Journées de Probabilités 2021
Malliavin-Stein Method for compound Hawkes processes

Mahmoud Khabou

Joint work with C.Hillairet, L.Huang and A.Réveillac

Institut de mathématiques de Toulouse

24 June 2021

Table of Contents

Counting process and intensity

Definition (Simple counting process)

A stochastic process $(H_t)_{t\in\mathbb{R}_+}$ is called a simple counting process if

- $H_t \geq 0$ for any $t \geq 0$,
- H is non-decreasing,
- H has jumps of size 1.

Giving a simple counting process is equivalent to giving an infinite increasing sequence of jumping times

$$0 < \tau_1 < \tau_2 < \cdots$$

Counting process and intensity

Definition (Simple counting process)

A stochastic process $(H_t)_{t\in\mathbb{R}_+}$ is called a simple counting process if

- $H_t \geq 0$ for any $t \geq 0$,
- H is non-decreasing,
- H has jumps of size 1.

Giving a simple counting process is equivalent to giving an infinite increasing sequence of jumping times

$$0 < \tau_1 < \tau_2 < \cdots$$

Definition (Intensity process)

The intensity/rate process $(\lambda_t)_{t\in\mathbb{R}_+}$ is the predictable process defined as

$$\lambda_t \mathrm{d}t = \mathbb{E}[H_{t+\mathrm{d}t} - H_t | \mathcal{F}_{t-}].$$

Hawkes compound process

Definition

Let H a simple counting process and $(X_k)_{k\in\mathbb{N}}$ *i.i.d* positive random variables.
A compound Hawkes process $(L_t)_{t\in\mathbb{R}_+}$ is defined as

$$L_t = \sum_{k=1}^{H_t} X_k,$$

where the intensity process λ of H follows the dynamics

$$\lambda_t = \mu + \int_0^{t^-} \phi(t-s)\mathrm{d}L_s = \mu + \sum_{\tau_k<t} \phi(t-\tau_k)X_k.$$

Where ϕ is a non-negative integrable kernel such that $\|\phi\|_1\mathbb{E}[X] < 1$.

Hawkes process defined as Poisson embedding

Compound Hawkes as thinning from a Poisson measure

Let $N(t, \theta, y)$ be a three component Poisson process of intensity $\mathrm{d}t\mathrm{d}\theta\mathrm{d}\nu(y)$ where $\nu(y)$ is the distribution of X. The following SDE

$$\begin{cases} L_t &= \int_{[0,t]\times\mathbb{R}_+^2} y \mathbb{1}_{\theta\leq\lambda_s} N(\mathrm{d}s, \mathrm{d}\theta, \mathrm{d}\nu(y)), \\ \lambda_t &= \mu + \int_{[0,t)\times\mathbb{R}_+^2} y\phi(t-s) \mathbb{1}_{\theta\leq\lambda_s} N(\mathrm{d}s, \mathrm{d}\theta, \mathrm{d}\nu(y)), \end{cases}$$

has a unique solution such that L is adapted and λ is predictable with respect to the Poisson filtration.

Simulation

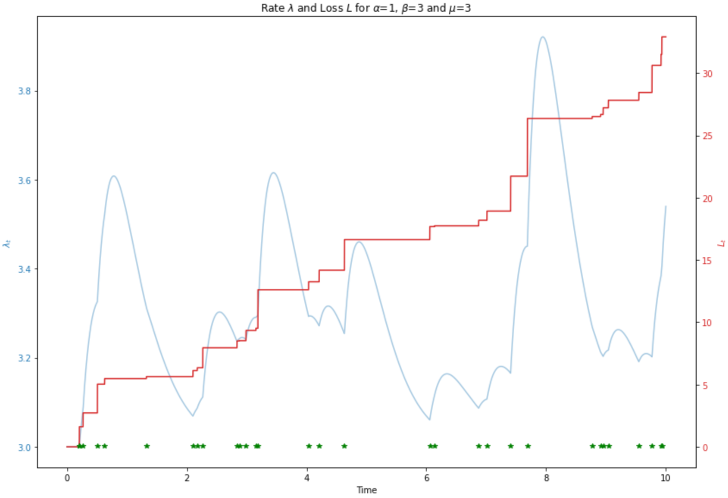

Figure 1: A simulation with $\phi(s) = \alpha s e^{-\beta s}$ and $X \sim Exp(1)$.

Simulation

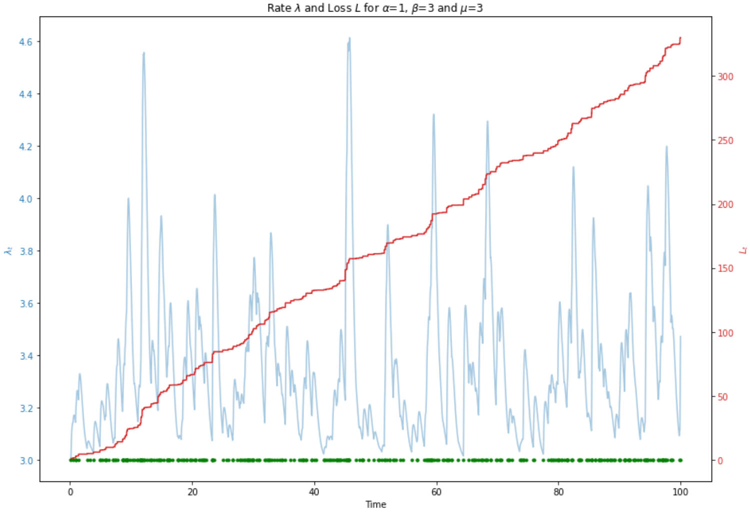

Figure 2: A simulation with $\phi(s) = \alpha s e^{-\beta s}$ and $X \sim Exp(1)$.

Table of Contents

Long time behaviour for L

- Do we have a CLT for L?

Long time behaviour for L

- Do we have a CLT for L?
- Set $m = \mathbb{E}[X]$. We define the martingale $M_T = L_T - m \int_0^T \lambda_t \mathrm{d}t$. If $X \equiv 1$, we have that

$$F_T := \frac{M_T}{\sqrt{T}} \underset{T \to +\infty}{\Longrightarrow} \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 = \frac{\mu}{1 - \|\phi\|_1}$.

Long time behaviour for L

- Do we have a CLT for L?
- Set $m = \mathbb{E}[X]$. We define the martingale $M_T = L_T - m \int_0^T \lambda_t \mathrm{d}t$. If $X \equiv 1$, we have that

$$F_T := \frac{M_T}{\sqrt{T}} \underset{T \to +\infty}{\Longrightarrow} \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 = \frac{\mu}{1 - \|\phi\|_1}$.

- Bacry *et al.* have also proved that for $Y_T = \frac{L_T - m \int_0^T \mathbb{E}[\lambda_t] \mathrm{d}t}{\sqrt{T}}$

$$Y_T \underset{T \to +\infty}{\Longrightarrow} \mathcal{N}(0, \tilde{\sigma}^2),$$

where $\tilde{\sigma}^2 = \frac{\mu}{(1 - \|\phi\|_1)^3}$.

Long time behaviour for L

- Do we have a CLT for L?
- Set $m = \mathbb{E}[X]$. We define the martingale $M_T = L_T - m\int_0^T \lambda_t \mathrm{d}t$. If $X \equiv 1$, we have that

$$F_T := \frac{M_T}{\sqrt{T}} \underset{T\to+\infty}{\Longrightarrow} \mathcal{N}(0,\sigma^2),$$

where $\sigma^2 = \frac{\mu}{1-\|\phi\|_1}$.

- Bacry *et al.* have also proved that for $Y_T = \frac{L_T - m\int_0^T \mathbb{E}[\lambda_t]\mathrm{d}t}{\sqrt{T}}$

$$Y_T \underset{T\to+\infty}{\Longrightarrow} \mathcal{N}(0,\tilde{\sigma}^2),$$

where $\tilde{\sigma}^2 = \frac{\mu}{(1-\|\phi\|_1)^3}$.

- Can we quantify the speed of convergence?

Stein's method

- A measure of the distance between two distributions V and G is the Wasserstein metric

$$d_W(V,G) = \sup_{f \in Lip} \left| \mathbb{E}\left[f(V) - f(G) \right] \right|,$$

with $Lip = \{f \in \mathcal{C}^1(\mathbb{R}), \|f'\|_\infty \leq 1\}$.

Stein's method

- A measure of the distance between two distributions V and G is the Wasserstein metric

$$d_W(V,G) = \sup_{f \in Lip} \left| \mathbb{E}\left[f(V) - f(G)\right] \right|,$$

with $Lip = \{f \in \mathcal{C}^1(\mathbb{R}), \|f'\|_\infty \leq 1\}$.

- If $G \sim \mathcal{N}(0, \gamma^2)$, C.Stein proved that

$$d_W(V,G) \leq \sup_{f \in \mathcal{F}_W} \left| \mathbb{E}\left[\gamma^2 f'(V) - Vf(V)\right] \right|,$$

with $\mathcal{F}_W = \{f \in \mathcal{C}^2(\mathbb{R}), \|f'\|_\infty \leq 1, \|f''\|_\infty \leq 2\}$.

Stein's method

- A measure of the distance between two distributions V and G is the Wasserstein metric

$$d_W(V,G) = \sup_{f \in Lip} \left| \mathbb{E}\left[f(V) - f(G)\right] \right|,$$

with $Lip = \{f \in \mathcal{C}^1(\mathbb{R}), \|f'\|_\infty \leq 1\}$.

- If $G \sim \mathcal{N}(0, \gamma^2)$, C.Stein proved that

$$d_W(V,G) \leq \sup_{f \in \mathcal{F}_W} \left| \mathbb{E}\left[\gamma^2 f'(V) - Vf(V)\right] \right|,$$

with $\mathcal{F}_W = \{f \in \mathcal{C}^2(\mathbb{R}), \|f'\|_\infty \leq 1, \|f''\|_\infty \leq 2\}$.

- How to obtain a bound if we plug in the normalized martingale F_T?

Table of Contents

Definition (Divergence operator)

Let L be a compound Hawkes process of intensity λ. Let $(Z_{t,x})_{(t,x)\in\mathbb{R}_+^2}$ be a predictable process. The divergence of Z is defined as

$$\delta(Z) = \int_{\mathbb{R}_+^3} Z_{t,x}\mathbb{1}_{\theta\leq\lambda_t}\left(N(\mathrm{d}t,\mathrm{d}\theta,\mathrm{d}\nu(x)) - \mathrm{d}t\mathrm{d}\theta\mathrm{d}\nu(x)\right),$$

whenever the expectations of the square of these integrals are finite.

Malliavin calculus for Hawkes processes

Definition (Divergence operator)

Let L be a compound Hawkes process of intensity λ. Let $(Z_{t,x})_{(t,x)\in\mathbb{R}_+^2}$ be a predictable process. The divergence of Z is defined as

$$\delta(Z) = \int_{\mathbb{R}_+^3} Z_{t,x}\mathbb{1}_{\theta\leq\lambda_t}\left(N(\mathrm{d}t,\mathrm{d}\theta,\mathrm{d}\nu(x)) - \mathrm{d}t\mathrm{d}\theta\mathrm{d}\nu(x)\right),$$

whenever the expectations of the square of these integrals are finite.

Example

$$\delta\left(\frac{x\mathbb{1}_{t\leq T}}{\sqrt{T}}\right) = \frac{M_T}{\sqrt{T}} = F_T.$$

Malliavin calculus for Hawkes processes

Definition (Shift operator)

Let $u \leq t$ and $x > 0$. The shift operator consists of adding an artificial jump of size x at time u to the Hawkes process and to its intensity:

$$\begin{cases} L_t \circ \epsilon_{(u,x)} = & L_u + x + \int_{(u,t]\times\mathbb{R}_+^2} y \mathbb{1}_{\theta \leq \lambda_s \circ \epsilon_{(u,x)}} N(\mathrm{d}s, \mathrm{d}\theta, \mathrm{d}\nu(y)), \\ \lambda_t \circ \epsilon_{(u,x)} = & \mu + \int_{(0,u)} \phi(t-s)\mathrm{d}L_s + y\phi(t-u) \\ & + \int_{(u,t)} \phi(t-s)\mathrm{d}(L_t \circ \epsilon_{(u,x)}). \end{cases}$$

We extend the operator naturally to any random variable $V \in \sigma(L_s, s \leq t)$.

Definition (Malliavin derivative)

Let $V \in \sigma(L_s, s \leq t)$. For $u \leq t$, the Malliavin derivative of V is defined as

$$D_{(u,x)}V = V \circ \epsilon_{(u,x)} - V.$$

Shifted loss and intensity

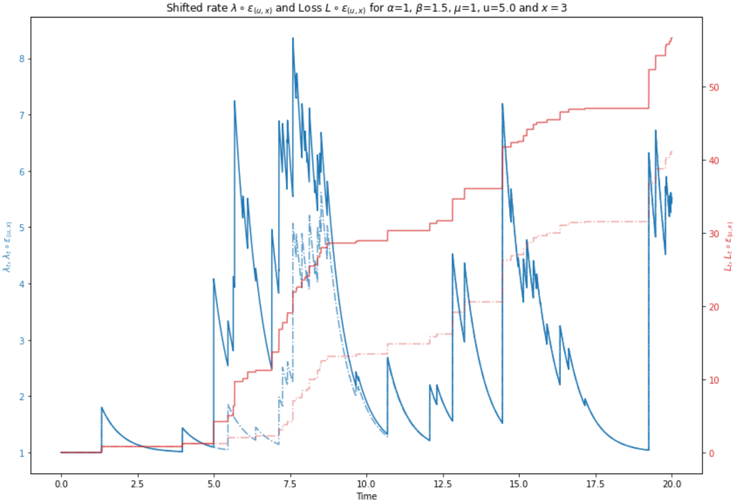

Figure 3: The effect of adding a jump of size $x = 3$ at time $u = 5$. The kernel is $\phi(s) = \alpha e^{-\beta s}$.

Malliavin derivatives of L and λ

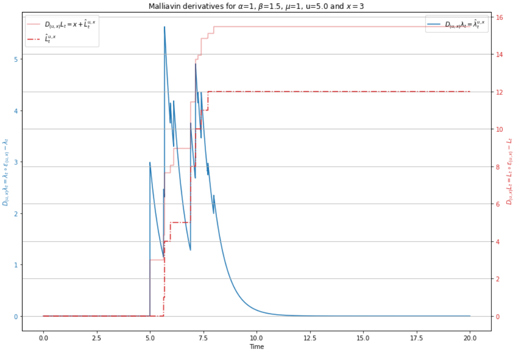

Figure 4: The processes $\hat{L}_t^{u,x}$ and $\hat{\lambda}_t^{u,x}$.

Malliavin derivatives of L and λ

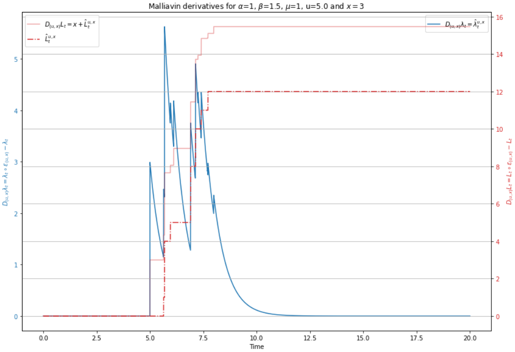

Figure 4: The processes $\hat{L}_t^{u,x}$ and $\hat{\lambda}_t^{u,x}$.

Derivative of the normalized martingale

$$D_{(u,x)}F_T = \frac{1}{\sqrt{T}}\left(x + \hat{M}_T^{u,x}\right)$$

Integration by parts

Duality

Let $(Z_{t,x})_{t\geq 0}$ be a predictable process and $V \in \sigma(L_s, s \geq 0)$. It holds that

$$\mathbb{E}[\delta(Z)V] = \mathbb{E}[\int_{\mathbb{R}_+^2} \lambda_t Z_{t,x} D_{(t,x)} V \mathrm{d}t\mathrm{d}\nu(x)].$$

Verification for $X \equiv 1$, $\phi \equiv 0$ and $Z_t = \mathbb{1}_{t\leq T}$:

$$\begin{aligned}
\mathbb{E}[(N_T - \mu T)f(N_T)] &= \sum_{n=0}^{+\infty} nf(n)e^{-\mu T}\frac{(\mu T)^n}{n!} - \mu T\mathbb{E}[f(N_T)],\\
&= \sum_{n=0}^{+\infty} f(n)e^{-\mu T}\frac{(\mu T)^n}{(n-1)!} - \mu T\mathbb{E}[f(N_T)],\\
&= \mathbb{E}[\mu T\,(f(N_T+1) - f(N_T))].
\end{aligned}$$

Table of Contents

General bound on the Wasserstein metric

- We combine Stein's method with Malliavin calculus: Nourdin-Peccati approach.

General bound on the Wasserstein metric

- We combine Stein's method with Malliavin calculus: Nourdin-Peccati approach.
- Remember that

$$d_W(V, G) \leq \sup_{f \in \mathcal{F}_W} \left| \mathbb{E}\left[\gamma^2 f'(V) - Vf(V)\right]\right|,$$

with $\mathcal{F}_W = \{f \in \mathcal{C}^2(\mathbb{R}), \|f'\|_\infty \leq 1, \|f''\|_\infty \leq 2\}$.

General bound on the Wasserstein metric

- We combine Stein's method with Malliavin calculus: Nourdin-Peccati approach.
- Remember that

$$d_W(V, G) \leq \sup_{f \in \mathcal{F}_W} \left| \mathbb{E}\left[\gamma^2 f'(V) - V f(V)\right]\right|,$$

  with $\mathcal{F}_W = \{f \in \mathcal{C}^2(\mathbb{R}), \|f'\|_\infty \leq 1, \|f''\|_\infty \leq 2\}$.
- If $V = \delta(Z)$, then $\mathbb{E}[Vf(V)]$ becomes

$$\mathbb{E}[\int_{\mathbb{R}_+^2} \lambda_t Z_{t,x} D_{(t,x)} f(V) \mathrm{d}t \mathrm{d}\nu(x)].$$

General bound on the Wasserstein metric

- We combine Stein's method with Malliavin calculus: Nourdin-Peccati approach.
- Remember that

$$d_W(V,G) \leq \sup_{f \in \mathcal{F}_W} \left| \mathbb{E}\left[\gamma^2 f'(V) - Vf(V)\right]\right|,$$

with $\mathcal{F}_W = \{f \in \mathcal{C}^2(\mathbb{R}), \|f'\|_\infty \leq 1, \|f''\|_\infty \leq 2\}$.

- If $V = \delta(Z)$, then $\mathbb{E}[Vf(V)]$ becomes

$$\mathbb{E}[\int_{\mathbb{R}_+^2} \lambda_t Z_{t,x} D_{(t,x)} f(V) \mathrm{d}t \mathrm{d}\nu(x)].$$

- Using a Taylor expansion it is possible to write

$$D_{(t,x)} f(V) = f'(V) D_{(t,x)} V + \frac{1}{2} f''(\bar{V}) |D_{(t,x)} V|^2.$$

General bound on the Wasserstein metric

- Hence

$$\mathbb{E}\left[\gamma^2 f'(V) - Vf(V)\right] = \mathbb{E}\left[\gamma^2 f'(V)\right] - \mathbb{E}\left[f'(V)\int_{\mathbb{R}_+^2} \lambda_t Z_t D_{(t,x)} V \mathrm{d}t\mathrm{d}\nu(x)\right] - \frac{1}{2}\mathbb{E}\left[f''(\bar{V})\int_{\mathbb{R}_+^2} \lambda_t |Z_t||D_{(t,x)} V|^2 \mathrm{d}t\mathrm{d}\nu(x)\right].$$

General bound

Let $G \sim \mathcal{N}(0,\gamma^2)$. For $V = \delta(Z)$ we have the bound

$$d_W(V,G) \leq \mathbb{E}[|\gamma^2 - \int_{\mathbb{R}_+^2} \lambda_t Z_{t,x} D_{(t,x)} V \mathrm{d}t\mathrm{d}\nu(x)|] + \mathbb{E}[\int_{\mathbb{R}_+^2} \lambda_t |Z_{t,x}||D_{(t,x)} V|^2 \mathrm{d}t\mathrm{d}\nu(x)].$$

Application to the normalized martingale

- Recall that $F_T = \frac{M_T}{\sqrt{T}} = \delta(\frac{\mathbb{1}_{t\leq T}}{\sqrt{T}})$.

Theorem

If ν has a moment of order 2 and $\int_0^{+\infty} s\phi(s)ds < +\infty$ then we have the following

$$d_W(F_T, G) \leq \frac{1}{T}\mathbb{E}[|\int_{\mathbb{R}_+}\int_0^T \lambda_t \hat{M}_T^{t,x} \mathrm{d}t\mathrm{d}\nu(x)|] + O(\frac{1}{\sqrt{T}}),$$

where $G \sim \mathcal{N}(0, \sigma^2)$.

Application to the normalized martingale

- Recall that $F_T = \frac{M_T}{\sqrt{T}} = \delta(\frac{\mathbb{1}_{t\leq T}}{\sqrt{T}})$.
- The Gaussian limit is centered and its variance is $\sigma^2 = \frac{\int x^2 \mathrm{d}\nu(x)\mu}{1-m\|\phi\|_1}$.

Theorem

If ν has a moment of order 2 and $\int_0^{+\infty} s\phi(s)ds < +\infty$ then we have the following

$$d_W(F_T, G) \leq \frac{1}{T}\mathbb{E}[| \int_{\mathbb{R}_+} \int_0^T \lambda_t \hat{M}_T^{t,x} \mathrm{d}t\mathrm{d}\nu(x)|] + O(\frac{1}{\sqrt{T}}),$$

where $G \sim \mathcal{N}(0, \sigma^2)$.

An explicit bound

- For a spacial choice of the kernel $\phi(s) = \alpha e^{-\beta s}$ $(m\alpha < \beta)$ or $\phi(s) = \alpha s e^{-\beta s}$ $(m\alpha < \beta^2)$, (λ, L) is (up to some extra manipulations) a Markov process. The same can be said about $(\hat{\lambda}_T^{u,x}, \hat{L}_T^{u,x})$.
- In these cases we have more explicit results on $\hat{M}_T^{t,x}$ which makes it possible to have a better bound.

Theorem (Explicit bound)

If ν has a third moment and ϕ is either an exponential or an Erlang function then we have that

$$d_W(F_T, G) = O(\frac{1}{\sqrt{T}}).$$

Approximation of L by a deterministic entity

- It is interesting to approximate L_t with a deterministic entity.

Approximation of L by a deterministic entity

- It is interesting to approximate L_t with a deterministic entity.
- Recall that $Y_T = \frac{L_T - m\int_0^T \mathbb{E}[\lambda_t]\mathrm{d}t}{\sqrt{T}}$. If the kernel is exponential or Erlang, $\int_0^T \mathbb{E}[\lambda_t]\mathrm{d}t$ is known explicitely.

Approximation of L by a deterministic entity

- It is interesting to approximate L_t with a deterministic entity.
- Recall that $Y_T = \frac{L_T - m\int_0^T \mathbb{E}[\lambda_t]\mathrm{d}t}{\sqrt{T}}$. If the kernel is exponential or Erlang, $\int_0^T \mathbb{E}[\lambda_t]\mathrm{d}t$ is known explicitely.
- Moreover we have that

$$Y_T = \frac{F_T}{1 - m\|\phi\|_1} + R_T$$

where $R_T \propto \frac{\lambda_T - \mathbb{E}[\lambda_T]}{\sqrt{T}}$ is a "small" remainder.

Approximation of L by a deterministic entity

- It is interesting to approximate L_t with a deterministic entity.
- Recall that $Y_T = \frac{L_T - m\int_0^T \mathbb{E}[\lambda_t]\mathrm{d}t}{\sqrt{T}}$. If the kernel is exponential or Erlang, $\int_0^T \mathbb{E}[\lambda_t]\mathrm{d}t$ is known explicitely.
- Moreover we have that

$$Y_T = \frac{F_T}{1 - m\|\phi\|_1} + R_T$$

where $R_T \propto \frac{\lambda_T - \mathbb{E}[\lambda_T]}{\sqrt{T}}$ is a "small" remainder.

- In this case as well we can prove that

$$d_W(Y_T, \tilde{G}) = O(\frac{1}{\sqrt{T}}),$$

where $\tilde{G} \sim \mathcal{N}(0, \tilde{\sigma}^2)$ with $\tilde{\sigma}^2 = \frac{\int x^2 \mathrm{d}\nu(x)\mu}{(1-m\|\phi\|_1)^3}$.

Thank you for your attention.