An Introduction to Hawkes Processes with Applications to Finance

Sophie LARUELLE

LAMA-UPEC

April 10, 2026

Outline

Outline

Simple Point Processes

Definition. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be some probability space. Let $(\tau_i)_{i\geq 1}$ be a sequence of non-negative random variables such that $\forall i \geq 1$, $\tau_i < \tau_{i+1}$. We call $(\tau_i)_{i\geq 1}$ a (simple) point process on $\mathbb{R}_+$.

Simple Point Processes

Definition. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be some probability space. Let $(\tau_i)_{i\geq 1}$ be a sequence of non-negative random variables such that $\forall i \geq 1$, $\tau_i < \tau_{i+1}$. We call $(\tau_i)_{i\geq 1}$ a (simple) point process on $\mathbb{R}_+$.

In particular, the variables τ_i can represent the times of occurrence of transactions, or arrival of limit orders in an order book, etc.

Simple Point Processes

Definition. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be some probability space. Let $(\tau_i)_{i\geq 1}$ be a sequence of non-negative random variables such that $\forall i \geq 1$, $\tau_i < \tau_{i+1}$. We call $(\tau_i)_{i\geq 1}$ a (simple) point process on $\mathbb{R}_+$.

In particular, the variables τ_i can represent the times of occurrence of transactions, or arrival of limit orders in an order book, etc.

We start counting events with index 1. If needed, we will assume that $\tau_0 = 0$.

Counting Processes and Duration

Definition (Counting process). Let $(\tau_i)_{i\geq 1}$ be a point process. The right-continuous process

$$N_t = \sum_{i\geq 1} \mathbb{1}_{\tau_i \leq t}$$

is called the counting process associated with $(\tau_i)_{i\geq 1}$.

Counting Processes and Duration

Definition (Counting process). Let $(\tau_i)_{i\geq 1}$ be a point process. The right-continuous process

$$N_t = \sum_{i\geq 1} \mathbb{1}_{\tau_i \leq t}$$

is called the counting process associated with $(\tau_i)_{i\geq 1}$.

Definition (Duration). The process $(\Delta\tau_i)_{i\geq 1}$ defined by

$$\forall i \geq 1, \quad \Delta\tau_i = \tau_i - \tau_{i-1}$$

is called the duration process associated with $(\tau_i)_{i\geq 1}$.

Representation of a simple point process

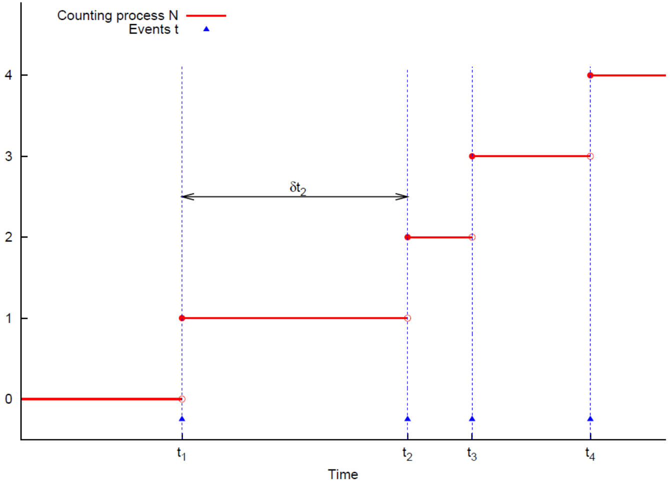


Figure: Illustration of a simple point process: events, counting process and duration process.

Intensity Process

Definition. Let N be a counting process adapted to a filtration $\mathcal{F}_t$. The left-continuous intensity process is defined as

$$\lambda(t|\mathcal{F}_t) = \lim_{h\downarrow 0} \mathbb{E}\left[\frac{N_{t+h} - N_t}{h} \mid \mathcal{F}_t\right],$$

or equivalently

$$\lambda(t|\mathcal{F}_t) = \lim_{h\downarrow 0} \frac{1}{h}\mathbb{P}\left(N_{t+h} - N_t > 0|\mathcal{F}_t\right).$$

Intensity Process

Definition. Let N be a counting process adapted to a filtration $\mathcal{F}_t$. The left-continuous intensity process is defined as

$$\lambda(t|\mathcal{F}_t) = \lim_{h \downarrow 0} \mathbb{E}\left[\frac{N_{t+h} - N_t}{h} \mid \mathcal{F}_t\right],$$

or equivalently

$$\lambda(t|\mathcal{F}_t) = \lim_{h \downarrow 0} \frac{1}{h} \mathbb{P}\left(N_{t+h} - N_t > 0 | \mathcal{F}_t\right).$$

Intensity depends on the choice of filtration, but we will always assume that the filtration used is the natural one for the process N, denoted $\mathcal{F}_t^N$. We will therefore write $\lambda(t)$ instead of $\lambda(t|\mathcal{F}_t^N)$.

Example: Homogeneous Poisson Process

Let $\lambda \in \mathbb{R}_+^*$. A Poisson process with constant rate λ is a point process defined by

$$\begin{aligned}
\mathbb{P}(N_{t+h} - N_t = 1 | \mathcal{F}_t) &= \lambda h + o(h), \quad h \to 0^+, \\
\mathbb{P}(N_{t+h} - N_t > 1 | \mathcal{F}_t) &= o(h), \quad h \to 0^+.
\end{aligned}$$

Example: Homogeneous Poisson Process

Let $\lambda \in \mathbb{R}_+^*$. A Poisson process with constant rate λ is a point process defined by

$$\begin{aligned}
\mathbb{P}(N_{t+h} - N_t = 1 | \mathcal{F}_t) &= \lambda h + o(h), \quad h \to 0^+, \\
\mathbb{P}(N_{t+h} - N_t > 1 | \mathcal{F}_t) &= o(h), \quad h \to 0^+.
\end{aligned}$$

- The intensity does not depend on the history of the process N, and the probability of occurrence of an event in $(t, t+h]$ is independent of $\mathcal{F}_t$.

Example: Homogeneous Poisson Process

Let $\lambda \in \mathbb{R}_+^*$. A Poisson process with constant rate λ is a point process defined by

$$\begin{aligned}\mathbb{P}(N_{t+h} - N_t = 1 | \mathcal{F}_t) &= \lambda h + o(h), \quad h \to 0^+, \\ \mathbb{P}(N_{t+h} - N_t > 1 | \mathcal{F}_t) &= o(h), \quad h \to 0^+.\end{aligned}$$

- The intensity does not depend on the history of the process N, and the probability of occurrence of an event in $(t, t+h]$ is independent of $\mathcal{F}_t$.
- Durations $(\Delta\tau_i)_{i\geq 1}$ of a homogeneous Poisson process are independent and identically distributed (i.i.d.) according to an exponential distribution with parameter λ.

Example: Homogeneous Poisson Process

Let $\lambda \in \mathbb{R}_+^*$. A Poisson process with constant rate λ is a point process defined by

$$\begin{aligned}
\mathbb{P}(N_{t+h} - N_t = 1 | \mathcal{F}_t) &= \lambda h + o(h), \quad h \to 0^+, \\
\mathbb{P}(N_{t+h} - N_t > 1 | \mathcal{F}_t) &= o(h), \quad h \to 0^+.
\end{aligned}$$

- The intensity does not depend on the history of the process N, and the probability of occurrence of an event in $(t, t+h]$ is independent of $\mathcal{F}_t$.
- Durations $(\Delta\tau_i)_{i\geq 1}$ of a homogeneous Poisson process are independent and identically distributed (i.i.d.) according to an exponential distribution with parameter λ.
- This process is Markovian and very often used in queueing theory (when the consumption and the service are assumed to be independent).

Stochastic Time Change

Definition. The integrated intensity function Λ is defined as

$$\forall i \geq 1, \quad \Lambda(\tau_{i-1}, \tau_i) = \int_{\tau_{i-1}}^{\tau_i} \lambda(s) ds.$$

Stochastic Time Change

Definition. The integrated intensity function Λ is defined as

$$\forall i \geq 1, \quad \Lambda(\tau_{i-1}, \tau_i) = \int_{\tau_{i-1}}^{\tau_i} \lambda(s) ds.$$

The following theorem illustrates the fact that any point process comes down to a Poisson process up to a time change. This result is the equivalent of the Dubins-Schwarz theorem for martingales, namely every continuous local $(\mathcal{F}_t)$-martingale M is of the form $M_t = B_{\langle M \rangle_t}$, $t \geq 0$, where $(B_s)_{s \geq 0}$ is a $\mathcal{F}_{\tau_s}$-Brownian motion (with $\tau_s = \inf\{t : \langle M \rangle_t > s\}$).

Stochastic Time Change

Definition. The integrated intensity function Λ is defined as

$$\forall i \geq 1, \quad \Lambda(\tau_{i-1}, \tau_i) = \int_{\tau_{i-1}}^{\tau_i} \lambda(s)ds.$$

The following theorem illustrates the fact that any point process comes down to a Poisson process up to a time change. This result is the equivalent of the Dubins-Schwarz theorem for martingales, namely every continuous local $(\mathcal{F}_t)$-martingale M is of the form $M_t = B_{\langle M \rangle_t}$, $t \geq 0$, where $(B_s)_{s\geq 0}$ is a $\mathcal{F}_{\tau_s}$-Brownian motion (with $\tau_s = \inf\{t : \langle M \rangle_t > s\}$).

Theorem (Time change theorem). Let N be a counting process on $\mathbb{R}_+$ such that $\int_0^\infty \lambda(s)ds = 1$. Let t_τ be the stopping time defined by

$$\int_0^{t_\tau} \lambda(s)ds = \tau.$$

Then the process $\tilde{N}_\tau = N_{t_\tau}$ is a homogeneous Poisson process with constant intensity $\lambda = 1$.

Outline

Outline

Definition of a linear self-exciting process

When the data that we want to model are not Markovian, namely they have some clusters, we do not use Poisson processes to model them. Thus we introduce a new type of processes recently used in microstructure and called Hawkes processes.

Definition of a linear self-exciting process

When the data that we want to model are not Markovian, namely they have some clusters, we do not use Poisson processes to model them. Thus we introduce a new type of processes recently used in microstructure and called Hawkes processes.

Definition. A general definition for a linear self-exciting process N reads

$$\lambda(t) = \lambda_0(t) + \int_{-\infty}^{t} \varphi(t-s)dN_s = \lambda_0(t) + \sum_{\tau_i < t} \varphi(t-\tau_i),$$

where $\lambda_0 : \mathbb{R} \mapsto \mathbb{R}_+$ is a deterministic base intensity and $\varphi : \mathbb{R}_+ \mapsto \mathbb{R}_+$ expresses the positive influence of the past events τ_i on the current value of the intensity process (φ is called the kernel).

Simple Hawkes process considered here

Hawkes (1971) proposes an exponential kernel $\varphi(t) = \sum_{j=1}^{P} \alpha_j e^{-\beta_j t} \mathbb{1}_{\mathbb{R}_+}(t)$, so that the intensity of the model becomes

$$\lambda(t) = \lambda_0(t) + \int_0^t \sum_{j=1}^{P} \alpha_j e^{-\beta_j (t-s)} dN_s = \lambda_0(t) + \sum_{\tau_i < t} \sum_{j=1}^{P} \alpha_j e^{-\beta_j (t-\tau_i)}.$$

Simple Hawkes process considered here

Hawkes (1971) proposes an exponential kernel $\varphi(t) = \sum_{j=1}^{P} \alpha_j e^{-\beta_j t} \mathbb{1}_{\mathbb{R}_+}(t)$, so that the intensity of the model becomes

$$\lambda(t) = \lambda_0(t) + \int_0^t \sum_{j=1}^{P} \alpha_j e^{-\beta_j (t-s)} dN_s = \lambda_0(t) + \sum_{\tau_i < t} \sum_{j=1}^{P} \alpha_j e^{-\beta_j (t-\tau_i)}.$$

The simplest version with $P = 1$ and $\lambda_0(t)$ constant is defined as

$$\lambda(t) = \lambda_0 + \int_0^t \alpha e^{-\beta (t-s)} dN_s = \lambda_0 + \sum_{\tau_i < t} \alpha e^{-\beta (t-\tau_i)}.$$

Sample path of a 1D-Hawkes process

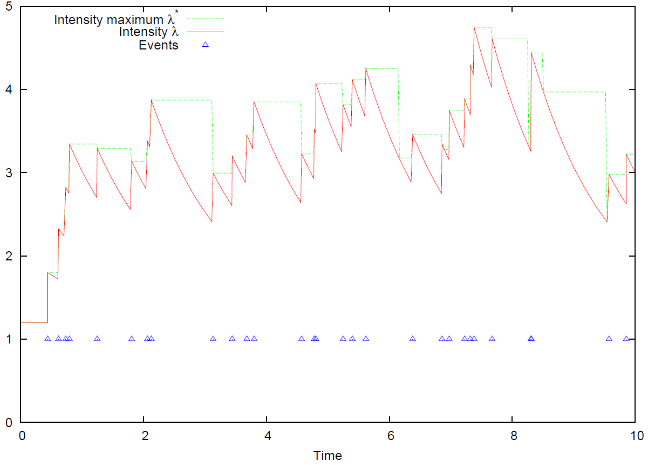


Figure: Simulation of a one-dimensional Hawkes process with parameters $P = 1$, $\lambda_0 = 1.2$, $\alpha = 0.6$, $\beta = 0.8$.

Stationarity : General case

Assuming stationarity gives $\mathbb{E}[\lambda(t)] = \mu$ constant. Thus,

$$\mu \quad = \quad \mathbb{E}[\lambda(t)] = \mathbb{E}\left[\lambda_0 + \int_{-\infty}^{t} \varphi(t-s)dN_s\right]$$

Stationarity : General case

Assuming stationarity gives $\mathbb{E}[\lambda(t)] = \mu$ constant. Thus,

$$\begin{aligned}
\mu &= \mathbb{E}[\lambda(t)] = \mathbb{E}\left[\lambda_0 + \int_{-\infty}^{t} \varphi(t-s)dN_s\right] \\
&= \lambda_0 + \mathbb{E}\left[\int_{-\infty}^{t} \varphi(t-s)\lambda(s)ds\right]
\end{aligned}$$

Stationarity : General case

Assuming stationarity gives $\mathbb{E}[\lambda(t)] = \mu$ constant. Thus,

$$\begin{aligned}
\mu &= \mathbb{E}[\lambda(t)] = \mathbb{E}\left[\lambda_0 + \int_{-\infty}^{t} \varphi(t-s)dN_s\right] \\
&= \lambda_0 + \mathbb{E}\left[\int_{-\infty}^{t} \varphi(t-s)\lambda(s)ds\right] \\
&= \lambda_0 + \int_{-\infty}^{t} \varphi(t-s)\mu ds
\end{aligned}$$

Stationarity : General case

Assuming stationarity gives $\mathbb{E}[\lambda(t)] = \mu$ constant. Thus,

$$\begin{aligned}
\mu &= \mathbb{E}[\lambda(t)] = \mathbb{E}\left[\lambda_0 + \int_{-\infty}^{t} \varphi(t-s)dN_s\right] \\
&= \lambda_0 + \mathbb{E}\left[\int_{-\infty}^{t} \varphi(t-s)\lambda(s)ds\right] \\
&= \lambda_0 + \int_{-\infty}^{t} \varphi(t-s)\mu ds \\
&= \lambda_0 + \mu \int_{0}^{\infty} \varphi(u)du
\end{aligned}$$

Stationarity : General case

Assuming stationarity gives $\mathbb{E}[\lambda(t)] = \mu$ constant. Thus,

$$\begin{aligned}
\mu &= \mathbb{E}[\lambda(t)] = \mathbb{E}\left[\lambda_0 + \int_{-\infty}^{t} \varphi(t-s)dN_s\right] \\
&= \lambda_0 + \mathbb{E}\left[\int_{-\infty}^{t} \varphi(t-s)\lambda(s)ds\right] \\
&= \lambda_0 + \int_{-\infty}^{t} \varphi(t-s)\mu ds \\
&= \lambda_0 + \mu \int_{0}^{\infty} \varphi(u)du
\end{aligned}$$

which gives

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

Stationarity: Exponential kernel

Stationarity condition for a 1D-Hawkes process.

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

Stationarity: Exponential kernel

Stationarity condition for a 1D-Hawkes process.

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

Average intensity of a stationary process : for the one-dimensional Hawkes process with $P = 1$, the unconditional expected value of the intensity process is given by

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

Outline

Thinning procedure

Lewis & Shedler (1979) proposes a "thinning procedure" that allows the simulation of a point process with bounded intensity.

Thinning procedure

Lewis & Shedler (1979) proposes a “thinning procedure” that allows the simulation of a point process with bounded intensity.

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 $\mu_0^* = \int_0^{T_0} \lambda^*(s)ds$.

Thinning procedure

Lewis & Shedler (1979) proposes a “thinning procedure” that allows the simulation of a point process with bounded intensity.

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 $\tau_1^*, \ldots, \tau_{N_{T_0}^*}^*$ be the points of the process in the interval $(0, T_0]$.
Suppose that for $0 \leq t \leq T_0$, $\lambda(t) \leq \lambda^*(t)$.

Thinning procedure

Lewis & Shedler (1979) proposes a "thinning procedure" that allows the simulation of a point process with bounded intensity.

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 $\tau^*_1, \ldots, \tau^*_{N^*_{T_0}}$ be the points of the process in the interval $(0, T_0]$. Suppose that for $0 \leq t \leq T_0$, $\lambda(t) \leq \lambda^*(t)$.

For $i = 1, \ldots, N^*(T_0)$, delete the points τ^*_i with probability $1 - \dfrac{\lambda(\tau^*_i)}{\lambda^*(\tau^*_i)}$.

Thinning procedure

Lewis & Shedler (1979) proposes a "thinning procedure" that allows the simulation of a point process with bounded intensity.

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 $\tau^*_1, \dots, \tau^*_{N^*_{T_0}}$ be the points of the process in the interval $(0, T_0]$. Suppose that for $0 \leq t \leq T_0$, $\lambda(t) \leq \lambda^*(t)$.

For $i = 1, \dots, N^*(T_0)$, delete the points τ^*_i with probability $1 - \dfrac{\lambda(\tau^*_i)}{\lambda^*(\tau^*_i)}$.

Then the remaining points form a non-homogeneous Poisson process $(N_t)_{t\geq 0}$ with rate function $\lambda(t)$ in the interval $(0, T_0]$.

Simulation of a Hawkes process

Ogata (1981) proposes an algorithm for the simulation of Hawkes processes. Let us denote $\mathcal{U}_{[0,1]}$ 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 $P = 1$.

Simulation of a Hawkes process

Ogata (1981) proposes an algorithm for the simulation of Hawkes processes. Let us denote $\mathcal{U}_{[0,1]}$ 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 $P = 1$.

1. **Initialization** : Set $\lambda^* \leftarrow \lambda_0$, $n \leftarrow 1$.

Simulation of a Hawkes process

Ogata (1981) proposes an algorithm for the simulation of Hawkes processes. Let us denote $\mathcal{U}_{[0,1]}$ 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 $P = 1$.

1. **Initialization** : Set $\lambda^* \leftarrow \lambda_0$, $n \leftarrow 1$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{\lambda^*} \ln U$.

Simulation of a Hawkes process

Ogata (1981) proposes an algorithm for the simulation of Hawkes processes. Let us denote $\mathcal{U}_{[0,1]}$ 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 $P = 1$.

1. **Initialization** : Set $\lambda^* \leftarrow \lambda_0$, $n \leftarrow 1$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{\lambda^*} \ln U$.
   If $s \leq T$, **Then** $\tau_1 \leftarrow s$, **Else** go to last step.

Simulation of a Hawkes process

Ogata (1981) proposes an algorithm for the simulation of Hawkes processes. Let us denote $\mathcal{U}_{[0,1]}$ 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 $P=1$.

1. **Initialization** : Set $\lambda^* \leftarrow \lambda_0$, $n \leftarrow 1$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{\lambda^*}\ln U$.
   If $s \leq T$, **Then** $\tau_1 \leftarrow s$, **Else** go to last step.
3. **General routine** : Set $n \leftarrow n+1$.

Simulation of a Hawkes process

Ogata (1981) proposes an algorithm for the simulation of Hawkes processes. Let us denote $\mathcal{U}_{[0,1]}$ 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 $P = 1$.

1. **Initialization** : Set $\lambda^* \leftarrow \lambda_0$, $n \leftarrow 1$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{\lambda^*} \ln U$.
   If $s \leq T$, **Then** $\tau_1 \leftarrow s$, **Else** go to last step.
3. **General routine** : Set $n \leftarrow n + 1$.
   a. **Update maximum intensity** : Set $\lambda^* \leftarrow \lambda(\tau_{n-1}) + \alpha$.
   λ^* exhibits a jump of size α as an event has just occurred. λ being left-continuous, this jump is not counted in $\lambda(\tau_{n-1})$, hence the explicit addition.

Simulation of a Hawkes process

Ogata (1981) proposes an algorithm for the simulation of Hawkes processes. Let us denote $\mathcal{U}_{[0,1]}$ 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 $P = 1$.

1. **Initialization** : Set $\lambda^* \leftarrow \lambda_0$, $n \leftarrow 1$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{\lambda^*} \ln U$.
   If $s \leq T$, **Then** $\tau_1 \leftarrow s$, **Else** go to last step.
3. **General routine** : Set $n \leftarrow n + 1$.
   a. **Update maximum intensity** : Set $\lambda^* \leftarrow \lambda(\tau_{n-1}) + \alpha$.
   λ^* exhibits a jump of size α as an event has just occurred. λ being left-continuous, this jump is not counted in $\lambda(\tau_{n-1})$, hence the explicit addition.
   b. **New event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow s - \frac{1}{\lambda^*} \ln U$.
   If $s \geq T$, **Then** go to the last step.

Simulation of a Hawkes process

Ogata (1981) proposes an algorithm for the simulation of Hawkes processes. Let us denote $\mathcal{U}_{[0,1]}$ 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 $P = 1$.

1. **Initialization** : Set $\lambda^* \leftarrow \lambda_0$, $n \leftarrow 1$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{\lambda^*} \ln U$.
   If $s \leq T$, **Then** $\tau_1 \leftarrow s$, **Else** go to last step.
3. **General routine** : Set $n \leftarrow n+1$.
   a. **Update maximum intensity** : Set $\lambda^* \leftarrow \lambda(\tau_{n-1}) + \alpha$.
   λ^* exhibits a jump of size α as an event has just occurred. λ being left-continuous, this jump is not counted in $\lambda(\tau_{n-1})$, hence the explicit addition.
   b. **New event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow s - \frac{1}{\lambda^*} \ln U$.
   If $s \geq T$, **Then** go to the last step.
   c. **Rejection test** : Generate $D \sim \mathcal{U}_{[0,1]}$.
   If $D \leq \frac{\lambda(s)}{\lambda^*}$, **Then** $\tau_n \leftarrow s$ and go through the general routine again, **Else** update $\lambda^* \leftarrow \lambda(s)$ and try a new date at step b. of the general routine.

Simulation of a Hawkes process

Ogata (1981) proposes an algorithm for the simulation of Hawkes processes. Let us denote $\mathcal{U}_{[0,1]}$ 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 $P=1$.

1. **Initialization** : Set $\lambda^* \leftarrow \lambda_0$, $n \leftarrow 1$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{\lambda^*} \ln U$.
   If $s \leq T$, **Then** $\tau_1 \leftarrow s$, **Else** go to last step.
3. **General routine** : Set $n \leftarrow n+1$.
   a. **Update maximum intensity** : Set $\lambda^* \leftarrow \lambda(\tau_{n-1}) + \alpha$.
   λ^* exhibits a jump of size α as an event has just occurred. λ being left-continuous, this jump is not counted in $\lambda(\tau_{n-1})$, hence the explicit addition.
   b. **New event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow s - \frac{1}{\lambda^*} \ln U$.
   If $s \geq T$, **Then** go to the last step.
   c. **Rejection test** : Generate $D \sim \mathcal{U}_{[0,1]}$.
   If $D \leq \frac{\lambda(s)}{\lambda^*}$, **Then** $\tau_n \leftarrow s$ and go through the general routine again, **Else** update $\lambda^* \leftarrow \lambda(s)$ and try a new date at step b. of the general routine.
4. **Output**: Retrieve the simulated process $(\tau_n)_n$ on $[0,T]$.

Example of simulations

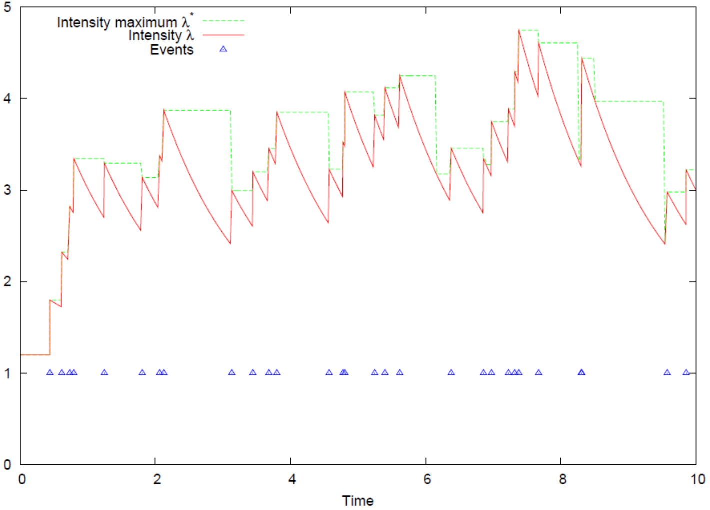


Figure: Simulation of a one-dimensional Hawkes process with parameters $P = 1$, $\lambda_0 = 1.2$, $\alpha_1 = 0.6$, $\beta_1 = 0.8$.

Testing the simulated process

For any consecutive events τ_{i-1} and τ_i

$$\Lambda(\tau_{i-1}, \tau_i) \;=\; \int_{\tau_{i-1}}^{\tau_i} \lambda(s)ds =$$

Testing the simulated process

For any consecutive events τ_{i-1} and τ_i

$$\Lambda(\tau_{i-1}, \tau_i) \quad = \quad \int_{\tau_{i-1}}^{\tau_i} \lambda(s)ds = \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k < s} \alpha e^{-\beta(s-\tau_k)} ds$$

Testing the simulated process

For any consecutive events τ_{i-1} and τ_i

$$
\begin{aligned}
\Lambda(\tau_{i-1},\tau_i) &= \int_{\tau_{i-1}}^{\tau_i} \lambda(s)ds = \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k < s} \alpha e^{-\beta(s-\tau_k)} ds \\
&= \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k \leq \tau_{i-1}} \alpha e^{-\beta(s-\tau_k)} ds
\end{aligned}
$$

Testing the simulated process

For any consecutive events τ_{i-1} and τ_i

$$\begin{aligned}
\Lambda(\tau_{i-1},\tau_i) &= \int_{\tau_{i-1}}^{\tau_i} \lambda(s)ds = \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k<s} \alpha e^{-\beta(s-\tau_k)} ds \\
&= \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k\leq\tau_{i-1}} \alpha e^{-\beta(s-\tau_k)} ds \\
&= \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \sum_{\tau_k\leq\tau_{i-1}} \frac{\alpha}{\beta}\left(e^{-\beta(\tau_{i-1}-\tau_k)} - e^{-\beta(\tau_i-\tau_k)}\right).
\end{aligned}$$

Testing the simulated process

For any consecutive events τ_{i-1} and τ_i

$$\begin{aligned}
\Lambda(\tau_{i-1},\tau_i) &= \int_{\tau_{i-1}}^{\tau_i} \lambda(s)ds = \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k<s} \alpha e^{-\beta(s-\tau_k)} ds \\
&= \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k\leq\tau_{i-1}} \alpha e^{-\beta(s-\tau_k)} ds \\
&= \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \sum_{\tau_k\leq\tau_{i-1}} \frac{\alpha}{\beta}\left(e^{-\beta(\tau_{i-1}-\tau_k)} - e^{-\beta(\tau_i-\tau_k)}\right).
\end{aligned}$$

This computation can be simplified with a recursive element. Let us denote

$$A_i = \sum_{\tau_k\leq\tau_i} e^{-\beta(\tau_i-\tau_k)}.$$

Testing the simulated process

For any consecutive events τ_{i-1} and τ_i

$$\begin{aligned}
\Lambda(\tau_{i-1},\tau_i) &= \int_{\tau_{i-1}}^{\tau_i} \lambda(s)ds = \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k<s} \alpha e^{-\beta(s-\tau_k)} ds \\
&= \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k\leq\tau_{i-1}} \alpha e^{-\beta(s-\tau_k)} ds \\
&= \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \sum_{\tau_k\leq\tau_{i-1}} \frac{\alpha}{\beta}\left(e^{-\beta(\tau_{i-1}-\tau_k)} - e^{-\beta(\tau_i-\tau_k)}\right).
\end{aligned}$$

This computation can be simplified with a recursive element. Let us denote

$$A_i = \sum_{\tau_k\leq\tau_i} e^{-\beta(\tau_i-\tau_k)}.$$

We observe that

$$A_i = 1 + e^{-\beta(\tau_i-\tau_{i-1})} \sum_{\tau_k\leq\tau_{i-1}} e^{-\beta(\tau_{i-1}-\tau_k)}$$

Testing the simulated process

For any consecutive events τ_{i-1} and τ_i

$$\begin{aligned}
\Lambda(\tau_{i-1},\tau_i) &= \int_{\tau_{i-1}}^{\tau_i} \lambda(s)ds = \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k<s} \alpha e^{-\beta(s-\tau_k)}ds \\
&= \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \int_{\tau_{i-1}}^{\tau_i} \sum_{\tau_k\le\tau_{i-1}} \alpha e^{-\beta(s-\tau_k)}ds \\
&= \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \sum_{\tau_k\le\tau_{i-1}} \frac{\alpha}{\beta}\left(e^{-\beta(\tau_{i-1}-\tau_k)} - e^{-\beta(\tau_i-\tau_k)}\right).
\end{aligned}$$

This computation can be simplified with a recursive element. Let us denote

$$A_i = \sum_{\tau_k\le\tau_i} e^{-\beta(\tau_i-\tau_k)}.$$

We observe that

$$A_i = 1 + e^{-\beta(\tau_i-\tau_{i-1})} \sum_{\tau_k\le\tau_{i-1}} e^{-\beta(\tau_{i-1}-\tau_k)} = 1 + e^{-\beta(\tau_i-\tau_{i-1})}A_{i-1}.$$

Testing the simulated process: Time change property

Finally, the integrated density can be written for all $i \in \mathbb{N}^*$,

$$\Lambda(\tau_{i-1}, \tau_i) = \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \frac{\alpha}{\beta}\left(1 - e^{-\beta(\tau_i - \tau_{i-1})}\right) A_{i-1},$$

where $(A_i)_{i \geq 0}$ is defined as above with $A_0 = 0$.

Testing the simulated process: Time change property

Finally, the integrated density can be written for all $i \in \mathbb{N}^*$,

$$\Lambda(\tau_{i-1}, \tau_i) = \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s)ds + \frac{\alpha}{\beta}\left(1 - e^{-\beta(\tau_i - \tau_{i-1})}\right) A_{i-1},$$

where $(A_i)_{i \geq 0}$ is defined as above with $A_0 = 0$.

Let define $(\vartheta_i)_{i \geq 0}$ as

$$\vartheta_0 = \int_0^{\tau_0} \lambda(s)ds = \Lambda(0, \tau_0),$$

$$\vartheta_i = \vartheta_{i-1} + \Lambda(\tau_{i-1}, \tau_i).$$

Testing the simulated process: Time change property

Finally, the integrated density can be written for all $i \in \mathbb{N}^*$,

$$\Lambda(\tau_{i-1}, \tau_i) = \int_{\tau_{i-1}}^{\tau_i} \lambda_0(s) ds + \frac{\alpha}{\beta} \left(1 - e^{-\beta(\tau_i - \tau_{i-1})}\right) A_{i-1},$$

where $(A_i)_{i \geq 0}$ is defined as above with $A_0 = 0$.

Let define $(\vartheta_i)_{i \geq 0}$ as

$$\vartheta_0 = \int_0^{\tau_0} \lambda(s) ds = \Lambda(0, \tau_0),$$

$$\vartheta_i = \vartheta_{i-1} + \Lambda(\tau_{i-1}, \tau_i).$$

The durations $\vartheta_i - \vartheta_{i-1} = \Lambda(\tau_{i-1}, \tau_i)$ are exponentially distributed.

Testing the simulated process

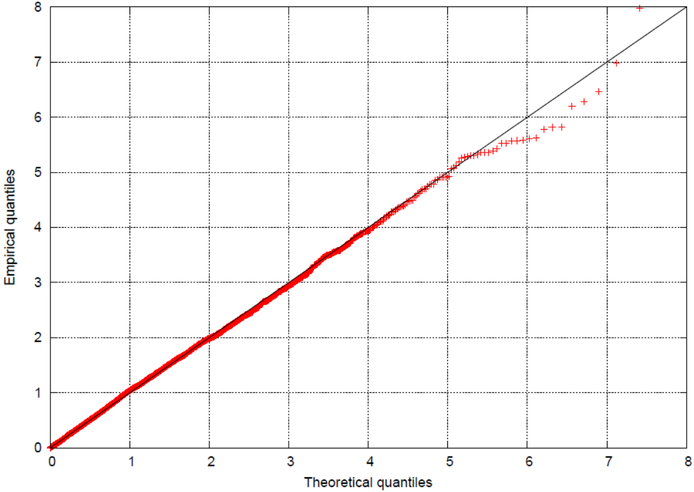


Figure: Quantile plot for one sample of simulated data of a one-dimensional Hawkes process with parameters $P = 1$, $\lambda_0 = 1.2$, $\alpha_1 = 0.6$, $\beta_1 = 0.8$, on an interval $[0, 10000]$.

Outline

Computation of the log-likelihood function

The log-likelihood of a simple point process N with intensity λ is written

$$\ln \mathcal{L}((N_t)_{t\in[0,T]}) = \int_0^T (1-\lambda(s))ds + \int_0^T \ln \lambda(s) dN_s,$$

Computation of the log-likelihood function

The log-likelihood of a simple point process N with intensity λ is written

$$\ln \mathcal{L}((N_t)_{t\in[0,T]}) = \int_0^T (1-\lambda(s))ds + \int_0^T \ln \lambda(s) dN_s,$$

which in the case of a Hawkes model can be explicitly computed as

$$\ln \mathcal{L}((\tau_i)_{i=1,\dots,n}) \quad = \quad \tau_n - \Lambda(0,\tau_n) + \sum_{i=1}^{n} \ln \lambda(\tau_i)$$

Computation of the log-likelihood function

The log-likelihood of a simple point process N with intensity λ is written

$$\ln \mathcal{L}((N_t)_{t\in[0,T]}) = \int_0^T (1-\lambda(s))ds + \int_0^T \ln \lambda(s) dN_s,$$

which in the case of a Hawkes model can be explicitly computed as

$$\begin{aligned} \ln \mathcal{L}((\tau_i)_{i=1,\dots,n}) &= \tau_n - \Lambda(0,\tau_n) + \sum_{i=1}^{n} \ln \lambda(\tau_i) \\ &= \tau_n - \Lambda(0,\tau_n) + \sum_{i=1}^{n} \ln \left(\lambda_0(\tau_i) + \sum_{k=1}^{i-1} \alpha e^{-\beta(\tau_i-\tau_k)} \right). \end{aligned}$$

Computation of the log-likelihood function

The log-likelihood of a simple point process N with intensity λ is written

$$\ln \mathcal{L}((N_t)_{t\in[0,T]}) = \int_0^T (1-\lambda(s))ds + \int_0^T \ln \lambda(s) dN_s,$$

which in the case of a Hawkes model can be explicitly computed as

$$\begin{aligned}\ln \mathcal{L}((\tau_i)_{i=1,\dots,n}) &= \tau_n - \Lambda(0,\tau_n) + \sum_{i=1}^{n} \ln \lambda(\tau_i) \\ &= \tau_n - \Lambda(0,\tau_n) + \sum_{i=1}^{n} \ln\left(\lambda_0(\tau_i) + \sum_{k=1}^{i-1} \alpha e^{-\beta(\tau_i-\tau_k)}\right).\end{aligned}$$

As noted by Ogata (1981), this log-likelihood function is easily computed with a recursive formula.

Computation of the log-likelihood function

The log-likelihood of a simple point process N with intensity λ is written

$$\ln \mathcal{L}((N_t)_{t\in[0,T]}) = \int_0^T (1-\lambda(s))ds + \int_0^T \ln \lambda(s) dN_s,$$

which in the case of a Hawkes model can be explicitly computed as

$$\begin{aligned} \ln \mathcal{L}((\tau_i)_{i=1,\ldots,n}) &= \tau_n - \Lambda(0,\tau_n) + \sum_{i=1}^{n} \ln \lambda(\tau_i) \\ &= \tau_n - \Lambda(0,\tau_n) + \sum_{i=1}^{n} \ln \left(\lambda_0(\tau_i) + \sum_{k=1}^{i-1} \alpha e^{-\beta(\tau_i-\tau_k)} \right). \end{aligned}$$

As noted by Ogata (1981), this log-likelihood function is easily computed with a recursive formula. We observe that

$$R_i = \sum_{k=1}^{i-1} e^{-\beta(\tau_i-\tau_k)}$$

Computation of the log-likelihood function

The log-likelihood of a simple point process N with intensity λ is written

$$\ln \mathcal{L}((N_t)_{t\in[0,T]}) = \int_0^T (1-\lambda(s))ds + \int_0^T \ln \lambda(s)dN_s,$$

which in the case of a Hawkes model can be explicitly computed as

$$\begin{aligned}
\ln \mathcal{L}((\tau_i)_{i=1,\dots,n}) &= \tau_n - \Lambda(0,\tau_n) + \sum_{i=1}^{n} \ln \lambda(\tau_i) \\
&= \tau_n - \Lambda(0,\tau_n) + \sum_{i=1}^{n} \ln \left(\lambda_0(\tau_i) + \sum_{k=1}^{i-1} \alpha e^{-\beta(\tau_i-\tau_k)} \right).
\end{aligned}$$

As noted by Ogata (1981), this log-likelihood function is easily computed with a recursive formula. We observe that

$$R_i = \sum_{k=1}^{i-1} e^{-\beta(\tau_i-\tau_k)} = e^{-\beta(\tau_i-\tau_{i-1})} \sum_{k=1}^{i-1} e^{-\beta(\tau_{i-1}-\tau_k)},$$

Computation of the log-likelihood function

$$R_i = e^{-\beta(\tau_i - \tau_{i-1})} \left(1 + \sum_{k=1}^{i-2} e^{-\beta(\tau_{i-1} - \tau_k)} \right)$$

Computation of the log-likelihood function

$$R_i = e^{-\beta(\tau_i - \tau_{i-1})}\left(1 + \sum_{k=1}^{i-2} e^{-\beta(\tau_{i-1} - \tau_k)}\right) = e^{-\beta(\tau_i - \tau_{i-1})}(1 + R_{i-1}).$$

Computation of the log-likelihood function

$$R_i = e^{-\beta(\tau_i - \tau_{i-1})} \left(1 + \sum_{k=1}^{i-2} e^{-\beta(\tau_{i-1} - \tau_k)}\right) = e^{-\beta(\tau_i - \tau_{i-1})}(1 + R_{i-1}).$$

The log-likelihood can thus be recursively computed with

$$\ln \mathcal{L}((\tau_i)_{i=1,\dots,n}) = \tau_n - \Lambda(0, \tau_n) + \sum_{i=1}^{n} \ln \left(\lambda_0(\tau_i) + \alpha R_i\right),$$

where $(R_i)_{i \geq 0}$ is defined above and $R_0 = 0$.

Computation of the log-likelihood function

$$R_i = e^{-\beta(\tau_i-\tau_{i-1})}\left(1+\sum_{k=1}^{i-2} e^{-\beta(\tau_{i-1}-\tau_k)}\right) = e^{-\beta(\tau_i-\tau_{i-1})}(1+R_{i-1}).$$

The log-likelihood can thus be recursively computed with

$$\ln \mathcal{L}((\tau_i)_{i=1,\dots,n}) = \tau_n - \Lambda(0,\tau_n) + \sum_{i=1}^{n} \ln\left(\lambda_0(\tau_i) + \alpha R_i\right),$$

where $(R_i)_{i\geq 0}$ is defined above and $R_0 = 0$.

Direct computation of $\Lambda(0,\tau_n)$ yields to

$$\begin{aligned}\ln \mathcal{L}((\tau_i)_{i=1,\dots,n}) &= \tau_n - \int_0^{\tau_n} \lambda_0(s)ds - \sum_{i=1}^{n} \frac{\alpha}{\beta}\left(1 - e^{-\beta(\tau_n-\tau_i)}\right) \\ &\quad + \sum_{i=1}^{n} \ln\left(\lambda_0(\tau_i)+\alpha R_i\right).\end{aligned}$$

Properties of the maximum-likelihood estimator

Ogata (1978) 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

Properties of the maximum-likelihood estimator

Ogata (1978) 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 $\theta^T = (\lambda_0, \alpha_1, \beta_1)$ as $T \to \infty$,

Properties of the maximum-likelihood estimator

Ogata (1978) 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 $\theta^T = (\lambda_0, \alpha_1, \beta_1)$ as $T \to \infty$,
- asymptotically normal, *i.e.*

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I^{-1}(\theta)),$$

where $I^{-1}(\theta) = \left(\mathbb{E}\left[\frac{1}{\lambda}\frac{\partial \lambda}{\partial \theta_i}\frac{\partial \lambda}{\partial \theta_j}\right]\right)_{i,j}$,

Properties of the maximum-likelihood estimator

Ogata (1978) 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 $\theta^T = (\lambda_0, \alpha_1, \beta_1)$ as $T \to \infty$,
- asymptotically normal, *i.e.*

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow{\mathcal{L}} \mathcal{N}(0, I^{-1}(\theta)),$$

where $I^{-1}(\theta) = \left(\mathbb{E}\left[\frac{1}{\lambda} \frac{\partial \lambda}{\partial \theta_i} \frac{\partial \lambda}{\partial \theta_j} \right] \right)_{i,j}$,

- asymptotically efficient, *i.e.* asymptotically reaches the lower bound of the variance.

Numerical estimation of a simulated process

T	λ_0	α	β
100	1.210 (0.370)	0.588 (0.164)	0.833 (0.442)
1000	1.185 (0.133)	0.590 (0.044)	0.787 (0.068)
10000	1.204 (0.045)	0.602 (0.016)	0.804 (0.023)
100000	1.202 (0.014)	0.600 (0.004)	0.800 (0.007)
True values	1.200	0.600	0.800

Table: Maximum likelihood estimation of a one-dimensional Hawkes process on simulated data. Each estimation is the average result computed on 100 samples of length $[0, T]$. Standard deviations are given in parentheses. These results are obtained with a simple Nelder-Mead simplex algorithm.

Outline

Outline

Multidimensionnal Hawkes process

Let $M \in \mathbb{N}^*$. Let $\{(\tau_i^m)_i\}_{m=1,\dots,M}$ be a M-dimensional point process. We will denote $N_t = (N_t^1, \dots, N_t^M)$ the associated counting process.

Multidimensionnal Hawkes process

Let $M \in \mathbb{N}^*$. Let $\{(\tau_i^m)_i\}_{m=1,\dots,M}$ be a M-dimensional point process. We will denote $N_t = (N_t^1, \dots, N_t^M)$ the associated counting process.

Definition. A multidimensional Hawkes process is defined with intensities λ^m, $m = 1, \dots, M$ given by

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

Multidimensionnal Hawkes process

Let $M \in \mathbb{N}^*$. Let $\{(\tau_i^m)_i\}_{m=1,\dots,M}$ be a M-dimensional point process. We will denote $N_t = (N_t^1, \dots, N_t^M)$ the associated counting process.

Definition. A multidimensional Hawkes process is defined with intensities λ^m, $m = 1, \dots, M$ given by

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

i.e. in its simplest version with $P = 1$ and $\lambda_0^m(t)$ constant

$$\begin{aligned} \lambda^m(t) &= \lambda_0^m + \sum_{n=1}^{M} \int_0^t \alpha^{mn} e^{-\beta^{mn}(t-s)} dN_s^n \\ &= \lambda_0^m + \sum_{n=1}^{M} \sum_{\tau_i^n < t} \alpha^{mn} e^{-\beta^{mn}(t-\tau_i^n)}. \end{aligned}$$

Stationarity condition

Set $P = 1$ to simplify the notations. Using vectorial notation, we have

$$\lambda(t) = \lambda_0 + \int_0^t \Phi(t-s)dN_s, \quad \text{where} \quad \Phi(t) = \left(\alpha^{mn} e^{-\beta^{mn} t}\right)_{m,n=1,\dots,M}.$$

Stationarity condition

Set $P = 1$ to simplify the notations. Using vectorial notation, we have

$$\lambda(t) = \lambda_0 + \int_0^t \Phi(t-s)dN_s, \quad \text{where} \quad \Phi(t) = \left(\alpha^{mn} e^{-\beta^{mn} t}\right)_{m,n=1,\dots,M}.$$

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

$$\mu = \left(I - \int_0^\infty \Phi(u)du\right)^{-1} \lambda_0.$$

Stationarity condition

Set $P = 1$ to simplify the notations. Using vectorial notation, we have

$$\lambda(t) = \lambda_0 + \int_0^t \Phi(t-s)dN_s, \quad \text{where} \quad \Phi(t) = \left(\alpha^{mn} e^{-\beta^{mn} t}\right)_{m,n=1,\ldots,M}.$$

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

$$\mu = \left(I - \int_0^\infty \Phi(u)du\right)^{-1} \lambda_0.$$

A sufficient condition for a multivariate Hawkes process to be stationary is that the spectral radius of the matrix

$$\Gamma = \int_0^\infty \Phi(u)du = \left(\frac{\alpha_{mn}}{\beta_{mn}}\right)_{m,n=1,\ldots,M}$$ be strictly smaller than 1.

Stationarity condition

Set $P = 1$ to simplify the notations. Using vectorial notation, we have

$$\lambda(t) = \lambda_0 + \int_0^t \Phi(t-s)dN_s, \quad \text{where} \quad \Phi(t) = \left(\alpha^{mn} e^{-\beta^{mn} t}\right)_{m,n=1,\ldots,M}.$$

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

$$\mu = \left(I - \int_0^\infty \Phi(u)du\right)^{-1} \lambda_0.$$

A sufficient condition for a multivariate Hawkes process to be stationary is that the spectral radius of the matrix

$$\Gamma = \int_0^\infty \Phi(u)du = \left(\frac{\alpha_{mn}}{\beta_{mn}}\right)_{m,n=1,\ldots,M} \quad \text{be strictly smaller than 1.}$$

We recall that the spectral radius of the matrix Φ is defined as

$$\rho(\Phi) = \max_{a \in \mathcal{S}(\Phi)} |a|, \quad \text{where } \mathcal{S}(\Phi) \text{ denotes the set of all eigenvalues of } \Phi.$$

Outline

Simulation of a multivariate Hawkes process

We generalize the 1D-algorithm in a multidimensional setting.

Simulation of a multivariate Hawkes process

We generalize the 1D-algorithm in a multidimensional setting. We recall that

- $\mathcal{U}_{[0,1]}$ denotes the uniform distribution on the interval [0, 1],
- $[0, T]$ is the time interval on which the process is to be simulated,

and we define

$$I^K(t) = \sum_{n=1}^{K} \lambda^n(t)$$

the sum of the intensities of the first K components of the multivariate process.

Simulation of a multivariate Hawkes process

We generalize the 1D-algorithm in a multidimensional setting. We recall that

- $\mathcal{U}_{[0,1]}$ denotes the uniform distribution on the interval $[0, 1]$,
- $[0, T]$ is the time interval on which the process is to be simulated,

and we define

$$I^K(t) = \sum_{n=1}^{K} \lambda^n(t)$$

the sum of the intensities of the first K components of the multivariate process.

$$I^M(t) = \sum_{n=1}^{M} \lambda^n(t)$$

is thus the total intensity of the multivariate process and we set $I^0 = 0$. The algorithm is then rewritten as follows.

Simulation of a multivariate Hawkes process

1. **Initialization** : Set $i \leftarrow 1$, $i_1 \leftarrow 1, \ldots, i_M \leftarrow 1$ and $I^* \leftarrow I^M(0) = \sum_{i=1}^{M} \lambda_0^i(0)$.

Simulation of a multivariate Hawkes process

1. **Initialization** : Set $i \leftarrow 1$, $i_1 \leftarrow 1, \ldots, i_M \leftarrow 1$ and $I^* \leftarrow I^M(0) = \sum_{i=1}^M \lambda_0^i(0)$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{I^*} \ln U$.

Simulation of a multivariate Hawkes process

1. **Initialization** : Set $i \leftarrow 1$, $i_1 \leftarrow 1, \ldots, i_M \leftarrow 1$ and $I^* \leftarrow I^M(0) = \sum_{i=1}^M \lambda_0^i(0)$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{I^*} \ln U$.
   a. **If** $s > T$ **Then** go to last step.

Simulation of a multivariate Hawkes process

1. **Initialization** : Set $i \leftarrow 1$, $i_1 \leftarrow 1, \dots, i_M \leftarrow 1$ and $I^* \leftarrow I^M(0) = \sum_{i=1}^{M} \lambda_0^i(0)$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{I^*} \ln U$.
   a. **If** $s > T$ **Then** go to last step.
   b. **Attribution Test** : Generate $D \sim \mathcal{U}_{[0,1]}$ and set $\tau_1^{n_0} \leftarrow s$ where n_0 is such that $\frac{I^{n_0-1}(0)}{I^*} < D \leq \frac{I^{n_0}(0)}{I^*}$.

Simulation of a multivariate Hawkes process

1. **Initialization** : Set $i \leftarrow 1$, $i_1 \leftarrow 1, \ldots, i_M \leftarrow 1$ and $I^* \leftarrow I^M(0) = \sum_{i=1}^M \lambda_0^i(0)$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{I^*} \ln U$.
   a. **If** $s > T$ **Then** go to last step.
   b. **Attribution Test** : Generate $D \sim \mathcal{U}_{[0,1]}$ and set $\tau_1^{n_0} \leftarrow s$ where n_0 is such that $\frac{I^{n_0-1}(0)}{I^*} < D \leq \frac{I^{n_0}(0)}{I^*}$.
   c. Set $\tau_1 \leftarrow \tau_1^{n_0}$.

Simulation of a multivariate Hawkes process

1. **Initialization** : Set $i \leftarrow 1$, $i_1 \leftarrow 1, \ldots, i_M \leftarrow 1$ and $I^* \leftarrow I^M(0) = \sum_{i=1}^M \lambda_0^i(0)$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{I^*} \ln U$.
   a. **If** $s > T$ **Then** go to last step.
   b. **Attribution Test** : Generate $D \sim \mathcal{U}_{[0,1]}$ and set $\tau_1^{n_0} \leftarrow s$ where n_0 is such that $\frac{I^{n_0-1}(0)}{I^*} < D \leq \frac{I^{n_0}(0)}{I^*}$.
   c. Set $\tau_1 \leftarrow \tau_1^{n_0}$.
3. **General routine** : Set $i^{n_0} \leftarrow i^{n_0} + 1$ and $i \leftarrow i + 1$.

Simulation of a multivariate Hawkes process

1. **Initialization** : Set $i \leftarrow 1$, $i_1 \leftarrow 1, \ldots, i_M \leftarrow 1$ and $I^* \leftarrow I^M(0) = \sum_{i=1}^{M} \lambda_0^i(0)$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{I^*} \ln U$.
   a. **If** $s > T$ **Then** go to last step.
   b. **Attribution Test** : Generate $D \sim \mathcal{U}_{[0,1]}$ and set $\tau_1^{n_0} \leftarrow s$ where n_0 is such that $\frac{I^{n_0-1}(0)}{I^*} < D \leq \frac{I^{n_0}(0)}{I^*}$.
   c. Set $\tau_1 \leftarrow \tau_1^{n_0}$.
3. **General routine** : Set $i^{n_0} \leftarrow i^{n_0} + 1$ and $i \leftarrow i + 1$.
   a. **Update maximum intensity:** Set $I^* \leftarrow I^M(t_{i-1}) + \sum_{n=1}^{M} \sum_{j=1}^{P} \alpha_j^{nn_0}$.

Simulation of a multivariate Hawkes process

1. **Initialization** : Set $i \leftarrow 1$, $i_1 \leftarrow 1, \dots, i_M \leftarrow 1$ and $I^* \leftarrow I^M(0) = \sum_{i=1}^{M} \lambda_0^i(0)$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{I^*} \ln U$.
   a. **If** $s > T$ **Then** go to last step.
   b. **Attribution Test** : Generate $D \sim \mathcal{U}_{[0,1]}$ and set $\tau_1^{n_0} \leftarrow s$ where n_0 is such that $\frac{I^{n_0-1}(0)}{I^*} < D \leq \frac{I^{n_0}(0)}{I^*}$.
   c. Set $\tau_1 \leftarrow \tau_1^{n_0}$.
3. **General routine** : Set $i^{n_0} \leftarrow i^{n_0} + 1$ and $i \leftarrow i + 1$.
   a. **Update maximum intensity:** Set $I^* \leftarrow I^M(t_{i-1}) + \sum_{n=1}^{M} \sum_{j=1}^{P} \alpha_j^{nn_0}$.
   b. **New event:** Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow s - \frac{1}{I^*} \ln U$. **If** $s > T$, **Then** go to the last step.

Simulation of a multivariate Hawkes process

1. **Initialization** : Set $i \leftarrow 1$, $i_1 \leftarrow 1, \ldots, i_M \leftarrow 1$ and $I^* \leftarrow I^M(0) = \sum_{i=1}^{M} \lambda_0^i(0)$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{I^*} \ln U$.
   a. **If** $s > T$ **Then** go to last step.
   b. **Attribution Test** : Generate $D \sim \mathcal{U}_{[0,1]}$ and set $\tau_1^{n_0} \leftarrow s$ where n_0 is such that $\frac{I^{n_0-1}(0)}{I^*} < D \leq \frac{I^{n_0}(0)}{I^*}$.
   c. Set $\tau_1 \leftarrow \tau_1^{n_0}$.
3. **General routine** : Set $i^{n_0} \leftarrow i^{n_0} + 1$ and $i \leftarrow i + 1$.
   a. **Update maximum intensity:** Set $I^* \leftarrow I^M(t_{i-1}) + \sum_{n=1}^{M} \sum_{j=1}^{P} \alpha_j^{nn_0}$.
   b. **New event:** Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow s - \frac{1}{I^*} \ln U$. **If** $s > T$, **Then** go to the last step.
   c. **Attribution-Rejection test:** Generate $D \sim \mathcal{U}_{[0,1]}$. **If** $D \leq \frac{I^M(s)}{I^*}$, **Then** set $\tau_{i_0^n}^{n_0} \leftarrow s$ where n_0 is such that $\frac{I^{n_0-1}(s)}{I^*} < D \leq \frac{I^{n_0}(s)}{I^*}$, and $\tau_i \leftarrow \tau_{i^{n_0}}^{n_0}$ and go through the general routine again, **Else** update $I^* \leftarrow I^M(s)$ and try a new date at step b. of the general routine.

Simulation of a multivariate Hawkes process

1. **Initialization** : Set $i \leftarrow 1$, $i_1 \leftarrow 1, \ldots, i_M \leftarrow 1$ and $I^* \leftarrow I^M(0) = \sum_{i=1}^{M} \lambda_0^i(0)$.
2. **First event** : Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow -\frac{1}{I^*} \ln U$.
   a. **If** $s > T$ **Then** go to last step.
   b. **Attribution Test** : Generate $D \sim \mathcal{U}_{[0,1]}$ and set $\tau_1^{n_0} \leftarrow s$ where n_0 is such that $\frac{I^{n_0-1}(0)}{I^*} < D \leq \frac{I^{n_0}(0)}{I^*}$.
   c. Set $\tau_1 \leftarrow \tau_1^{n_0}$.
3. **General routine** : Set $i^{n_0} \leftarrow i^{n_0} + 1$ and $i \leftarrow i + 1$.
   a. **Update maximum intensity:** Set $I^* \leftarrow I^M(t_{i-1}) + \sum_{n=1}^{M} \sum_{j=1}^{P} \alpha_j^{nn_0}$.
   b. **New event:** Generate $U \sim \mathcal{U}_{[0,1]}$ and set $s \leftarrow s - \frac{1}{I^*} \ln U$. **If** $s > T$, **Then** go to the last step.
   c. **Attribution-Rejection test:** Generate $D \sim \mathcal{U}_{[0,1]}$. **If** $D \leq \frac{I^M(s)}{I^*}$, **Then** set $\tau_{i_0^n}^{n_0} \leftarrow s$ where n_0 is such that $\frac{I^{n_0-1}(s)}{I^*} < D \leq \frac{I^{n_0}(s)}{I^*}$, and $\tau_i \leftarrow \tau_{i^{n_0}}^{n_0}$ and go through the general routine again, **Else** update $I^* \leftarrow I^M(s)$ and try a new date at step b. of the general routine.
4. **Output**: Retrieve the simulated process $((\tau_i^n)_i)_{n=1,\ldots,M}$ on $[0, T]$.

Sample paths of a bivariate Hawkes process

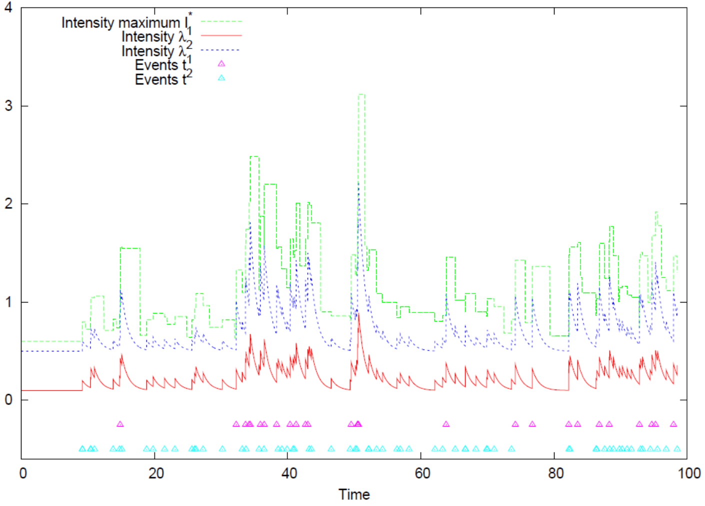


Figure: Simulation of a two-dimensional Hawkes process with P = 1 and $\lambda_0^1 = 0.1$, $\alpha_1^{11} = 0.2$, $\beta_1^{11} = 1.0$, $\alpha_1^{12} = 0.1$, $\beta_1^{12} = 1.0$, $\lambda_0^2 = 0.5$, $\alpha_1^{21} = 0.5$, $\beta_1^{21} = 1.0$, $\alpha_1^{22} = 0.1$, $\beta_1^{22} = 1.0$.

Testing the simulated data

The integrated intensity of the m-th coordinate of a multidimensional Hawkes process between two consecutive events t_{i-1}^m and t_i^m of type m is computed as

$$\Lambda^m(\tau_{i-1}^m, \tau_i^m) = \int_{\tau_{i-1}^m}^{\tau_i^m} \lambda^m(s)ds$$

Testing the simulated data

The integrated intensity of the m-th coordinate of a multidimensional Hawkes process between two consecutive events t_{i-1}^m and t_i^m of type m is computed as

$$\begin{aligned}\Lambda^m(\tau_{i-1}^m, \tau_i^m) &= \int_{\tau_{i-1}^m}^{\tau_i^m} \lambda^m(s)ds \\ &= \int_{\tau_{i-1}^m}^{\tau_i^m} \lambda_0^m(s)ds + \int_{\tau_{i-1}^m}^{\tau_i^m} \sum_{n=1}^{M} \sum_{\tau_k^n < s} \alpha^{mn} e^{-\beta^{nm}(s-\tau_k^n)} ds\end{aligned}$$

Testing the simulated data

The integrated intensity of the m-th coordinate of a multidimensional Hawkes process between two consecutive events t_{i-1}^m and t_i^m of type m is computed as

$$\begin{aligned}
\Lambda^m(\tau_{i-1}^m, \tau_i^m) &= \int_{\tau_{i-1}^m}^{\tau_i^m} \lambda^m(s)ds \\
&= \int_{\tau_{i-1}^m}^{\tau_i^m} \lambda_0^m(s)ds + \int_{\tau_{i-1}^m}^{\tau_i^m} \sum_{n=1}^{M} \sum_{\tau_k^n < s} \alpha^{mn} e^{-\beta^{nm}(s-\tau_k^n)} ds \\
&= \int_{\tau_{i-1}^m}^{\tau_i^m} \lambda_0^m(s)ds + \int_{\tau_{i-1}^m}^{\tau_i^m} \sum_{n=1}^{M} \sum_{\tau_k^n < \tau_{i-1}^m} \alpha^{mn} e^{-\beta^{nm}(s-\tau_k^n)} ds \\
&\quad + \int_{\tau_{i-1}^m}^{\tau_i^m} \sum_{n=1}^{M} \sum_{\tau_{i-1}^m \leq \tau_k^n < s} \alpha^{mn} e^{-\beta^{nm}(s-\tau_k^n)} ds
\end{aligned}$$

$$
\begin{aligned}
\Lambda^m(\tau_{i-1}^m, \tau_i^m) &= \int_{\tau_{i-1}^m}^{\tau_i^m} \lambda_0^m(s)ds \\
&+ \sum_{n=1}^{M} \sum_{\tau_k^n < \tau_{i-1}^m} \frac{\alpha^{mn}}{\beta^{mn}} \left[e^{-\beta^{nm}(\tau_{i-1}^m - \tau_k^n)} - e^{-\beta^{nm}(\tau_i^m - \tau_k^n)} \right] \\
&+ \sum_{n=1}^{M} \sum_{\tau_{i-1}^m \leq \tau_k^n < \tau_i^m} \frac{\alpha^{mn}}{\beta^{mn}} \left[1 - e^{-\beta^{nm}(\tau_i^m - \tau_k^n)} \right].
\end{aligned}
$$

$$
\begin{aligned}
\Lambda^m(\tau_{i-1}^m, \tau_i^m) &= \int_{\tau_{i-1}^m}^{\tau_i^m} \lambda_0^m(s)ds \\
&+ \sum_{n=1}^{M} \sum_{\tau_k^n < \tau_{i-1}^m} \frac{\alpha^{mn}}{\beta^{mn}} \left[e^{-\beta^{nm}(\tau_{i-1}^m - \tau_k^n)} - e^{-\beta^{nm}(\tau_i^m - \tau_k^n)} \right] \\
&+ \sum_{n=1}^{M} \sum_{\tau_{i-1}^m \leq \tau_k^n < \tau_i^m} \frac{\alpha^{mn}}{\beta^{mn}} \left[1 - e^{-\beta^{nm}(\tau_i^m - \tau_k^n)} \right].
\end{aligned}
$$

This computation can be simplified with a recursive element. Let us denote

$$
A_{i-1}^{mn} = \sum_{\tau_k^n < \tau_{i-1}^m} e^{-\beta^{nm}(\tau_{i-1}^m - \tau_k^n)}.
$$

We observe that

$$A_{i-1}^{mn} = \sum_{\tau_k^n < \tau_{i-1}^m} e^{-\beta^{mn}(\tau_{i-1}^m - \tau_k^n)}$$

We observe that

$$
\begin{aligned}
A_{i-1}^{mn} &= \sum_{\tau_k^n < \tau_{i-1}^m} e^{-\beta^{mn}(\tau_{i-1}^m - \tau_k^n)} \\
&= e^{-\beta^{mn}(\tau_{i-1}^m - \tau_{i-2}^m)} \sum_{\tau_k^n < \tau_{i-2}^m} e^{-\beta^{mn}(\tau_{i-2}^m - \tau_k^n)} \\
&\quad + \sum_{\tau_{i-2}^m \le \tau_k^n < \tau_{i-1}^m} e^{-\beta^{mn}(\tau_{i-1}^m - \tau_k^n)}
\end{aligned}
$$

We observe that

$$
\begin{aligned}
A_{i-1}^{mn} &= \sum_{\tau_k^n < \tau_{i-1}^m} e^{-\beta^{mn}(\tau_{i-1}^m - \tau_k^n)} \\
&= e^{-\beta^{mn}(\tau_{i-1}^m - \tau_{i-2}^m)} \sum_{\tau_k^n < \tau_{i-2}^m} e^{-\beta^{mn}(\tau_{i-2}^m - \tau_k^n)} \\
&\quad + \sum_{\tau_{i-2}^m \le \tau_k^n < \tau_{i-1}^m} e^{-\beta^{mn}(\tau_{i-1}^m - \tau_k^n)} \\
&= e^{-\beta^{mn}(\tau_{i-1}^m - \tau_{i-2}^m)} A_{i-2}^{mn} + \sum_{\tau_{i-2}^m \le \tau_k^n < \tau_{i-1}^m} e^{-\beta^{mn}(\tau_{i-1}^m - \tau_k^n)}.
\end{aligned}
$$

The integrated density can thus be written $\forall i \in \mathbb{N}^*$,

$$\begin{aligned}
\Lambda^m(\tau_{i-1}^m, \tau_i^m) &= \int_{\tau_{i-1}^m}^{\tau_i^m} \lambda_0^m(s)ds \\
&\quad + \sum_{n=1}^{M} \frac{\alpha^{mn}}{\beta^{mn}} \left[\left(1 - e^{-\beta^{mn}(\tau_i^m - \tau_{i-1}^m)} A_{i-1}^{mn}\right) \right. \\
&\quad \left. + \sum_{\tau_{i-1}^m \leq \tau_k^n < \tau_i^m} \left(1 - e^{-\beta^{mn}(\tau_i^m - \tau_k^n)}\right) \right].
\end{aligned}$$

where A is defined as previously with $A_0^{mn} = 0$.

The integrated density can thus be written $\forall i \in \mathbb{N}^*$,

$$\begin{aligned}\Lambda^m(\tau_{i-1}^m, \tau_i^m) &= \int_{\tau_{i-1}^m}^{\tau_i^m} \lambda_0^m(s)ds \\ &+ \sum_{n=1}^{M} \frac{\alpha^{mn}}{\beta^{mn}} \Bigg[\left(1 - e^{-\beta^{mn}(\tau_i^m - \tau_{i-1}^m)} A_{i-1}^{mn}\right) \\ &+ \sum_{\tau_{i-1}^m \leq \tau_k^n < \tau_i^m} \left(1 - e^{-\beta^{mn}(\tau_i^m - \tau_k^n)}\right) \Bigg].\end{aligned}$$

where A is defined as previously with $A_0^{mn} = 0$.

Time change property. As for the one-dimensional case, the durations $\vartheta_i^m - \vartheta_{i-1}^m = \Lambda^m(\tau_{i-1}^m, \tau_i^m)$ are exponentially distributed with parameter 1. See *e.g.* Bowsher (2007).

Testing the simulated data

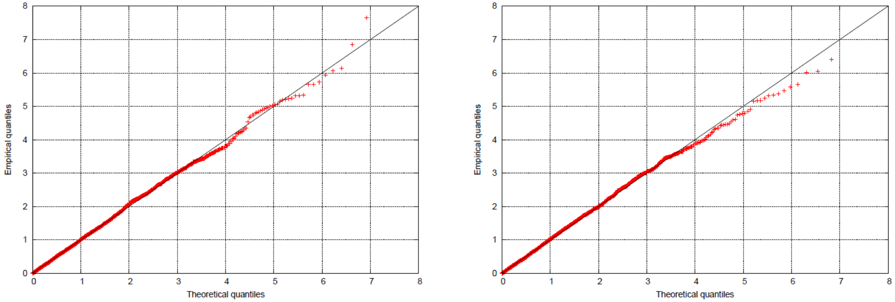


Figure: Quantile plots for one sample of simulated data of a two-dimensional Hawkes process with $P = 1$ and parameters given as in the simulation. *(Left)* $m = 1$. *(Right)* $m = 2$.

Outline

Computation of the log-likelihood function

The log-likelihood of a multidimensional Hawkes process can be computed as the sum of the likelihood of each coordinate, *i.e.* is written

$$\ln \mathcal{L}(\{\tau_i\}_{i=1,\ldots,N}) = \sum_{m=1}^{M} \ln \mathcal{L}^m(\{\tau_i\}),$$

Computation of the log-likelihood function

The log-likelihood of a multidimensional Hawkes process can be computed as the sum of the likelihood of each coordinate, *i.e.* is written

$$\ln \mathcal{L}(\{\tau_i\}_{i=1,\dots,N}) = \sum_{m=1}^{M} \ln \mathcal{L}^m(\{\tau_i\}),$$

where each term is defined by

$$\ln \mathcal{L}^m(\{\tau_i\}) = \int_0^T (1 - \lambda^m(s))ds + \int_0^T \ln \lambda^m(s) dN_s^m.$$

Computation of the log-likelihood function

The log-likelihood of a multidimensional Hawkes process can be computed as the sum of the likelihood of each coordinate, *i.e.* is written

$$\ln \mathcal{L}(\{\tau_i\}_{i=1,\dots,N}) = \sum_{m=1}^{M} \ln \mathcal{L}^m(\{\tau_i\}),$$

where each term is defined by

$$\ln \mathcal{L}^m(\{\tau_i\}) = \int_0^T (1 - \lambda^m(s))ds + \int_0^T \ln \lambda^m(s) dN_s^m.$$

In the case of a multidimensional Hawkes process, denoting $\{\tau_i\}_{i=1,\dots,N}$ the ordered pool of all events $\{\{\tau_i^m\}_{m=1,\dots,M}\}$, this log-likelihood can be computed as

$$\begin{aligned} \ln \mathcal{L}^m(\{\tau_i\}) &= T - \Lambda^m(0,T) \\ &\quad + \sum_{i=1}^{N} z_i^m \ln \left(\lambda_0^m(\tau_i) + \sum_{n=1}^{M} \sum_{j=1}^{P} \sum_{\tau_k^n < \tau_i} \alpha_j^{mn} e^{-\beta_j^{mn}(\tau_i - \tau_k^n)} \right), \end{aligned}$$

where z_i^m is equal to 1 if the event τ_i is of type m, 0 otherwise.

As in the one dimensional case, this can be computed in a recursive way.

As in the one dimensional case, this can be computed in a recursive way. We observe that (with $P = 1$)

$$R_l^{mn} = \sum_{\tau_k^n < \tau_l^m} e^{-\beta^{mn}(\tau_l^m - \tau_k^n)}$$

As in the one dimensional case, this can be computed in a recursive way. We observe that (with $P = 1$)

$$R_l^{mn} = \sum_{\tau_k^n < \tau_l^m} e^{-\beta^{mn}(\tau_l^m - \tau_k^n)}$$

$$= \begin{cases} e^{-\beta^{mn}(\tau_l^m - \tau_l^m - 1)} R_{l-1}^{mn} + \displaystyle\sum_{\tau_{l-1}^m \leq \tau_k^n < \tau_l^m} e^{-\beta^{mn}(\tau_l^m - \tau_k^n)} & \text{if} \quad m \neq n, \\ e^{-\beta^{mn}(\tau_l^m - \tau_l^m - 1)}(1 + R_{l-1}^{mn}) & \text{if} \quad m = n. \end{cases}$$

As in the one dimensional case, this can be computed in a recursive way. We observe that (with $P=1$)

$$R_l^{mn} = \sum_{\tau_k^n<\tau_l^m} e^{-\beta^{mn}(\tau_l^m-\tau_k^n)}$$

$$= \begin{cases} e^{-\beta^{mn}(\tau_l^m-\tau_l^m-1)} R_{l-1}^{mn} + \sum_{\tau_{l-1}^m \le \tau_k^n < \tau_l^m} e^{-\beta^{mn}(\tau_l^m-\tau_k^n)} & \text{if} \quad m \neq n, \\ e^{-\beta^{mn}(\tau_l^m-\tau_l^m-1)}(1+R_{l-1}^{mn}) & \text{if} \quad m = n. \end{cases}$$

The final expression of the log-likelihood may be written

$$\begin{aligned} \ln \mathcal{L}^m(\{\tau_i\}) &= T - \sum_{i=1}^{N}\sum_{n=1}^{M} \frac{\alpha^{mn}}{\beta^{mn}}\left(1-e^{-\beta^{mn}(T-\tau_i)}\right) \\ &\quad + \sum_{\tau_l^m} \ln\left(\lambda_0^m(\tau_l^m) + \sum_{n=1}^{M} \alpha^{mn} R_l^{mn}\right), \end{aligned}$$

where R_l^{mn} is defined above and $R_0^{mn}=0$.

Numerical estimation of a simulated process

T	λ_0^1	α^{11}	$\beta 11$	α^{12}	β^{12}	λ_0^2	α^{21}	β^{21}	α^{22}	β^{22}
100	0.614	0.510	0.369	1.482	1.710	0.518	0.337	0.600	1.605	2.595
	(0.372)	(0.268)	(0.269)	(1.216)	(3.172)	(0.272)	(0.206)	(0.365)	(2.051)	(6.586)
500	0.516	0.505	0.268	1.043	0.865	0.518	0.264	0.507	0.814	1.048
	(0.112)	(0.085)	(0.080)	(0.214)	(0.479)	(0.120)	(0.080)	(0.084)	(0.278)	(0.221)
1000	0.507	0.492	0.254	1.018	0.761	0.513	0.255	0.488	0.794	1.003
	(0.079)	(0.054)	(0.052)	(0.122)	(0.203)	(0.092)	(0.052)	(0.061)	(0.387)	(0.152)
True	0.500	0.500	0.250	1.000	0.750	0.500	0.250	0.500	0.750	1.000

Table: Maximum likelihood estimation of a two-dimensional Hawkes process on simulated data. Each estimation is the average result computed on 100 samples of length $[0, T]$. Standard deviations are given in parentheses.

Outline

Outline

A one-dimensional price model

Bacry, Delattre, Hoffmann & Muzy (2013) propose a one-dimensional case where the price p is written

$$p(t) = N^1(t) - N^2(t),$$

A one-dimensional price model

Bacry, Delattre, Hoffmann & Muzy (2013) propose a one-dimensional case where the price p is written

$$p(t) = N^1(t) - N^2(t),$$

where N^i, $i \in \{1, 2\}$ is a Hawkes process with intensities λ_i, $i \in \{1, 2\}$ such that

$$\lambda_1(t) = \lambda_0 + \int_{-\infty}^{t} \alpha e^{-\beta(t-s)} dN_s^2,$$

$$\lambda_2(t) = \lambda_0 + \int_{-\infty}^{t} \alpha e^{-\beta(t-s)} dN_s^1.$$

A one-dimensional price model

Bacry, Delattre, Hoffmann & Muzy (2013) propose a one-dimensional case where the price p is written

$$p(t) = N^1(t) - N^2(t),$$

where N^i, $i \in \{1,2\}$ is a Hawkes process with intensities λ_i, $i \in \{1,2\}$ such that

$$\lambda_1(t) = \lambda_0 + \int_{-\infty}^{t} \alpha e^{-\beta(t-s)} dN_s^2,$$

$$\lambda_2(t) = \lambda_0 + \int_{-\infty}^{t} \alpha e^{-\beta(t-s)} dN_s^1.$$

- No self-excitation of upward (resp. downward) jumps on following upward (resp. downward) jumps.

A one-dimensional price model

Bacry, Delattre, Hoffmann & Muzy (2013) propose a one-dimensional case where the price p is written

$$p(t) = N^1(t) - N^2(t),$$

where N^i, $i \in \{1, 2\}$ is a Hawkes process with intensities λ_i, $i \in \{1, 2\}$ such that

$$\lambda_1(t) = \lambda_0 + \int_{-\infty}^{t} \alpha e^{-\beta(t-s)} dN_s^2,$$

$$\lambda_2(t) = \lambda_0 + \int_{-\infty}^{t} \alpha e^{-\beta(t-s)} dN_s^1.$$

- No self-excitation of upward (resp. downward) jumps on following upward (resp. downward) jumps.
- Only cross-excitation terms are kept, enforcing the mean-reversion empirically observed on the price p.

A one-dimensional price model

Bacry, Delattre, Hoffmann & Muzy (2013) propose a one-dimensional case where the price p is written

$$p(t) = N^1(t) - N^2(t),$$

where N^i, $i \in \{1, 2\}$ is a Hawkes process with intensities λ_i, $i \in \{1, 2\}$ such that

$$\lambda_1(t) = \lambda_0 + \int_{-\infty}^{t} \alpha e^{-\beta(t-s)} dN_s^2,$$

$$\lambda_2(t) = \lambda_0 + \int_{-\infty}^{t} \alpha e^{-\beta(t-s)} dN_s^1.$$

- No self-excitation of upward (resp. downward) jumps on following upward (resp. downward) jumps.
- Only cross-excitation terms are kept, enforcing the mean-reversion empirically observed on the price p.
- Cross-excitation is set to be symmetric.

An analytical expression for the variance of the price

A volatility signature plot plots the realized variance as a function of the sampling period

$$RV(\tau) = \frac{1}{\tau} \sum_{i=1}^{I} (\hat{p}(i\tau) - \hat{p}((i-1)\tau))^2,$$

where I is the number of observations $\hat{p}$.

An analytical expression for the variance of the price

A volatility signature plot plots the realized variance as a function of the sampling period

$$RV(\tau) = \frac{1}{\tau}\sum_{i=1}^{I}(\hat{p}(i\tau) - \hat{p}((i-1)\tau))^2,$$

where I is the number of observations $\hat{p}$. Bacry et al. (2011) shows that the theoretical signature plot of the stationary model can be theoretically computed as

$$C(\tau) = \frac{1}{\tau}\mathbb{E}[p(\tau)^2] = \Lambda\left(\kappa^2 + (1-\kappa^2)\frac{1-e^{-\gamma\tau}}{\gamma\tau}\right),$$

An analytical expression for the variance of the price

A volatility signature plot plots the realized variance as a function of the sampling period

$$RV(\tau) = \frac{1}{\tau} \sum_{i=1}^{I} (\hat{p}(i\tau) - \hat{p}((i-1)\tau))^2,$$

where I is the number of observations $\hat{p}$. Bacry et al. (2011) shows that the theoretical signature plot of the stationary model can be theoretically computed as

$$C(\tau) = \frac{1}{\tau}\mathbb{E}[p(\tau)^2] = \Lambda \left(\kappa^2 + (1-\kappa^2)\frac{1-e^{-\gamma\tau}}{\gamma\tau} \right),$$

where

$$\Lambda = \frac{2\lambda_0}{1-\alpha/\beta}, \quad \kappa = \frac{1}{1+\alpha/\beta} \quad \text{and} \quad \gamma = \alpha + \beta.$$

See Bacry et al. (2011), Appendix 1.

Likehood functions

Consequently the likelihood functions $\theta \to L_1(\theta)$ and $\theta \to L_2(\theta)$ generated by the observation of $N^1(t)$ and $N^2(t)$ over $[0, T]$ are given by

$$\begin{aligned} L_1(\theta) &= \sum_{0 \le \tau_i^{(1)} < N^1(T)} \ln\left(\lambda_0 + \sum_{0 \le \tau_j^{(2)} < N^2(T)} \alpha e^{-\beta(\tau_i^{(1)} - \tau_j^{(2)})}\right) \\ &\quad -(\lambda_0 - 1)T - \sum_{0 \le \tau_j^{(2)} < N^2(T)} \frac{\alpha}{\beta}\left(1 - e^{-\beta(T - \tau_j^{(2)})}\right), \end{aligned}$$

Likehood functions

$$L_2(\theta) = \sum_{0\le\tau_i^{(2)}<N^2(T)} \ln\left(\lambda_0 + \sum_{0\le\tau_j^{(1)}<N^2(T)} \alpha e^{-\beta(\tau_i^{(2)}-\tau_j^{(1)})}\right)$$
$$-(\lambda_0-1)T - \sum_{0\le\tau_j^{(1)}<N^1(T)} \frac{\alpha}{\beta}\left(1-e^{-\beta(T-\tau_j^{(1)})}\right).$$

and the MLE of θ based on the observation of $p(t)$ over $[0, T]$ is thus given by

$$\hat{\theta}_{MLE} = \mathrm{argmin}_\theta\left(L_1(\theta)+L_2(\theta)\right).$$

Of course, the minimization must be performed under the constraints

$$\lambda_0 > 0, \quad \alpha > 0, \quad \beta > 0, \quad \text{and the stability condition } \frac{\alpha}{\beta} < 1.$$

Fitting the volatility signature plot : Simulated data

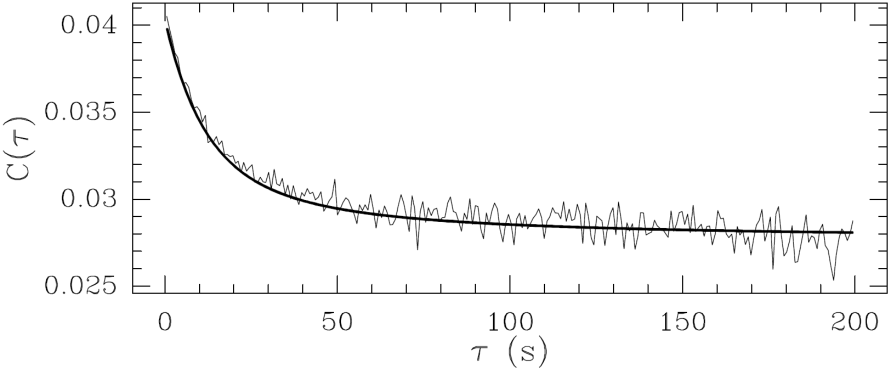


Figure: Numerical simulation of a 1D price model using mutually exciting Hawkes processes with parameters $\lambda_0 = 0.16$, $\alpha = 0.024$, $\beta = 0.11$. Estimated signature plot $C(\tau)$ with the corresponding expected analytical shape. τ is expressed in seconds. Reproduced from Bacry et al. (2011).

Fitting the volatility signature plot : Market data

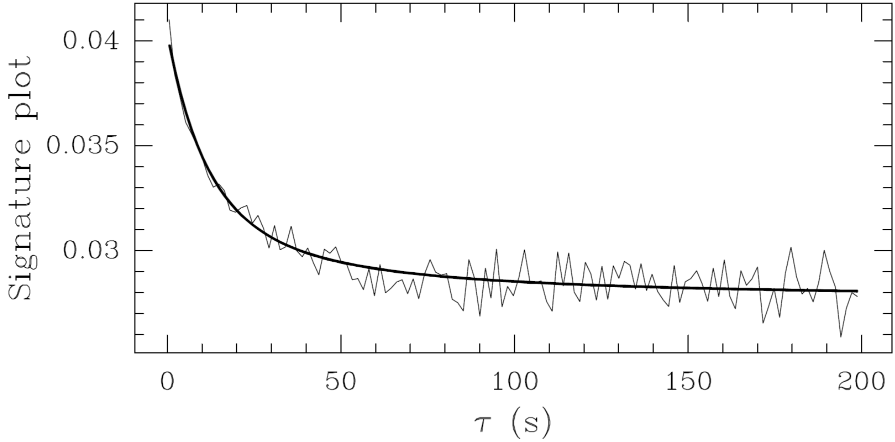


Figure: Euro-Bund contract on the 3rd of November 2009 (contract maturity 12/2009) from 8am to 10pm. Associated mean daily signature plot. The solid line represents the fit using the regression estimator for the 1D Hawkes model. Reproduced from Bacry et al. (2011).

Outline

A two-dimensional model

Bacry et al. (2013) also propose a bivariate version of the model

$$p^1(t) = N^1(t) - N^2(t),$$
$$p^2(t) = N^3(t) - N^4(t),$$

A two-dimensional model

Bacry et al. (2013) also propose a bivariate version of the model

$$p^1(t) = N^1(t) - N^2(t),$$

$$p^2(t) = N^3(t) - N^4(t),$$

in which $\mathbf{N} = (N^i)_{i=1,\dots,4}$ is a Hawkes process with intensity

$$\lambda(t) = \lambda_0 + \int_0^t \begin{pmatrix} 0 & \varphi^{12} & \varphi^{13} & 0 \\ \varphi^{12} & 0 & 0 & \varphi^{13} \\ \varphi^{31} & 0 & 0 & \varphi^{34} \\ 0 & \varphi^{31} & \varphi^{34} & 0 \end{pmatrix} (t-s) d\mathbf{N}_s,$$

where $\varphi^{ij}(t-s) = \alpha^{ij} e^{-\beta^{ij}(t-s)}$.

A two-dimensional model

Bacry et al. (2013) also propose a bivariate version of the model

$$p^1(t) = N^1(t) - N^2(t),$$

$$p^2(t) = N^3(t) - N^4(t),$$

in which $\mathbf{N} = (N^i)_{i=1,\dots,4}$ is a Hawkes process with intensity

$$\lambda(t) = \lambda_0 + \int_0^t \begin{pmatrix} 0 & \varphi^{12} & \varphi^{13} & 0 \\ \varphi^{12} & 0 & 0 & \varphi^{13} \\ \varphi^{31} & 0 & 0 & \varphi^{34} \\ 0 & \varphi^{31} & \varphi^{34} & 0 \end{pmatrix} (t-s) d\mathbf{N}_s,$$

where $\varphi^{ij}(t-s) = \alpha^{ij} e^{-\beta^{ij}(t-s)}$.

- self-exciting terms are ruled out : $\forall i$, $\varphi^{ii} = 0$;

A two-dimensional model

Bacry et al. (2013) also propose a bivariate version of the model

$$p^1(t) = N^1(t) - N^2(t),$$

$$p^2(t) = N^3(t) - N^4(t),$$

in which $\mathbf{N} = (N^i)_{i=1,\dots,4}$ is a Hawkes process with intensity

$$\lambda(t) = \lambda_0 + \int_0^t \begin{pmatrix} 0 & \varphi^{12} & \varphi^{13} & 0 \\ \varphi^{12} & 0 & 0 & \varphi^{13} \\ \varphi^{31} & 0 & 0 & \varphi^{34} \\ 0 & \varphi^{31} & \varphi^{34} & 0 \end{pmatrix} (t-s) d\mathbf{N}_s,$$

where $\varphi^{ij}(t-s) = \alpha^{ij} e^{-\beta^{ij}(t-s)}$.

- self-exciting terms are ruled out : $\forall i$, $\varphi^{ii} = 0$;
- upward and downward effects are assumed to be symmetric within the processes p^1 and p^2 ($\varphi^{13} = \varphi^{24}$, $\varphi^{31} = \varphi^{42}$) ;

A two-dimensional model

Bacry et al. (2013) also propose a bivariate version of the model

$$p^1(t) = N^1(t) - N^2(t),$$
$$p^2(t) = N^3(t) - N^4(t),$$

in which $\mathbf{N} = (N^i)_{i=1,\ldots,4}$ is a Hawkes process with intensity

$$\lambda(t) = \lambda_0 + \int_0^t \begin{pmatrix} 0 & \varphi^{12} & \varphi^{13} & 0 \\ \varphi^{12} & 0 & 0 & \varphi^{13} \\ \varphi^{31} & 0 & 0 & \varphi^{34} \\ 0 & \varphi^{31} & \varphi^{34} & 0 \end{pmatrix} (t-s) d\mathbf{N}_s,$$

where $\varphi^{ij}(t-s) = \alpha^{ij} e^{-\beta^{ij}(t-s)}$.

- self-exciting terms are ruled out : $\forall i$, $\varphi^{ii} = 0$;
- upward and downward effects are assumed to be symmetric within the processes p^1 and p^2 ($\varphi^{13} = \varphi^{24}$, $\varphi^{31} = \varphi^{42}$) ;
- prices p^1 and p^2 influence each other in a positive way, not a negative one ($\varphi^{14} = \varphi^{41} = \varphi^{23} = \varphi^{32} = 0$).

An explicit expression for the covariance matrix

For this model, Bacry et al. (2013) show that an explicit form of the correlation coefficient

$$\rho(\tau) = \mathrm{Corr}(p^1(t+\tau) - p^1(t), p^2(t+\tau) - p^2(t)) = \frac{C_{12}(\tau)}{C_{11}(\tau)},$$

where $C_{ij}(\tau) = \mathrm{Cov}(p^i(t+\tau) - p^i(t), p^j(t+\tau) - p^j(t))$, can be explicitly computed, although the expected result is quite cumbersome (see Bacry et al. (2011), Proposition 3.1).

An explicit expression for the covariance matrix

For this model, Bacry et al. (2013) show that an explicit form of the correlation coefficient

$$\rho(\tau) = \mathrm{Corr}(p^1(t+\tau) - p^1(t), p^2(t+\tau) - p^2(t)) = \frac{C_{12}(\tau)}{C_{11}(\tau)},$$

where $C_{ij}(\tau) = \mathrm{Cov}(p^i(t+\tau) - p^i(t), p^j(t+\tau) - p^j(t))$, can be explicitly computed, although the expected result is quite cumbersome (see Bacry et al. (2011), Proposition 3.1).

Moreover, they get

$$\rho(\tau) = \frac{R(Q_2 - Q_1)}{4\Lambda}\tau + O(\tau^2) \quad \text{as } \tau \to 0,$$

$$\rho(\tau) = \frac{2\Gamma_{13}(1+\Gamma_{12})}{1+\Gamma_{13}^2 + 2\Gamma_{12} + \Gamma_{12}^2} \quad \text{as } \tau \to \infty,$$

where $\Gamma_{ij} = \alpha^{ij}/\beta^{ij}$, $\Lambda = \lambda_0/(1-\Gamma_{12}-\Gamma_{13})$, $R = \beta\lambda_0//(\Gamma_{12}+\Gamma_{13}-1)$, $Q_1 = -\lambda_0(\Gamma_{12}^2 + \Gamma_{12} - \Gamma_{13}^2)/((((\Gamma_{12}+1)^2 - \Gamma_{13}^2)(1-\Gamma_{12}-\Gamma_{13}))$ and $Q_1 = -\lambda_0\Gamma_{13}/((((\Gamma_{12}+1)^2 - \Gamma_{13}^2)(1-\Gamma_{12}-\Gamma_{13}))$.

Addressing the Epps effect : Simulated data

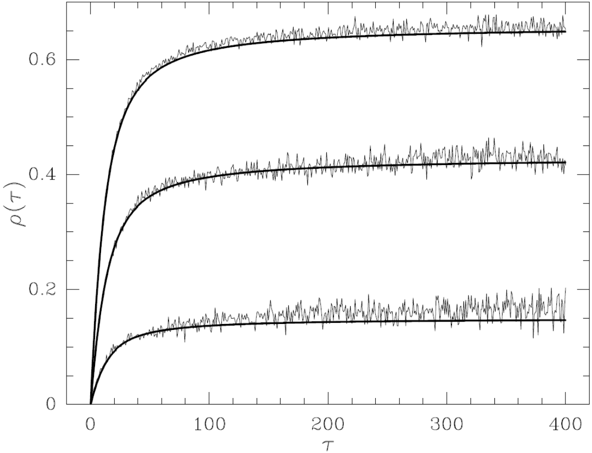


Figure: Epps effect for simulation samples. The estimated correlation coefficients are compared to the expected analytical curves (solid lines). From top top bottom the cross correlation between the two asset returns increases from 0.15 to 0.65. Reproduced from Bacry et al. (2011).

Addressing the Epps effect : Market data

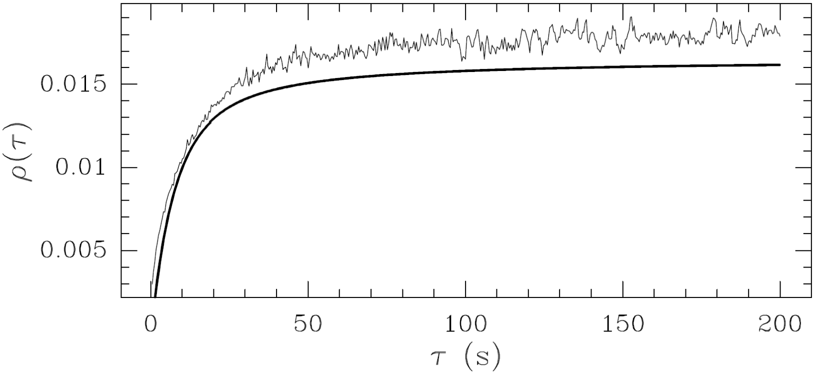


Figure: Epps effect between Euro-Bund et Euro-Bobl returns as measured by the covariance $C_{12}(\tau)$ between the two asset returns at different scales τ. The heavy lines represent a fit of the empirical curves according to the 2D Hawkes model. Reproduced from Bacry et al. (2011).

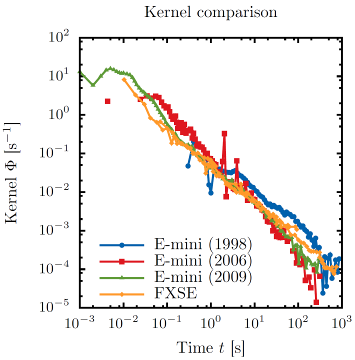
Outline

Power-law kernels

One expects the involved kernels not to be exponential functions.

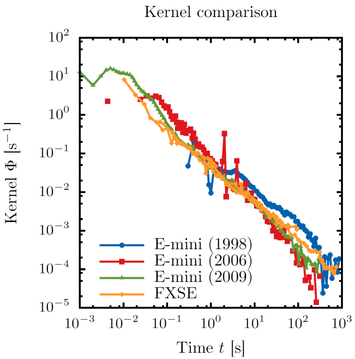
Power-law kernels

One expects the involved kernels not to be exponential functions. Indeed, empirical self-exciting kernels are closer to a power-law than to an exponential.



Power-law kernels

One expects the involved kernels not to be exponential functions. Indeed, empirical self-exciting kernels are closer to a power-law than to an exponential.



In [3], they use non-parametric estimation for the kernels to confirm this power-law behavior and the existence of non-negligible diagonal and anti-diagonal terms.

Scaling limit

Jaisson and Rosenbaum (2015) build a version of the $2D$-Hawkes model that converges at large scales towards a model displaying volatility clustering.

Scaling limit

Jaisson and Rosenbaum (2015) build a version of the $2D$-Hawkes model that converges at large scales towards a model displaying volatility clustering.

Let $\varphi = \varphi_1 - \varphi_2$. Under some assumptions, the sequence of Hawkes-based price models $P_t^T = \frac{1}{T}\left(N^1(Tt) - N^2(Tt)\right)$, converges in law, for the Skorohod topology, toward a Heston-type process P on $[0,1]$ defined by

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

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

Scaling limit

Jaisson and Rosenbaum (2015) build a version of the $2D$-Hawkes model that converges at large scales towards a model displaying volatility clustering.

Let $\varphi = \varphi_1 - \varphi_2$. Under some assumptions, the sequence of Hawkes-based price models $P_t^T = \frac{1}{T}\left(N^1(Tt) - N^2(Tt)\right)$, converges in law, for the Skorohod topology, toward a Heston-type process P on $[0,1]$ defined by

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

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

This work can be seen as the very first step towards an "across scales" unified model that would fit both microstructure stylized facts of price (*i.e.*, point process with strong mean reversion) and "diffusive" stylized facts (volatility clustering and multifractality).

Marked Hawkes processes

Fauth and Tudor (2012) proposed to describe bid and ask prices of an asset (or a couple of assets) within the framework of marked multivariate Hawkes model.

Marked Hawkes processes

Fauth and Tudor (2012) proposed to describe bid and ask prices of an asset (or a couple of assets) within the framework of marked multivariate Hawkes model.

Motivated by the empirical observation that, as the transaction volumes increase, the inter-trade durations decrease, the authors proposed to consider, in addition to an exponential Hawkes kernel, a multiplicative mark that corresponds to a power-law function of the volumes.

Marked Hawkes processes

Fauth and Tudor (2012) proposed to describe bid and ask prices of an asset (or a couple of assets) within the framework of marked multivariate Hawkes model.

Motivated by the empirical observation that, as the transaction volumes increase, the inter-trade durations decrease, the authors proposed to consider, in addition to an exponential Hawkes kernel, a multiplicative mark that corresponds to a power-law function of the volumes.

Their model thus describes the events corresponding to an increase/decrease of the bid/ask as a $4D$-Hawkes process marked by transaction volumes. By adding suitable constraints in order to avoid infinite spread, they calibrated the model on FX rates data by a maximum likelihood approach.

Marked Hawkes processes

Fauth and Tudor (2012) proposed to describe bid and ask prices of an asset (or a couple of assets) within the framework of marked multivariate Hawkes model.

Motivated by the empirical observation that, as the transaction volumes increase, the inter-trade durations decrease, the authors proposed to consider, in addition to an exponential Hawkes kernel, a multiplicative mark that corresponds to a power-law function of the volumes.

Their model thus describes the events corresponding to an increase/decrease of the bid/ask as a $4D$-Hawkes process marked by transaction volumes. By adding suitable constraints in order to avoid infinite spread, they calibrated the model on FX rates data by a maximum likelihood approach.

They have shown that their model is consistent with empirical data by reproducing the signature plots of the considered assets and the behavior of the high-frequency pair correlation function.

Outline

Leverage effect

The leverage effect accounts for the one-way correlation between volatility and returns.

Leverage effect

The leverage effect accounts for the one-way correlation between volatility and returns.

Definition. The leverage effect is defined, for $\tau > 0$, as

$$|\mathrm{Cov}(dP_t, \sigma_{t-\tau}^2)| < |\mathrm{Cov}(dP_{t-\tau}, \sigma_t^2)|$$

where $\mathrm{Cov}(dP_{t-\tau}, \sigma_t^2)$ is the covariance between past returns and future volatility (lagged by τ) and, conversely, $\mathrm{Cov}(dP_t, \sigma_{t-\tau}^2)$ is the covariance between past volatility and future returns.

Leverage effect

The leverage effect accounts for the one-way correlation between volatility and returns.

Definition. The leverage effect is defined, for $\tau > 0$, as

$$|\mathrm{Cov}(dP_t, \sigma^2_{t-\tau})| < |\mathrm{Cov}(dP_{t-\tau}, \sigma^2_t)|$$

where $\mathrm{Cov}(dP_{t-\tau}, \sigma^2_t)$ is the covariance between past returns and future volatility (lagged by τ) and, conversely, $\mathrm{Cov}(dP_t, \sigma^2_{t-\tau})$ is the covariance between past volatility and future returns. Intuitively, the leverage effect

can be understood by recognizing that large negative returns induce market panic, thereby increasing future volatility, while, high volatility does not necessarily lead to large future returns.

Zumbach effect

The Zumbach effect accounts for the asymmetry in the correlation between high frequency volatility and low frequency volatility.

Zumbach effect

The Zumbach effect accounts for the asymmetry in the correlation between high frequency volatility and low frequency volatility.

Definition. Let define aggregated returns of day t as

$$R_{t,\tau} := \begin{cases} P^{\text{close}}_{t-1} / P^{\text{open}}_{t-1+\tau} - 1 & \text{for } \tau < 0 \\ P^{\text{close}}_{t+1+\tau} / P^{\text{open}}_{t+1} - 1 & \text{for } \tau > 0 \end{cases}.$$

Low frequency volatility is then represented by $R^2_{t,\tau}$ and high frequency volatility of day t by σ_t.

Zumbach effect

The Zumbach effect accounts for the asymmetry in the correlation between high frequency volatility and low frequency volatility.

Definition. Let define aggregated returns of day t as

$$R_{t,\tau} := \begin{cases} P^{\text{close}}_{t-1}/P^{\text{open}}_{t-1+\tau} - 1 & \text{for } \tau < 0 \\ P^{\text{close}}_{t+1+\tau}/P^{\text{open}}_{t+1} - 1 & \text{for } \tau > 0 \end{cases}.$$

Low frequency volatility is then represented by $R^2_{t,\tau}$ and high frequency volatility of day t by σ_t. For $\tau > 0$, the Zumbach effect is expressed by

$$\mathrm{Cov}(\sigma_t^2, R^2_{t,\tau}) < \mathrm{Cov}(\sigma_t^2, R^2_{t,-\tau})$$

that is, past low frequency volatility has more influence on future high frequency volatility than conversely.

Zumbach effect

The Zumbach effect accounts for the asymmetry in the correlation between high frequency volatility and low frequency volatility.

Definition. Let define aggregated returns of day t as

$$R_{t,\tau} := \begin{cases} P^{\text{close}}_{t-1}/P^{\text{open}}_{t-1+\tau} - 1 & \text{for } \tau < 0 \\ P^{\text{close}}_{t+1+\tau}/P^{\text{open}}_{t+1} - 1 & \text{for } \tau > 0 \end{cases}.$$

Low frequency volatility is then represented by $R^2_{t,\tau}$ and high frequency volatility of day t by σ_t. For $\tau > 0$, the Zumbach effect is expressed by

$$\text{Cov}(\sigma_t^2, R^2_{t,\tau}) < \text{Cov}(\sigma_t^2, R^2_{t,-\tau})$$

that is, past low frequency volatility has more influence on future high frequency volatility than conversely.

Intuitively, the Zumbach effect captures the notion that a trending market (characterized by an increase of R^2_t over several days) heightens market participants' uncertainty, leading to increased volatility, while increased volatility does not necessarily result in a trending market.

Definition of QHawkes

A quadratic Hawkes process $(N_t)_{t\geq 0}$ is an inhomogenous Poisson process whose intensity is dependent on both past activity (dN_t) and past price returns defined as $dP_t := \varepsilon_t \psi dN_t$, where $\varepsilon_t = \pm 1$ is an unbiased random sign, independently chosen at each price change, and ψ is the tick size (size of elementary price changes).

Definition of QHawkes

A quadratic Hawkes process $(N_t)_{t\geq 0}$ is an inhomogenous Poisson process whose intensity is dependent on both past activity (dN_t) and past price returns defined as $dP_t := \varepsilon_t \psi dN_t$, where $\varepsilon_t = \pm 1$ is an unbiased random sign, independently chosen at each price change, and ψ is the tick size (size of elementary price changes). The Quadratic Hawkes process intensity is now defined as

$$\lambda_t = \lambda_0 + \frac{1}{\psi}\int_{-\infty}^{t} L(t-s)dP_s + \frac{1}{\psi^2}\int_{-\infty}^{t}\int_{-\infty}^{t} Q(t-s,t-u)dP_s dP_u,$$

where L is called the leverage kernel and Q the quadratic feedback kernel. Note that L and Q must be such that the quadratic form in dP_s is positive definite.

Definition of QHawkes

A quadratic Hawkes process $(N_t)_{t\geq 0}$ is an inhomogenous Poisson process whose intensity is dependent on both past activity (dN_t) and past price returns defined as $dP_t := \varepsilon_t \psi dN_t$, where $\varepsilon_t = \pm 1$ is an unbiased random sign, independently chosen at each price change, and ψ is the tick size (size of elementary price changes). The Quadratic Hawkes process intensity is now defined as

$$\lambda_t = \lambda_0 + \frac{1}{\psi}\int_{-\infty}^{t} L(t-s)dP_s + \frac{1}{\psi^2}\int_{-\infty}^{t}\int_{-\infty}^{t} Q(t-s, t-u)dP_s dP_u,$$

where L is called the leverage kernel and Q the quadratic feedback kernel. Note that L and Q must be such that the quadratic form in dP_s is positive definite.

Since $dP_t^2 = \psi^2 dN_t$, a purely diagonal Q recovers the standard Hawkes kernel. New effects arise when considering non-diagonal contributions to Q.

A special case: the ZHawkes model

As an interesting special case that captures the Zumbach effect (*i.e.* the correlation between future volatility and past trends),

A special case: the ZHawkes model

As an interesting special case that captures the Zumbach effect (*i.e.* the correlation between future volatility and past trends), Blanc et al. proposed the following ZHawkes specification

$$L \equiv 0, \quad Q(s,u) = \phi(s)\delta(s-u) + z(s)z(u),$$

i.e. a quadratic kernel that is diagonal plus a rank-one contribution.

A special case: the ZHawkes model

As an interesting special case that captures the Zumbach effect (*i.e.* the correlation between future volatility and past trends), Blanc et al. proposed the following ZHawkes specification

$$L \equiv 0, \quad Q(s,u) = \phi(s)\delta(s-u) + z(s)z(u),$$

i.e. a quadratic kernel that is diagonal plus a rank-one contribution.

In this case, the intensity of the QHawkes becomes

$$\lambda_t = \lambda_0 + H_t + Z_t^2,$$

where $H_t = \int_{-\infty}^t \phi(t-s)dN_s$ represents the Hawkes component of the intensity, $N_t - N_s := \frac{1}{\psi^2}(P_t - P_{t-})^2$, whereas $Z_t = \frac{1}{\psi}\int_{-\infty}^t z(t-s)dP_s$ represents the trend-induced (Zumbach) component.

Multivariate QHawkes processes

- Modeling multivariate volatility requires accounting for cross-time-reversal asymmetry effects and co-jumps occurrences, in addition to the characteristics of univariate financial time series.

Multivariate QHawkes processes

- Modeling multivariate volatility requires accounting for cross-time-reversal asymmetry effects and co-jumps occurrences, in addition to the characteristics of univariate financial time series.

- The QHawkes model is a suitable choice for modeling multivariate volatility as it inherently generates the characteristics of univariate financial time series. Additionally, its ability to explicitly delineate the influence of the past on future activity is particularly relevant for understanding cross-time-reversal asymmetry effects.

Multivariate QHawkes processes

- Modeling multivariate volatility requires accounting for cross-time-reversal asymmetry effects and co-jumps occurrences, in addition to the characteristics of univariate financial time series.

- The QHawkes model is a suitable choice for modeling multivariate volatility as it inherently generates the characteristics of univariate financial time series. Additionally, its ability to explicitly delineate the influence of the past on future activity is particularly relevant for understanding cross-time-reversal asymmetry effects.

- Multivariate QHawkes processes seem to consistently extend the relevant features of the univariate QHawkes: fat tails of the returns distribution, volatility clustering, cross-time-reversal asymmetry.

Outline

First Level-I model with Hawkes process

The Level-I book description concerns the events that exclusively occur at the best bid and best ask levels of the order book.

First Level-I model with Hawkes process

The Level-I book description concerns the events that exclusively occur at the best bid and best ask levels of the order book.

Large (2007) formalized the ability of LOB to replenish after being depleted by a large trade (book resiliency) by quantifying it by the way large trades alter future intensities of order occurrences. To this end, he introduced the response kernel

$$G^{ij}(t) = \mathbb{E}\left[dN^i(t)|\mathcal{F}_0, dN^j(t) = 1\right] - \mathbb{E}\left[dN^i(t)|\mathcal{F}_0\right].$$

First Level-I model with Hawkes process

The Level-I book description concerns the events that exclusively occur at the best bid and best ask levels of the order book.

Large (2007) formalized the ability of LOB to replenish after being depleted by a large trade (book resiliency) by quantifying it by the way large trades alter future intensities of order occurrences. To this end, he introduced the response kernel

$$G^{ij}(t) = \mathbb{E}\left[dN^i(t)|\mathcal{F}_0, dN^j(t) = 1\right] - \mathbb{E}\left[dN^i(t)|\mathcal{F}_0\right].$$

With data from LSE, he modelled as a $10D$-Hawkes processes:

- MOs and LOs that move the mid-prices (4 components with bid/ask),
- MOs and LOs that leave the book unchanged (4 components),
- cancel orders (2 components).

Large used MLE within the class of exponential kernels in order to estimate Ψ defined by the causal solution of $\Phi(t) + \Psi(t) * \Phi(t) = \Psi(t)$.

A second Level-I model

A similar analysis was conducted by Bacry and *al.* (2016) where the dynamics of event occurrence has then been modelled as a $8D$-Hawkes process.

A second Level-I model

A similar analysis was conducted by Bacry and *al.* (2016) where the dynamics of event occurrence has then been modelled as a $8D$-Hawkes process. More precisely

$$N_t = \left(P_t^{(a)}, P_t^{(b)}, T_t^{(a)}, T_t^{(b)}, L_t^{(a)}, L_t^{(b)}, C_t^{(a)}, C_t^{(b)}\right),$$

A second Level-I model

A similar analysis was conducted by Bacry and *al.* (2016) where the dynamics of event occurrence has then been modelled as a $8D$-Hawkes process. More precisely

$$N_t = \left(P_t^{(a)}, P_t^{(b)}, T_t^{(a)}, T_t^{(b)}, L_t^{(a)}, L_t^{(b)}, C_t^{(a)}, C_t^{(b)}\right),$$

where

- $P^{(a)}$ (resp. $P^{(b)}$) counts the number of upward (resp. downward) mid-price moves;

A second Level-I model

A similar analysis was conducted by Bacry and *al.* (2016) where the dynamics of event occurrence has then been modelled as a $8D$-Hawkes process. More precisely

$$N_t = \left(P_t^{(a)}, P_t^{(b)}, T_t^{(a)}, T_t^{(b)}, L_t^{(a)}, L_t^{(b)}, C_t^{(a)}, C_t^{(b)}\right),$$

where

- $P^{(a)}$ (resp. $P^{(b)}$) counts the number of upward (resp. downward) mid-price moves;
- $T^{(a)}$ (resp. $T^{(b)}$) counts the number of market orders at the ask (resp. bid) that do not move the price;

A second Level-I model

A similar analysis was conducted by Bacry and *al.* (2016) where the dynamics of event occurrence has then been modelled as a $8D$-Hawkes process. More precisely

$$N_t = \left(P_t^{(a)}, P_t^{(b)}, T_t^{(a)}, T_t^{(b)}, L_t^{(a)}, L_t^{(b)}, C_t^{(a)}, C_t^{(b)}\right),$$

where

- $P^{(a)}$ (resp. $P^{(b)}$) counts the number of upward (resp. downward) mid-price moves;
- $T^{(a)}$ (resp. $T^{(b)}$) counts the number of market orders at the ask (resp. bid) that do not move the price;
- $L^{(a)}$ (resp. $L^{(b)}$) counts the number of limit orders at the ask (resp. bid) that do not move the price;

A second Level-I model

A similar analysis was conducted by Bacry and *al.* (2016) where the dynamics of event occurrence has then been modelled as a $8D$-Hawkes process. More precisely

$$N_t = \left(P_t^{(a)}, P_t^{(b)}, T_t^{(a)}, T_t^{(b)}, L_t^{(a)}, L_t^{(b)}, C_t^{(a)}, C_t^{(b)}\right),$$

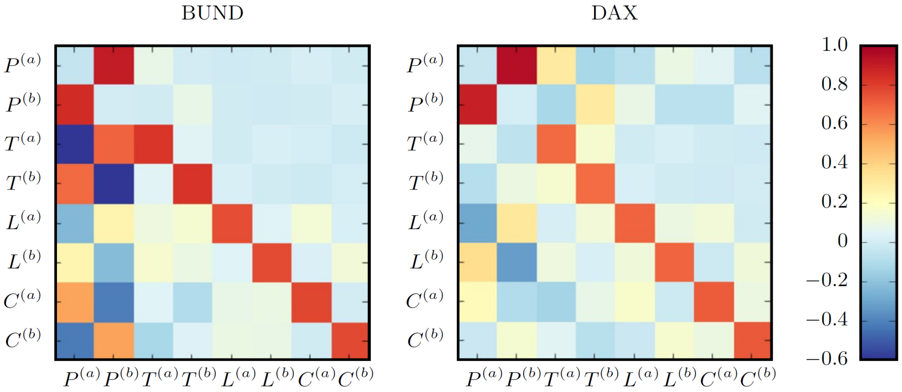
where

- $P^{(a)}$ (resp. $P^{(b)}$) counts the number of upward (resp. downward) mid-price moves;
- $T^{(a)}$ (resp. $T^{(b)}$) counts the number of market orders at the ask (resp. bid) that do not move the price;
- $L^{(a)}$ (resp. $L^{(b)}$) counts the number of limit orders at the ask (resp. bid) that do not move the price;
- $C^{(a)}$ (resp. $C^{(b)}$) counts the number of cancel orders at the ask (resp. bid) that do not move the price.

Results of this second model

- The book event dynamics is mainly self-exciting, except for the mid-price changes for which cross-excitation effects are strongly dominating (see the following figure): one can see that the resulting matrix is mainly diagonal except in the price sub-block, that is anti-diagonal.
- Price changes events are mainly triggered by price change events, but this effect is mostly anti-diagonal.
- The market is highly endogenous and all the dominating kernels are slowly decreasing, well described by a power-law behavior.
- They also provided evidence of some inhibitory effects which result from negative values of some Hawkes kernels. For example it was observed that, for a large tick asset, an upward price move not only triggers forthcoming trades at bid but also inhibits trades on the ask side.

Color map of the matrix of Hawkes kernel norms



Empirical determination of the matrix of Hawkes kernel norms in the 8D-model of level-I book events. This figure is reproduced from Bacry and *al.* (2016).

Outline

Full order book model I

Muni Toke (2010) introduces a 2-agents model as follows

Full order book model I

Muni Toke (2010) introduces a 2-agents model as follows

- A liquidity provider:
  - Arrivals of LO are modeled by using either a pure homogeneous Poisson process or a $1D$-Hawkes process with an exponential kernel.

Full order book model I

Muni Toke (2010) introduces a 2-agents model as follows

- A liquidity provider:
  - Arrivals of LO are modeled by using either a pure homogeneous Poisson process or a $1D$-Hawkes process with an exponential kernel.
  - The price-levels of new limit orders are randomly chosen by first choosing the (bid or ask) side with probability $1/2$ and then by sampling their values from a Student distribution.

Full order book model I

Muni Toke (2010) introduces a 2-agents model as follows

- A liquidity provider:
  - Arrivals of LO are modeled by using either a pure homogeneous Poisson process or a $1D$-Hawkes process with an exponential kernel.
  - The price-levels of new limit orders are randomly chosen by first choosing the (bid or ask) side with probability $1/2$ and then by sampling their values from a Student distribution.
- A liquidity taker:
  - Modeled using either a pure homogeneous Poisson process or a 1-dimensional Hawkes process with an exponential kernel.

Full order book model I

Muni Toke (2010) introduces a 2-agents model as follows

- A liquidity provider:
  - Arrivals of LO are modeled by using either a pure homogeneous Poisson process or a $1D$-Hawkes process with an exponential kernel.
  - The price-levels of new limit orders are randomly chosen by first choosing the (bid or ask) side with probability $1/2$ and then by sampling their values from a Student distribution.
- A liquidity taker:
  - Modeled using either a pure homogeneous Poisson process or a 1-dimensional Hawkes process with an exponential kernel.

All the volumes of both limit and market orders are i.i.d. and exponentially distributed variables. Hawkes processes use exponential kernels which are estimated parametrically using MLE.

Results of this model

- Inter-trade times of MOs are shown to be much better modelled by a $1D$-Hawkes process (MM) than by a pure Homogeneous Poisson process (HP).

Results of this model

- Inter-trade times of MOs are shown to be much better modelled by a $1D$-Hawkes process (MM) than by a pure Homogeneous Poisson process (HP).
- Inter-time of LOs are shown to be much better modelled by a $1D$-Hawkes process (LL) than by HP.

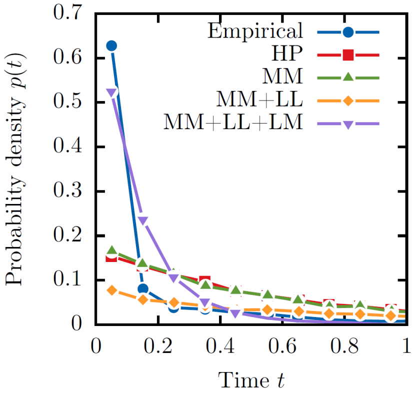
Results of this model

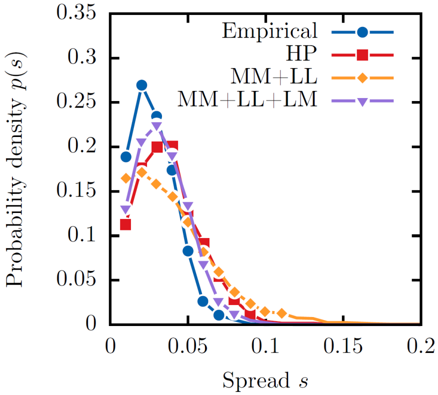
- Inter-trade times of MOs are shown to be much better modelled by a $1D$-Hawkes process (MM) than by a pure Homogeneous Poisson process (HP).
- Inter-time of LOs are shown to be much better modelled by a $1D$-Hawkes process (LL) than by HP.
- Inter-time between a given MO and the first arrival time of a LO after that MO is much better reproduced by a $2D$-Hawkes (MM+LL+LM) model with self-exciting kernels for both market (MM) and limit orders (LL) and a single cross-exciting kernel which corresponds to the influence of past MOs on future LOs (LM).

Results of this model

- Inter-trade times of MOs are shown to be much better modelled by a $1D$-Hawkes process (MM) than by a pure Homogeneous Poisson process (HP).
- Inter-time of LOs are shown to be much better modelled by a $1D$-Hawkes process (LL) than by HP.
- Inter-time between a given MO and the first arrival time of a LO after that MO is much better reproduced by a $2D$-Hawkes (MM+LL+LM) model with self-exciting kernels for both market (MM) and limit orders (LL) and a single cross-exciting kernel which corresponds to the influence of past MOs on future LOs (LM).
- Finally the right panel of next figure shows the so-obtained distribution of the bid-ask spread for empirical data, the pure Homogenous Poisson (HP) model, the MM+LL Hawkes model and the MM+LL+LM model.

Empirical density function





Empirical density function of the distribution of the duration between a given market order and the first arrival of a limit order after that market order. Reproduced from Muni Toke (2010).

Full order book model II

Abergel and Lu (2018) introduced a point-process-based order book model, building on the 12 event types

$$E = \{L^0_{buy}, L^0_{sell}, C^0_{buy}, C^0_{sell}, M^0_{buy}, M^0_{sell}, \\ L^1_{buy}, L^1_{sell}, C^1_{buy}, C^1_{sell}, M^1_{buy}, M^1_{sell}\},$$

where L stands for LO, M for MO, C for cancellation, 0 for orders that does not change the mid-price and 1 for orders that impact the mid-price.

Full order book model II

Abergel and Lu (2018) introduced a point-process-based order book model, building on the 12 event types

$$\begin{aligned} E = \{ & L^0_{buy}, L^0_{sell}, C^0_{buy}, C^0_{sell}, M^0_{buy}, M^0_{sell}, \\ & L^1_{buy}, L^1_{sell}, C^1_{buy}, C^1_{sell}, M^1_{buy}, M^1_{sell} \}, \end{aligned}$$

where L stands for LO, M for MO, C for cancellation, 0 for orders that does not change the mid-price and 1 for orders that impact the mid-price.

- They consider both linear and non-linear Hawkes processes with exponential kernels.

Full order book model II

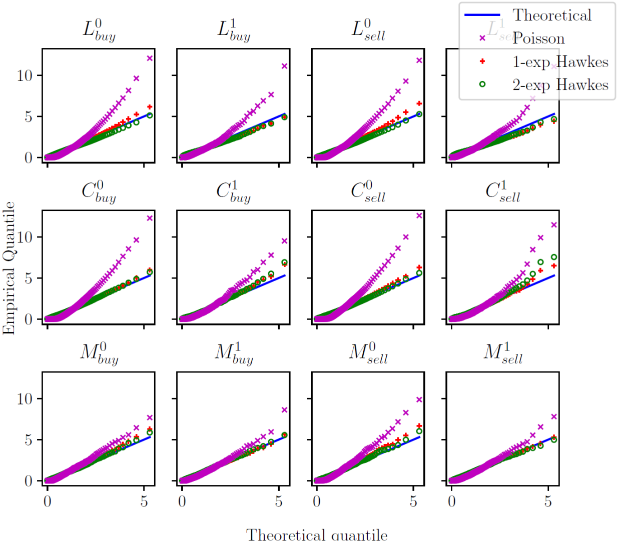
Abergel and Lu (2018) introduced a point-process-based order book model, building on the 12 event types

$$\begin{aligned} E \;=\; \{ & L^0_{buy}, L^0_{sell}, C^0_{buy}, C^0_{sell}, M^0_{buy}, M^0_{sell}, \\ & L^1_{buy}, L^1_{sell}, C^1_{buy}, C^1_{sell}, M^1_{buy}, M^1_{sell} \}, \end{aligned}$$

where L stands for LO, M for MO, C for cancellation, 0 for orders that does not change the mid-price and 1 for orders that impact the mid-price.

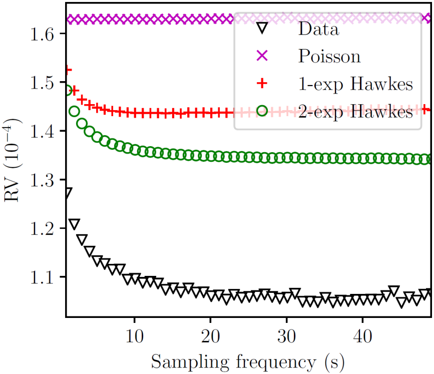
- They consider both linear and non-linear Hawkes processes with exponential kernels.
- They consider two criteria to evaluate the performances of their model: a goodness-of-fit criterion for the distribution of forward recurrence times, and a criterion based on the signature plot generated by the model.

Q-Q plot goodness of fit tests of linear order book models.



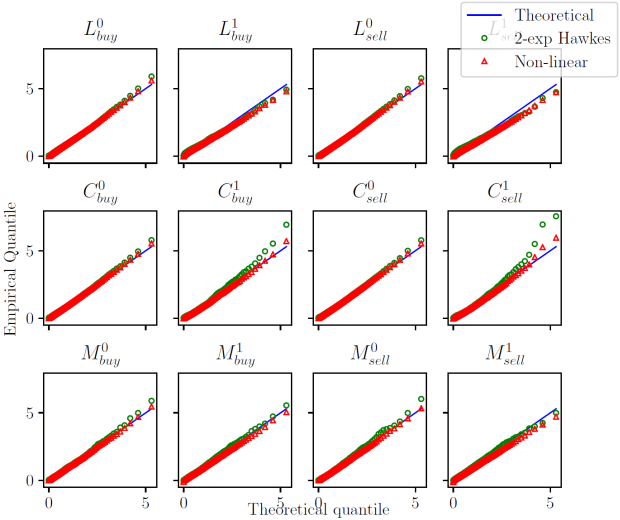
Reproduced from Abergel and Lu (2018).

Signature plots of linear order book models.



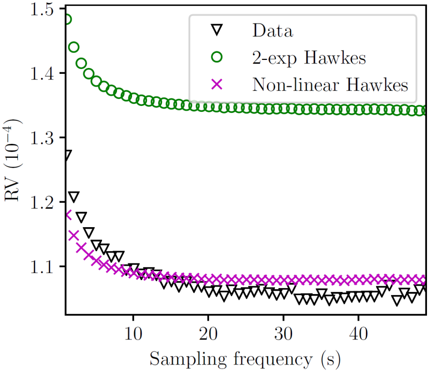
Mean signature plots of simulated price compared with real data for ALLIANZ SE. Reproduced from Abergel and Lu (2018).

Q-Q plot goodness of fit tests of non-linear order book models.



Reproduced from Abergel and Lu (2018).

Signature plots of non-linear order book models.



Mean signature plots of simulated price compared with real data for ALLIANZ SE. Reproduced from Abergel and Lu (2018).

References

Abergel, F. and Lu, X., Limit order book modelling with high dimensional Hawkes processes, *Quantitative Finance*, 18(2):249–264, 2018.

Bacry, E., Delattre, S., Hoffmann, M., Muzy, J. F., Modeling microstructure noise with mutually exciting point processes, *Quantitative finance*, 13(1):65–77, 2013.

Bacry, E., Dayri, K. and Muzy, J.F., Non-parametric kernel estimation for symmetric Hawkes processes. Application to high frequency financial data., *Eur. Phys. J. B*, 85:157, 2012.

Bacry, E. and Muzy, J.F., Hawkes model for price and trades high-frequency dynamics, *Quantitative Finance*, 14(7):1–20, 2014.

Bacry, E., Jaisson, T. and Muzy, J.F., Estimation of slowly decreasing Hawkes kernels: Application to high frequency order book modelling, *Quantitative Finance*, 16(8):1179–1201, 2016.

Blanc P., Donier J., and Bouchaud J.-P., Quadratic hawkes processes for financial prices, *Quantitative Finance*, 17(2):171–188, 2017.

Bowsher, C. G., Modeling security market events in continuous time: Intensity based, multivariate point process models, *Journal of Econometrics*, 141(2):876–912, 2007.

Hawkes, A. G., Spectra of some self-exciting and mutually exciting point processes, *Biometrika*, 58(1):83–90, 1971.

Jaisson, T. and Rosenbaum, M., Limit theorems for nearly unstable hawkes processes, *The Annals of Applied Probability*, 25(2):600–631, 2015.

References

- Kyle, A., Continuous auctions and insider trading, *Econometrica*, 53(6):1315–1336, 1985.
- Large, J., Measuring the resiliency of an electronic limit order book, *Journal of Financial Markets*, 10(1):1–25, 2007.
- Laruelle, S. and Lehalle, C.-A., *Market Microstructure in Practice*, World Scientific, 2018.
- Lewis, P. A. W., Shedler, G. S., Simulation of nonhomogeneous Poisson processes by thinning, *Naval Research Logistics Quarterly*, 26(3), 403–413, 1979.
- Muni Toke, I., Market making in an order book model and its impact on the bid-ask spread, *Econophysics of Order-Driven Markets, New Economic Windows*, Springer-Verlag Milan, pp.49–64, 2011.
- Ogata, Y., The asymptotic behaviour of maximum likelihood estimators for stationary point processes, *Annals of the Institute of Statistical Mathematics*, 30(1), 243–261, 1978.
- Ogata, Y., On Lewis' simulation method for point processes, *IEEE Transactions on Information Theory*, 27(1), 23–31, 1981.
- Sahalia, Y.A., Cacho-Diaz, J. and Laeven, R.J.A., Modeling financial contagion using mutually exciting jump processes, *Journal of Financial Economics*, 117(3):585–606, 2015.
- Zumbach, G., Time reversal invariance in finance, *Quantitative Finance*, 9(5):505–515, 2009.
- Zumbach, G., Volatility conditional on price trends, *Quantitative Finance*, 10(4):431–442, 2010.