

- The occurrence of a buy (resp. sell) order has an exciting effect on the stream of buy (resp. sell) orders, with a typical half-life of $\frac{log(2)}{\beta^{self}} \approx 24$ seconds.

- The zero value of α^{cross} tends to indicate that there is no influence of buy orders on sell orders, and conversely.

4.3 Hawkes model for price and trades high-frequency dynamics

The aim of E. Bacry and J.F. Muzy in [4] is to define a realistic continuous time microstructure price model that accounts for the impact of market orders. The price changes are thus represented by a 4-dimensional point process:

$$P_t = \begin{pmatrix} T_t^- \\ T_t^+ \\ N_t^- \\ N_t^+ \end{pmatrix}$$

Where T_t^+ represents the accumulated number of market orders arrived before time t representing the best ask and T_t^- the best bid. And N_t^+ (resp. N_t^-) represents the number of upward (resp. downward) price jumps at time t. The conditional intensity vector associated to P_t is denoted by:

$$\lambda_t = \begin{pmatrix} \lambda_t^{T^-} \\ \lambda_t^{T^+} \\ \lambda_t^{N^-} \\ \lambda_t^{N^+} \end{pmatrix}$$

4.3.1 The model

The model in [4] consists in considering that P_t is a Hawkes' process with a Hawkes' kernel Φ_t. Φ_t is a 4×4 matrix whose elements are the causal positive functions explaining the influence of a component over another component. We can decompose it as four 2×2 matrices in the following way:

$$\Phi_t = \begin{pmatrix} \Phi_t^T & \Phi_t^F \\ \Phi_t^I & \Phi_t^N \end{pmatrix}$$

where

- Φ^T (influence of T on λ^T) accounts for the trade correlations (e.g. splitting, herding, etc...)

- Φ^I (influence of T on λ^N) accounts for the impact of a single trade on the price.

- Φ^N (influence of N on λ^N) accounts for the influence of past changes in price on future changes in price (due to cancel limit orders only since price changes due to market orders are explicitly taken into account by Φ^I.
- Φ^F (influence of N on λ^T) accounts for the feedback influence of the price moves on the trades.

As shown in [4] there are symmetries between the bid-ask sides for trades and up-down directions for price jumps. The aforementioned matrices can therefore be naturally written as:

$$\Phi_t^T = \begin{pmatrix} \Phi_t^{T,s} & \Phi_t^{T,c} \\ \Phi_t^{T,c} & \Phi_t^{T,s} \end{pmatrix}, \Phi_t^I = \begin{pmatrix} \Phi_t^{I,s} & \Phi_t^{I,c} \\ \Phi_t^{I,c} & \Phi_t^{I,s} \end{pmatrix}$$

$$\Phi_t^N = \begin{pmatrix} \Phi_t^{N,s} & \Phi_t^{N,c} \\ \Phi_t^{N,c} & \Phi_t^{N,s} \end{pmatrix}, \Phi_t^F = \begin{pmatrix} \Phi_t^{F,s} & \Phi_t^{F,c} \\ \Phi_t^{F,c} & \Phi_t^{F,s} \end{pmatrix}$$

In order to illustrate the 4-dimensional process we will display the price path by defining the following quantity:

$$X_t = N_t^+ - N_t^-$$

and the cumulative trade process path:

$$U_t = T_t^+ - T_t^-$$

As shown in [4], these processes converge to correlated Brownian Motions. This can be seen in the following figure:

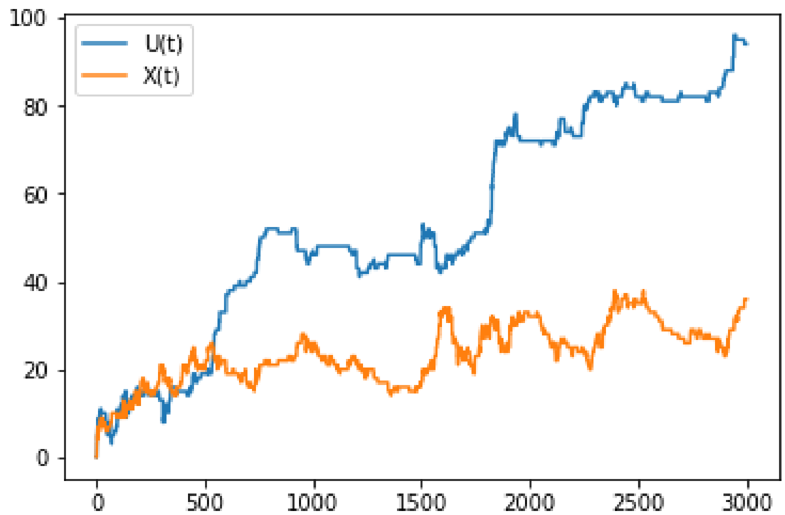

Figure 8: Paths for the cumulative trade and the price processes for the Eurostoxx data

4.3.2 Estimation using the EM algorithm

Using the EMV algorithm described above, we can estimate the Hawkes' kernels. We can first see that the kernel estimation shows that the symmetry hypothesis holds. This is easier to see when we plot the kernel norms.

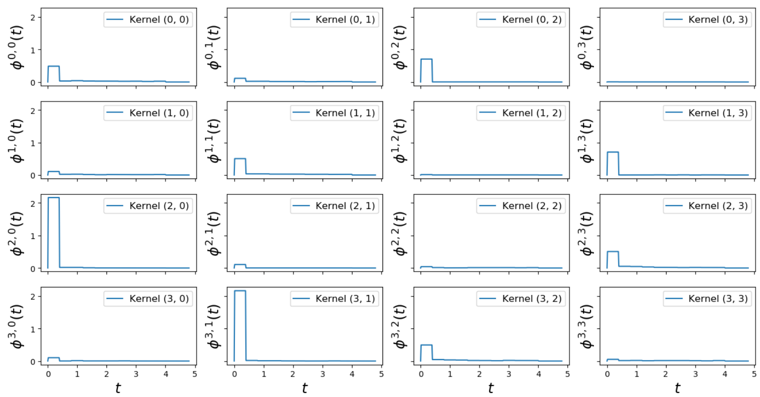

Figure 9: EM estimation of the Hawkes' kernels of Eurostoxx

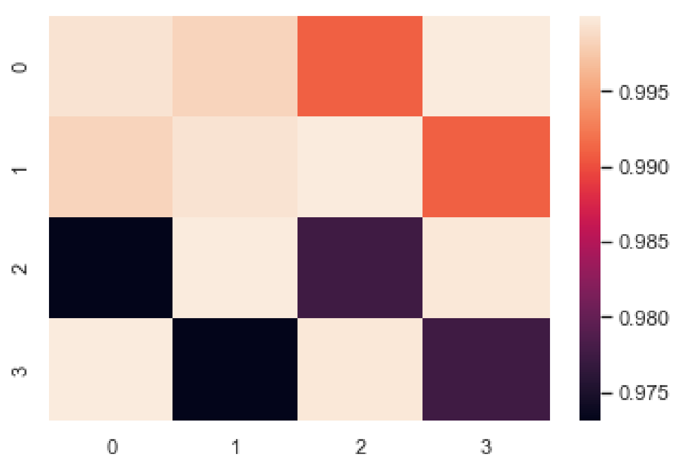

Figure 10: EM estimation of the Hawkes' kernels of Eurostoxx

The results are in accord with the findings in [4]. We can also use the non-parametric estimation method described in [5] while imposing the symmetry conditions to have better results. The kernels using this method are as follows:

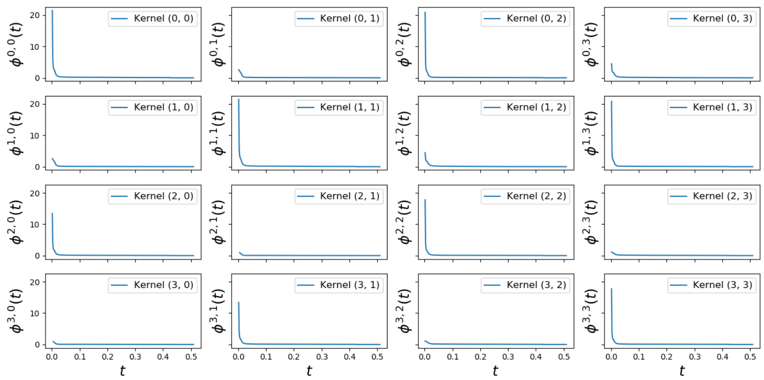

Figure 11: Estimation of the Hawkes' kernels of Eurostoxx using the method described in [5]

The method described in [5] has been proven to be reliable for several different kernels. We therefore compute the intensity by resimulating a process using the same event timestamps as in the data.

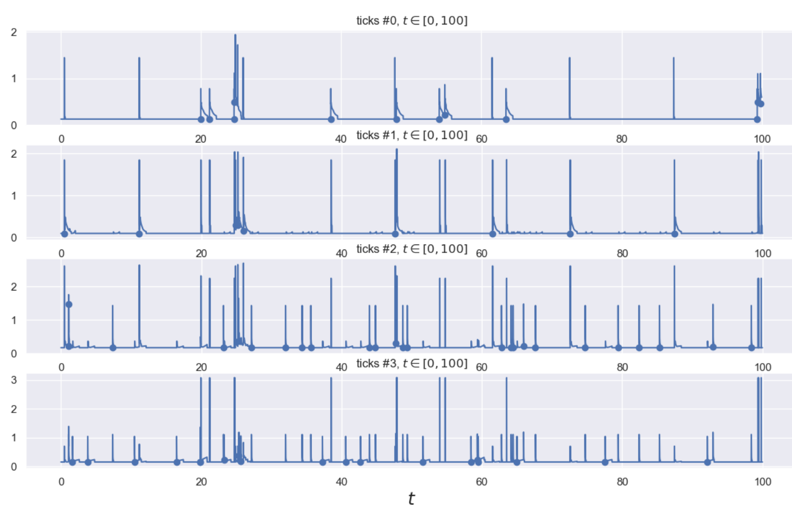

Figure 12: Intensity of the estimated Hawkes' process of Eurostoxx using the method described in [5]

4.3.3 Goodness of fit

As described in [15], we can use the Quantile-Quantile plots to have an idea about the goodness of fit of the model. But QQ-plots can only be plotted for unidimensional data. Unfortunately the use of the Kolmogorov-Smirnov statistical test (as indicated in [15]) is also impossible because the test needs to be adapted for multivariate processes. Using [19] a 2-dimensional version of the test has been implemented in python, but the multivariate version might be presented in an forthcoming work using [8] and [16].

4.4 Market Impact for the non-labelled data

Most markets do not provide labelled data. The order flows are anonymous and we thus cannot detect meta-orders. [4] leverages on the model built for labelled data to define one for non-labelled data. The idea is to numerically estimate the so-called "response function" R_t that is defined as the variation of the price from time 0 to time t knowing there was a trade at time 0. Thus it can be written as:

$$R_t = \mathbf{E}[N_t^+ - N_t^- | dT_0^+ = 1], \text{ for t} \geq 0$$

It has been proven in [4] that in the case of an impulsive impact kernel, R_t can be rewritten as:

$$R_t = I(1 - \int_0^t \Delta\xi'_u \, du) \, \forall t > 0$$

where $\hat{\Delta\xi}'_z$ is a a function of the Laplace transform of the estimated kernel imbalance and the mean intensity.

This first figure has been obtained by using the general formula in [4] defined by:

Theorem 6. Response formula
By defining:

$$R_t^s = \int_0^t \mathbf{E}[dN_u^+ | dT_0^+ = 1] - \Lambda^N \, du$$

$$R_t^s = \int_0^t \mathbf{E}[dN_u^- | dT_0 + = 1] - \Lambda^N \, du$$

and $dR_t^s = r_t^s dt$, $dR_t^c = r_t^c dt$. *Then:*

$$\begin{pmatrix} \hat{r}_z^s \\ \hat{r}_z^c \end{pmatrix} = (\mathbf{I} + \hat{D})(\mathbf{I} + \hat{D}^*)(\hat{E}^I)^* / \Lambda^T$$

where the $(.)^*$ *designates the conjugate,* $\hat{(.)}$ *the Laplace transform and* D *is as defined in equation (18) of [4].*

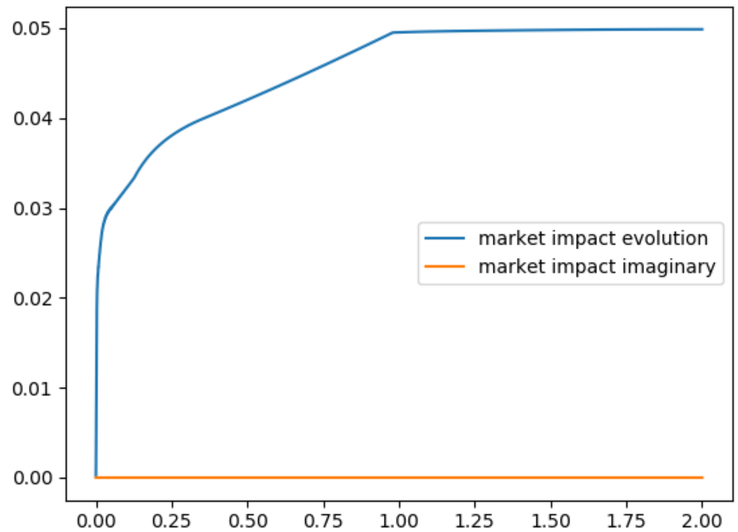

Figure 13: Response R_t plot using the general formula

Once the parameters estimated, [4] shows that it is possible to obtain the shape of the market impact of some particular order as shown in section 7.4. It is worth noting that for a certain range of Laplace parameters z, we have:

$$\hat{MI}_z = z^{\nu} \hat{A}_z$$

where $\hat{MI}_z$ is the Laplace transform of the maret impact and A_z is the point process corresponding to trade arrivals.

Under anonymous data conditions, the market impact behavior reads:

$$MI_t \approx t^{-\nu} \text{if } t >> T$$

$$MI_t \approx t^{1-\nu} \text{if } t << T$$

In this case the impact profile is as follows:

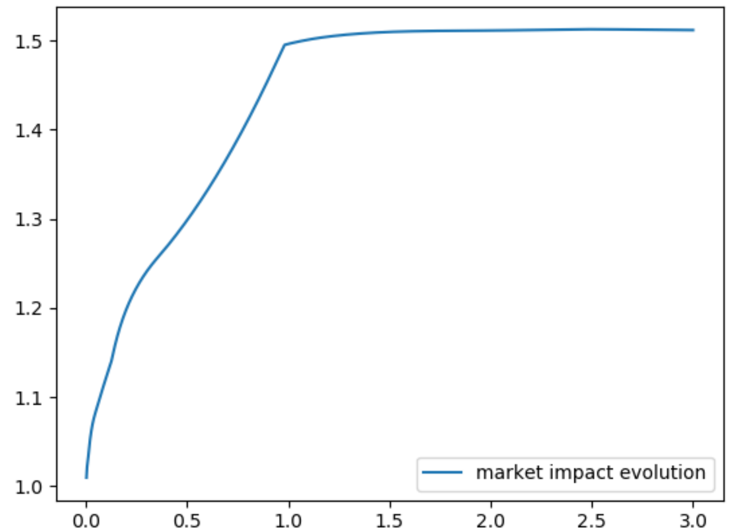

Figure 14: Market impact for the anonymous Eurostoxx data using the formulas presented in [4]

5 Conclusion

Through this paper we have reviewed the theory behind self-exciting point processes and the several state-of-the-art models used for market impact profile estimation. The model developed in [4] accounts for the market price microstructure and the market impact, and has the advantage of allowing analytically closed formulas. Our initial aim was to compare this model with other cutting-edge models such as in [1] or [9]. Unfortunately the data was unlabelled and the identification of meta-orders was beyond the scope of this work. Models such us the square-root models or simple computations (e.g. VWAP) were therefore impossible for us. The kernel estimates weren't back-tested by plotting the Quantile-Quantile plots or through Kolmogorov-Smirnov statistics, which means there was no explicit validation of our own implementation of the Hawkes' model. This should be corrected in future works.

We thus decided to try extending the model following the propositions in [4]. The extension of this model consists in the introduction of two additional point processes that will account for the limit orders. The kernel estimation for this new point process has been done and is as follows:

While interesting, this extended framework is unfortunately too intricate to fit within the scope of this project. The limit order book theory is a very thriving field with a plethora of theories (a review can be found in [13]) and a serious

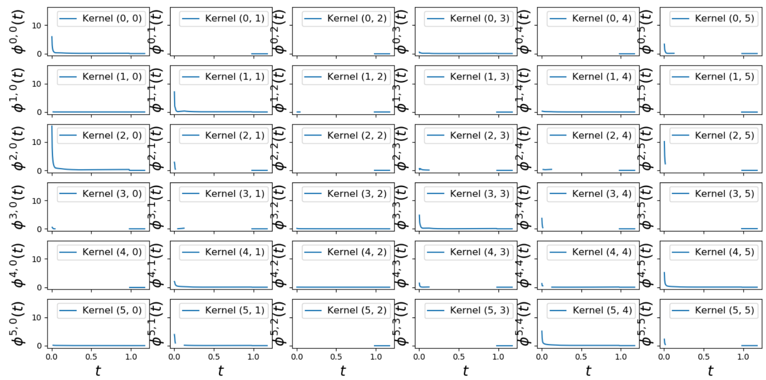

Figure 15: Kernel estimation on the Eurostoxx data for a 6-dimensional Hawkes' process taking into account limit orders

implementation might come in forthcoming works. However, the kernel estimation method could be bettered by leveraging on statistical and deep learning methods to estimate the kernel functions. Using neural networks to learn and correctly estimate the various kernels $\Phi^{?,?}$ might prove to be more accurate.

References

[1] A. Fruth A. Alfonsi and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quantitative Finance*, 10(2), 2010.

[2] J.P. Bouchaud B. Tóth, Z. Eisler. The square-root impact law also holds for option markets. *Wilmott*, 2016(85):430–453, 2016.

[3] J.P. Bouchaud. Price impact. *Encyclopedia of Quantitative Finance*, 2017.

[4] J.-F. Muzy E. Bacry. Hawkes model for price and trades high-frequency dynamics. 2013.

[5] J.-F. Muzy E. Bacry, T. Jaisson. Estimation of slowly decreasing hawkes kernels: Application to high frequency order book modelling. 2014.

[6] P. Brantingham E. Lewis, G. Mohler and A. Bertozzi. Self-exciting point process models of civilian deaths in iraq.

[7] E. Bacry et al. tick : a python library for statistical learning, with a particular emphasis on time-dependent modeling. *ArXiv e-prints (juil. 2017)*, eprint : 1707.03003.

[8] G. Fasano A. Franceschini. A multidimensional version of the kolmogorov–smirnov test. *Monthly Notices of the Royal Astronomical Society*, vol. 225(1), 1987.

[9] J. Gatheral. No-dynamic-arbitrage and market impact. *Quantitative Finance*, 7(10):749–759, 2010.

[10] P.J. Brantingham F.P. Schoenberg G.O. Mohler, M.B. Short and G.E. Tita. Self-exciting point process modeling of crime.

[11] A.G. Hawkes. Point spectra of some mutually exciting point processes. *Biometrika*, 1971(58):83–90, 1971.

[12] P. Hewlett. Clustering of order arrivals, price impact and trade path optimization.

[13] I. Lemhadri. Market impact in a latent order book. *HAL*, (¡hal-01711192v2¿), 2018.

[14] P. A. W. Lewis and G. S. Shedler. Simulation of nonhomogeneous poisson processes by thinning. 1979.

[15] F. Lorenzen. Analysis of order clustering using high frequency data: A point process approach.

[16] J.D. Loudin and H.E. Miettinen. A multivariate method for comparing n-dimensional distributions. *PHYSTAT2003, SLAC, Stanford, California*, 2003.

[17] Y. Ogata. On lewis' simulation method for point processes. *Ieee Transactions On Information Theory*, (27):23–31, 1981.

[18] V Rivoirard P. Reynaud-Bouret, C. Tuleau-Malot and F. Grammont. Spike trains as (in)homogeneous poisson processes or hawkes processes: nonparametric adaptive estimation and goodness-of-fit tests.

[19] J. A. Peacock. Two-dimensional goodness-of-fit testing in astronomy. *Monthly Notices of the Royal Astronomical Society*, vol. 202:615–627, 1983.

[20] R.D. Peng. Applications of multi-dimensional point process methodology to wildfire hazard assessment.