Limit theorems for nearly unstable Hawkes processes

Thibault Jaisson and Mathieu Rosenbaum

École Polytechnique and University Paris 6

19 March 2014

Outline

Outline

Definition

Hawkes process

- A Hawkes process $(N_t)_{t\geq 0}$ is a self exciting point process, whose intensity at time t, denoted by λ_t, is of the form

$$\lambda_t = \mu + \sum_{0<J_i<t} \phi(t - J_i) = \mu + \int_{(0,t)} \phi(t-s)dN_s,$$

  where μ is a positive real number, ϕ a regression kernel and the J_i are the points of the process before time t.
- These processes have been introduced in 1971 by Hawkes in the purpose of modeling earthquakes and their aftershocks.

Hawkes processes in finance

Hawkes processes as a tool for modeling :

- Midquotes and transaction prices : Bowsher (07), Bauwens and Hautsch (04), Hewlett (06), Bacry, Delattre, Hoffmann, Muzy (13).
- Order books : Large (07).
- Daily data analysis : Embrechts, Liniger, Lin (11).
- Financial contagion : Aït-Sahalia, Cacho-Diaz, Laeven (10).
- Credit Risk : Errais, Giesecke, Goldberg (10).

Popularity of Hawkes processes in finance

Two main reasons for the popularity of Hawkes processes

- These processes represent a very natural and tractable extension of Poisson processes. In fact, comparing point processes and conventional time series, Poisson processes are often viewed as the counterpart of iid random variables whereas Hawkes processes play the role of autoregressive processes.
- Another explanation for the appeal of Hawkes processes is that it is often easy to give a convincing interpretation to such modeling. To do so, the branching structure of Hawkes processes is quite helpful.

Hawkes processes as a population model

Poisson cluster representation

- Under the assumption $\|\phi\|_1 < 1$, where $\|\phi\|_1$ denotes the L^1 norm of ϕ, Hawkes processes can be represented as a population process where migrants arrive according to a Poisson process with parameter μ.
- Then each migrant gives birth to children according to a non homogeneous Poisson process with intensity function ϕ, these children also giving birth to children according to the same non homogeneous Poisson process, see Hawkes (74).
- Now consider for example the classical case of buy (or sell) market orders. Then migrants can be seen as exogenous orders whereas children are viewed as orders triggered by other orders.

Stability condition

The condition $\|\phi\|_1 < 1$

- The assumption $\|\phi\|_1 < 1$ is crucial in the study of Hawkes processes.
- If one wants to get a stationary intensity with finite first moment, then the condition $\|\phi\|_1 < 1$ is required.
- This condition is also necessary in order to obtain classical ergodic properties for the process.
- For these reasons, this condition is often called stability condition in the Hawkes literature.

$\|\phi\|_1$ in practice

Degree of endogeneity of the market

- From a practical point of view, a lot of interest has been recently devoted to the parameter $\|\phi\|_1$.
- For example, Hardiman, Bercot and Bouchaud (13) and Filimonov and Sornette (12,13) use the branching interpretation of Hawkes processes on midquote data in order to measure the so-called degree of endogeneity of the market, defined by $\|\phi\|_1$.

$\|\phi\|_1$ in practice

Degree of endogeneity of the market

- The parameter $\|\phi\|_1$ corresponds to the average number of children of an individual, $\|\phi\|_1^2$ to the average number of grandchildren of an individual,... Therefore, if we call cluster the descendants of a migrant, then the average size of a cluster is given by $\sum_{k\geq 1}\|\phi\|_1^k = \|\phi\|_1/(1-\|\phi\|_1)$.
- Thus, the average proportion of endogenously triggered events is $\|\phi\|_1/(1-\|\phi\|_1)$ divided by $1+\|\phi\|_1/(1-\|\phi\|_1)$, which is equal to $\|\phi\|_1$.

$\|\phi\|_1$ in practice

Unstable Hawkes processes

- This branching ratio can be measured using parametric and non parametric estimation methods for Hawkes processes, see Ogata (78,83) for likelihood based methods and Reynaud-Bouret and Schbath (10), Al Dayri *et al.* (11) and Bacry and Muzy (13) for functional estimators of the function ϕ.
- In Hardiman, Bercot and Bouchaud (13), very stable estimations of $\|\phi\|_1$ are reported for the E mini S&P futures between 1998 and 2012, the results being systematically close to one.
- This is also the case for Bund and Dax futures in Al Dayri *et al.* (11) and various other assets in Filimonov and Sornette (12).

About $\|\phi\|_1$ close to one

An intuitive explanation

- This stylized statistical result should definitely worry users of Hawkes processes. Indeed, it is rarely suitable to apply a statistical model where the parameters are pushed to their limits.
- In fact, these obtained values for $\|\phi\|_1$ on empirical data are not really surprising.

About $\|\phi\|_1$ close to one

An intuitive explanation

- Indeed, one of the most well documented stylized fact in high frequency finance is the persistence (or long memory) in flows and market activity measures, see for example Bouchaud *et al* (04), Lillo *et al* (04).
- Usual Hawkes processes, in the same way as autoregressive processes, can only exhibit short range dependence, failing to reproduce this classical empirical feature.

Dealing with $\|\phi\|_1$ close to one

The Brémaud-Massoulié approach

- In spite of their relative inadequacy with market data, Hawkes processes possess so many appealing properties that one could still try to apply them in some specific situations.
- In Hardiman, Bercot and Bouchaud (13), it is suggested to use the without ancestors version of Hawkes processes introduced by Brémaud and Massoulié (01).
- For such processes, $\|\phi\|_1 = 1$ but, in order to preserve stationarity and a finite expectation for the intensity, one needs to have $\mu = 0$.

Dealing with $\|\phi\|_1$ close to 1

The Brémaud-Massoulié approach

- This is probably a relevant approach. However setting the parameter μ to 0 is not completely satisfying since this parameter has a nice interpretation (exogenous orders).
- Moreover it is not found to be equal to zero in practice, see Hardiman, Bercot and Bouchaud (13).
- Finally, a time-varying μ is an easy way to reproduce seasonalities observed on the market, see Bacry and Muzy (13) (however, for simplicity, we work here with a constant $\mu > 0$).

Aim of this work

Limiting behavior of Hawkes processes

- These empirical measures of $\|\phi\|_1$, close to 1, are the starting point of this work.
- Our aim is to study the behavior at large time scales of nearly unstable Hawkes processes, which correspond to these estimations.
- More precisely, we consider a sequence of Hawkes processes observed on $[0, T]$, where T goes to infinity.

Limiting behavior of Hawkes processes in the classical case

Law of large numbers and central limit theorem

- In the case of a fixed kernel (not depending on T) with norm strictly smaller than one, scaling limits of Hawkes processes have been investigated in Bacry *et al.* (13).
- They obtain a deterministic limit for the properly normalized sequence of Hawkes processes, as it is the case for suitably rescaled Poisson processes.
- In their price model consisting in the difference of two Hawkes processes, a Brownian motion (with some volatility) is found at the limit.
- These two results are in fact quite intuitive. Indeed, in the same way as Poisson processes and autoregressive models, Hawkes processes enjoy short memory properties.

In this work

Limiting behavior of nearly unstable Hawkes processes

- We show that when the Hawkes processes are nearly unstable, these weakly dependent-like behaviors are no longer observed at intermediate time scales.
- To do so, we consider that the kernels of the Hawkes processes depend on T.
- More precisely, we translate the near instability condition into the assumption that the norm of the kernels tends to 1 as the observation scale T goes to infinity.

Outline

Assumptions and asymptotic framework

Sequence of Hawkes processes

- We consider a sequence of point processes $(N_t^T)_{t \geq 0}$ indexed by T. We have $N_0^T = 0$ and the process is observed on the time interval $[0, T]$. Furthermore, our asymptotic setting is that the observation scale T goes to infinity.
- The intensity process (λ_t^T) is defined for $t \geq 0$ by

$$\lambda_t^T = \mu + \int_0^t \phi^T(t-s)dN_s^T,$$

where μ is a positive real number and ϕ^T a non negative measurable function on $\mathbb{R}^+$ which satisfies $\|\phi^T\|_1 < 1$.

Assumptions and asymptotic framework

Sequence of Hawkes processes

- For any $0 \leq a < b \leq T$ and $A \in \mathcal{F}_a^T$

$$\mathbb{E}[(N_b^T - N_a^T)1_A] = \mathbb{E}[\int_a^b \lambda_s^T 1_A ds].$$

- If we denote by $(J_n^T)_{n\geq 1}$ the jump times of (N_t^T), the process

$$N_{t\wedge J_n^T}^T - \int_0^{t\wedge J_n^T} \lambda_s^T ds$$

is a martingale.

Assumptions and asymptotic framework

Assumptions on the kernel

- For $t \in \mathbb{R}^+$, we assume

$$\phi^T(t) = a_T \phi(t),$$

where $(a_T)_{T \geq 0}$ converges to 1 such that, $a_T < 1$ and ϕ is a non negative measurable function such that

$$\int_0^{+\infty} \phi(s)ds = 1 \text{ and } \int_0^{+\infty} s\phi(s)ds = m < \infty.$$

- The case where $\|\phi^T\|_1$ is larger than 1 corresponds to the situation where the stability condition is violated. Our framework is a way to get close to instability. Therefore we call our processes nearly unstable Hawkes processes.

Observation scales

Degeneracy of the parameters

- In our framework, two parameters degenerate at infinity : T and $(1 - a_T)^{-1}$.
- The relationship between these two sequences will determine the scaling behavior of the sequence of Hawkes processes.
- Recall that it is shown in Bacry *et al.* (13) that when $\|\phi\|_1$ is fixed and smaller than one, after appropriate scaling, the limit of the sequence of Hawkes processes is deterministic, as it is for example the case for Poisson processes.
- In our setting, if $1 - a_T$ tends "slowly" to zero, we can expect the same result. Indeed, we may have T large enough so that we reach the asymptotic regime and for such T, a_T is still sufficiently far from unity.

The case a_T converges slowly to one

Theorem

Assume $T(1 - a_T) \to +\infty$. Then the sequence of Hawkes processes is asymptotically deterministic, in the sense that the following convergence in L^2 holds :

$$\sup_{v \in [0,1]} \frac{1 - a_T}{T} |N^T_{Tv} - \mathbb{E}[N^T_{Tv}]| \to 0.$$

Intermediate cases

Unstable behavior

- On the contrary, if $1 - a_T$ tends too rapidly to zero, the situation is likely to be quite intricate.
- Indeed, for given T, the Hawkes process may already be very close to instability whereas T is not large enough to reach the asymptotic regime.
- The last case, which is probably the most interesting one, is the intermediate case, where $1 - a_T$ tends to zero in such a manner that a non deterministic scaling limit is obtained, while not being in the preceding degenerate setting.

Non degenerate limit for nearly unstable Hawkes processes

Martingale process

- Let M^T be the martingale process associated to N^T, that is, for $t \geq 0$,

$$M_t^T = N_t^T - \int_0^t \lambda_s^T ds.$$

- We also set ψ^T the function defined on $\mathbb{R}^+$ by

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

- We can show that

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

Non degenerate limit for nearly unstable Hawkes processes

Rescaling

- We rescale our processes so that they are defined on $[0,1]$. To do that, we consider for $t \in [0,1]$, we write

$$\lambda^T_{tT} = \mu + \int_0^{tT} \psi^T(Tt-s)\mu ds + \int_0^{tT} \psi^T(Tt-s)dM^T_s.$$

- For the scaling in space, a natural multiplicative factor is $(1-a_T)$. Indeed, in the stationary case, the expectation of λ^T_t is $\mu/(1-\|\phi^T\|_1)$. Thus, the order of magnitude of the intensity is $(1-a_T)^{-1}$. This is why we define

$$C^T_t = \lambda^T_{tT}(1-a_T).$$

Non degenerate limit for nearly unstable Hawkes processes

The function ψ^T

- The asymptotic behavior of C_t^T is closely linked to that of ψ^T.
- Remark that the function defined for $x \geq 0$ by

$$\rho^T(x) = T \frac{\psi^T(Tx)}{\|\psi^T\|_1}$$

is the density of the random variable

$$X^T = \frac{1}{T} \sum_{i=1}^{I^T} X_i,$$

where the (X_i) are iid random variables with density ϕ and I^T is a geometric random variable with parameter $1 - a_T$.

Non degenerate limit for nearly unstable Hawkes processes

The function ψ^T

- The characteristic function of the random variable X^T, denoted by $\widehat{\rho}^T$, satisfies :

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

  where $\hat{\phi}$ denotes the characteristic function of X_1.

- Since

$$\int_0^{+\infty} s\phi(s)ds = m < \infty,$$

  the function $\hat{\phi}$ is continuously differentiable with first derivative at point zero equal to *im*.

Non degenerate limit for nearly unstable Hawkes processes

The function ψ^T

- Therefore, using that a_T and $\hat{\phi}(\frac{z}{T})$ both tend to 1 as T goes to infinity, $\widehat{\rho}^T(z)$ is equivalent to

$$\frac{1}{1-\frac{izm}{T(1-a_T)}}.$$

- Thus, we precisely see here that the suitable regime so that we get a non trivial limiting law for X^T is that there exists $\lambda > 0$ such that

$$T(1-a_T) \underset{T\to+\infty}{\to} \lambda. \qquad (1)$$

- From now on, we assume (1) holds.

Non degenerate limit for nearly unstable Hawkes processes

Proposition

The sequence X^T converges in law towards an exponential random variable with parameter λ/m.

Non degenerate limit for nearly unstable Hawkes processes

Intuition for the result

- We set $u_T = T(1-a_T)/\lambda$ (so that u_T goes to 1). We have

$$\psi^T(Tx) = \rho^T(x)\frac{a_T}{\lambda u_T} \approx \frac{\lambda}{m}e^{-x\frac{\lambda}{m}}\frac{1}{\lambda} = \frac{1}{m}e^{-x\frac{\lambda}{m}}.$$

- Also,

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

with $B_t^T = \frac{1}{\sqrt{T}}\sqrt{u_T}\int_0^{tT}\frac{dM_s^T}{\sqrt{\lambda_s^T}}$.

Non degenerate limit for nearly unstable Hawkes processes

Intuition for the result

- The process B_t^T will be shown to converge towards a Brownian motion.
- So, heuristically replacing B^T by a Brownian motion B and $\psi^T(Tx)$ by $\frac{1}{m}e^{-x\frac{\lambda}{m}}$ in the preceding equation, we get

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

This gives :

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

which is the SDE satisfied by a CIR process.

Non degenerate limit for nearly unstable Hawkes processes

Theorem

The sequence of renormalized Hawkes intensities (C_t^T) converges in law, for the Skorohod topology, towards the law of the unique strong solution of the following Cox Ingersoll Ross stochastic differential equation on $[0,1]$:

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

Non degenerate limit for nearly unstable Hawkes processes

Theorem (continued)

Furthermore, the sequence of renormalized Hawkes process

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

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

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

Non degenerate limit for nearly unstable Hawkes processes

Discussion

- CIR type processes also appear as limit in law for critical branching processes with immigration.
- Our theorem implies that when $\|\phi\|_1$ is close to 1, if the observation time T is suitably chosen (that is of order $1/(1 - \|\phi\|_1)$), a non degenerate behavior (neither explosive, nor deterministic) can be obtained for a rescaled Hawkes process.
- This can for example be useful for the statistical estimation of the parameters of a Hawkes process.

Non degenerate limit for nearly unstable Hawkes processes

Discussion

- Indeed, designing an estimating procedure based on the properties of Hawkes processes at fine scales is a very hard task : Non parametric methods are difficult to use and suffer from various instabilities, whereas parametric approaches are of course very sensitive to model specifications.
- Considering an intermediate scale, where the process behaves like a CIR model, one can use statistical methods specifically developed in order to estimate CIR parameters.
- Of course, only the parameters λ, m and μ can be recovered this way.

Non degenerate limit for nearly unstable Hawkes processes

Discussion

- CIR processes are a very classical way to model stochastic (squared) volatilities in finance, see the celebrated Heston model.
- Also, it is widely acknowledged that there exists a linear relationship between the cumulated order flow and the integrated squared volatility, see for example Wyart *et al.* (08). Therefore, our setting where $\|\phi\|_1$ is close to 1 and the limiting behavior obtained seem in good agreement with market data.

Non degenerate limit for nearly unstable Hawkes processes

Discussion

- For the stationary version of a Hawkes process, one can show that the variance of N^T_T is of order $T(1 - \|\phi^T\|_1)^{-3}$.
- Thus, if $T(1 - a_T)$ tends to zero, that is $\|\phi^T\|_1$ goes rapidly to one, then the variance of $\frac{(1-a_T)}{T} N^T_T$ blows up as T goes to infinity. This situation is therefore very different from the one studied here.

Non degenerate limit for nearly unstable Hawkes processes

Discussion

- The assumption $\int_0^{+\infty} s\phi(s)ds < +\infty$ is crucial in order to approximate ψ^T by an exponential function.
- Let us now consider the fat tail case where ϕ is of order $\frac{1}{x^{1+\alpha}}$, $0 < \alpha < 1$, as x goes to infinity.
- After heuristic computations, we may expect that provided $T^{\alpha}(1 - a_T) \to \lambda$, the sequence

$$C_t^T = T^{1-\alpha}\lambda_{tT}^T(1 - a_T)$$

converges to the law of

$$X_t = \mu \int_0^t \phi^{\alpha}_{\frac{\sigma}{\lambda}}(t - s)ds + \int_0^t \phi^{\alpha}_{\frac{\sigma}{\lambda}}(t - s)\frac{1}{\sqrt{\lambda}}\sqrt{X_s}dB_s.$$

Ideas about the proof of the main theorem

Main steps

- Recall that

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

- We write

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

Ideas about the proof of the main theorem

Main steps

- Then we show that the following pseudo SDE holds :

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

with

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

Ideas about the proof of the main theorem

Main steps

- We show the ucp convergence to zero of the process U_t^T. To do so, we need an accurate control of the moments of the increments of the process, which is quite intricate.
- Then we pass to the limit in the sequence of pseudo SDEs.
- To do so, we use convergence theorems due to Jakubowsi, Mémin and Pagès (89) and Kurtz and Protter (91).

Outline

The model

Bidimensional Hawkes process

- We consider the model for the midprice of Bacry *et al* (11) :

$$P_t^T = N_t^{T+} - N_t^{T-},$$

with (N^{T+}, N^{T-}) a bidimensional Hawkes process with intensity

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

The model

Assumption

- We assume

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

where $(a_T)_{T \geq 0}$ is a sequence of positive numbers converging to one such that for all T, $a_T < 1$ and ϕ_1 and ϕ_2 such that

$$\int_0^{+\infty} \phi_1(s) + \phi_2(s) ds = 1 \text{ and } \int_0^{+\infty} s\big(\phi_1(s) + \phi_2(s)\big) ds = m.$$

The model

Properties of the model

- The preceding model takes into account the discreteness and the negative autocorrelation of the prices at the microstructure level.
- Moreover, it is shown in Bacry *et al.* that when one considers this price at large time scales, the stability condition implies that after suitable renormalization, it converges towards a Brownian motion (with a given volatility).

Convergence to a Heston model

Theorem

Let $\phi = \phi_1 - \phi_2$. The renormalized process

$$P_t^T = \frac{1}{T}(N_{Tt}^{T+} - N_{Tt}^{T-})$$

converges in law, for the Skorohod topology, towards a Heston type process P on $[0, 1]$ defined by :

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

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