Hawkes Processes with Stochastic Excitations

Young Lee, Kar Wai Lim, Cheng Soon Ong

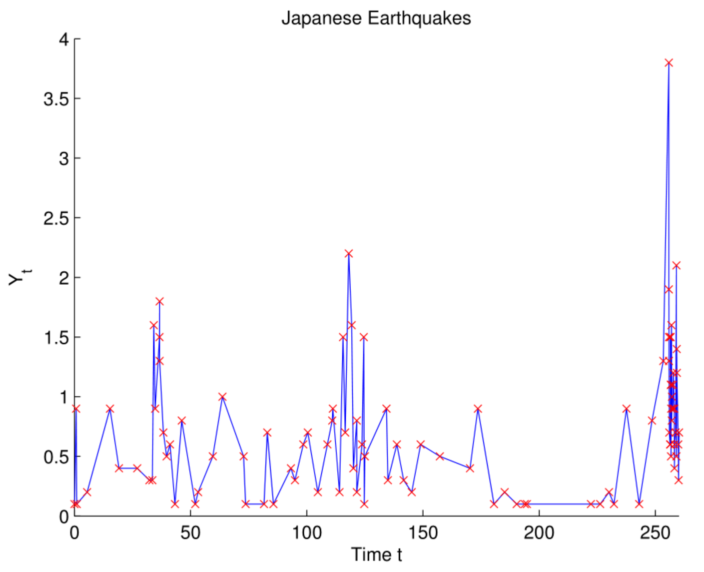
NICTA, London School of Economics, Australian National University

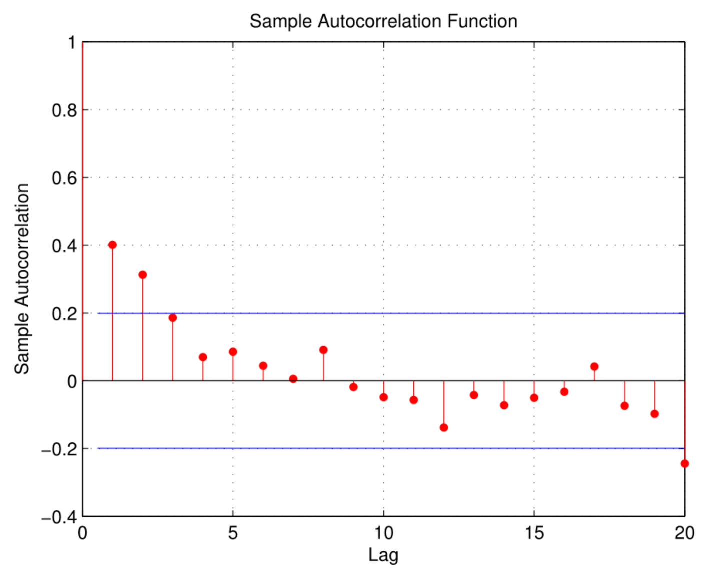
Contributions/Highlights

1. A fully Bayesian framework that utilises Stochastic Differential Equations (SDEs) to model the excitatory relationships of a Hawkes process.
2. The SDEs allow the levels of excitation (Y) to be correlated, a feature that cannot be tackled by existing models using constant or i.i.d. levels of excitation.
3. A novel simulation algorithm for the Stochastic Hawkes, drawing the levels of excitation as needed, following discretization with unequal periods.
4. A hybrid MCMC algorithm of Metropolis-Hastings (MH) and Gibbs sampler, made possible using the branching representation of Hawkes processes.
5. Synthetic experiments show that Stochastic Hawkes is more flexible and can model i.i.d. excitations, while the Hawkes process with i.i.d. excitations fail to fit the stochastic excitations.
6. Correlation in Japanese earthquakes' magnitude (ETAS) are better modelled by Stochastic Hawkes compared to classical Hawkes process.

Why use SDEs to model the levels of excitations?

– More flexible and can capture correlation in the level of excitations.
– Example: Japanese Earthquakes Data from ETAS (year 1951 – 1952, see Di Giacomo et al, 2015)





– This shows there is correlation in the dataset.

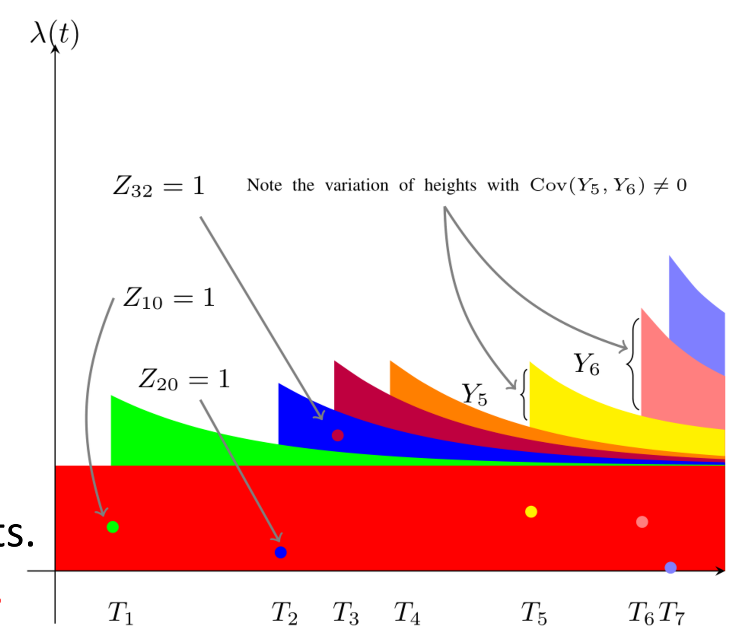
Stochastic Hawkes

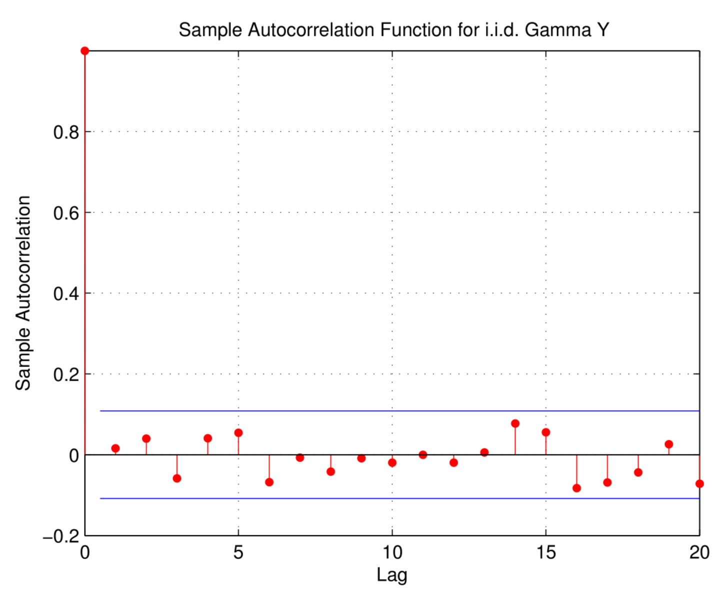
- **Intensity Function**

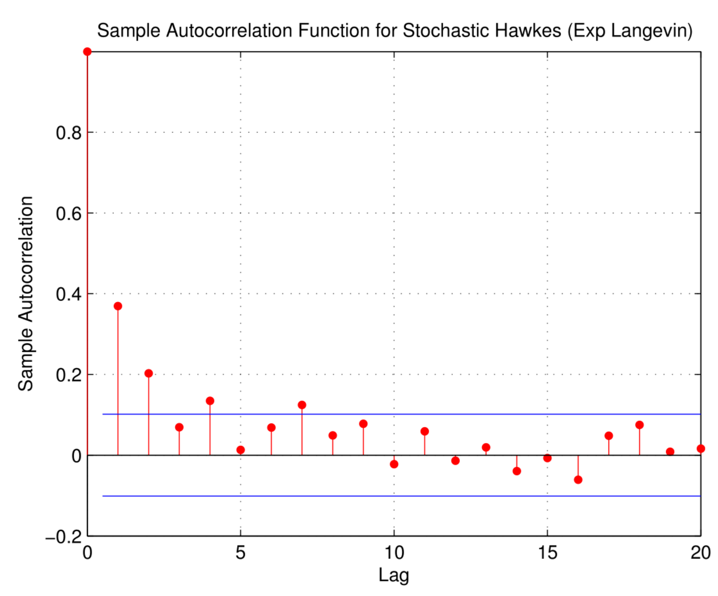
– Intensity function is of the following form:

$$\lambda(t) = \hat{\lambda}_0(t) + \sum_{i:t>T_i} Y(T_i)\,\nu(t - T_i)$$

– There are three kinds of Y.
  – Classical. Y = constant.
  – Random excitations. Y = i.i.d. elements.
  – Stochastic excitations. Y follows SDEs.







– i.i.d. elements do not exhibit correlation (left, as expected), but Stochastic Y induce correlation which are observed on real data (right).

- **Stochastic Excitations**

– We present two SDEs for modelling Y.
– Geometric Brownian Motion (GBM):

$$Y_\cdot = \int_0^\cdot \left(\mu + \frac{1}{2}\sigma^2\right) Y_t\,dt + \int_0^\cdot \sigma Y_t\,dB_t$$

where B_t is a Brownian motion. μ and σ are parameters.

– Exponential Langevin:

$$Y_\cdot = \exp\left(\int_0^\cdot k(\mu - Y_t)\,dt + \int_0^\cdot \sigma\,dB_t\right)$$

where B_t is a Brownian motion. μ and σ are parameters.

Simulation of Stochastic Hawkes

– Exact simulation of Stochastic Hawkes extending Dassios and Zhao (2013).
– By discretizing the SDEs using unequal periods from the event times, the levels of excitation Y are simulated as required by Stochastic Hawkes.
– The algorithms are as follow:

Algorithm 1 Simulation of Stochastic Hawkes

1. We firstly set $T_0 = 0$, $\lambda_0^{(1)} = \lambda_0 - a$, and given Y_0.
2. For $i = 1, 2, \dots$ and while $T_i < T$:
   (a) Draw $S_i^{(0)} = -\frac{1}{a}\log U(0,1)$.
   (b) Draw $u \sim U(0,1)$. Set $S_i^{(1)} = -\frac{1}{\delta}\log\left(1 - \delta/\lambda_{T_{i-1}}^{(1)}\log u\right)$. Note we set $S_k^{(1)} := \infty$ when the log term is undefined.
   (c) Set $T_i = T_{i-1} + \min\left(S_i^{(0)}, S_i^{(1)}\right)$.
   (d) Sample Y_{T_i} (refer to Algorithm 2)
   (e) Update $\lambda_{T_i}^{(1)} = \lambda_{T_{i-1}}^{(1)} e^{-\delta(T_i - T_{i-1})} + Y_{T_i}$.

Algorithm 2 Simulation of Stochastic Y_i

1. Given Y_{i-1} and $\{T_{i-1}, T_i\}$
2. If $Y \sim$ Geometric Brownian Motion, then
   (a) Sample Y_i through

$$u \sim N(0, \sigma^2(T_i - T_{i-1})),$$
$$Y_i = Y_{i-1}\exp\left(\mu(T_i - T_{i-1}) + u\right)$$

   If $Y \sim$ Exponential Langevin, then
   (a) Sample Y_i using

$$u \sim N\left(0, \frac{\sigma^2}{2k}\left(1 - e^{-2k(T_i - T_{i-1})}\right)\right),$$
$$Y_i = \exp\left(\log Y_{i-1} e^{-k(T_i - T_{i-1})} + \mu(1 - e^{-k(T_i - T_{i-1})}) + u\right)$$

There are two Representations of Point Processes

- **Intensity Representation**

– A point process can be defined by its intensity function (see `Stochastic Hawkes').

- **Branching Representation**

– Alternatively, can use the branching representation for point process, the benefit of this is that the likelihood function of the event times can be simplified.
– All events are classified into `immigrants' or `offsprings':
  – Immigrant means the event time is generated from the base intensity function.
  – Offspring means the event time is generated from the intensity excitation of other event times.
– These classifications are captured by indicators Z_{ij} . (see details in the paper)
– Likelihood of event times T_i (left: without branching; right: with branching):

$$e^{-\Lambda_T}\prod_{i=1}^{N_T} a + (\lambda_0 - a)e^{-\delta t} + \sum_{j\in\mathcal{J}_i}\left[Y_j e^{-\delta(T_i - T_j)}\right] \Longrightarrow e^{-\Lambda_T}\prod_{i=1}^{N_T}\left(a + (\lambda_0 - a)e^{-\delta t}\right)^{Z_{i0}}\prod_{j\in\mathcal{J}_i}\left[Y_j e^{-\delta(T_i - T_j)}\right]^{Z_{ij}}$$

- **Hybrid of Metropolis-Hastings and Gibbs Sampling**

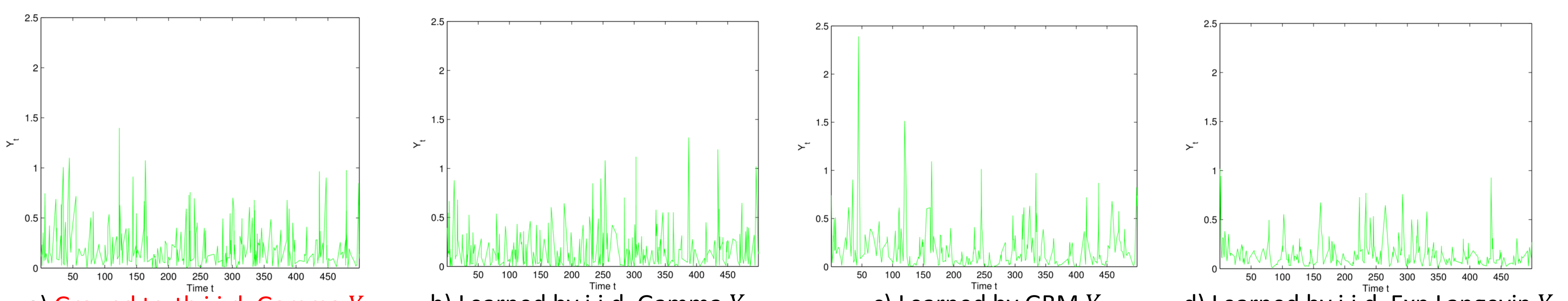
– The employment of branching representation enables the use of Gibbs sampling to learn Z, μ, and σ.
– Other parameters a, λ_0, and Y are learned with the vanilla MH algorithm.

Experiments and Results

- **Synthetic Validation**

– Using the exact simulation algorithm, the event times and the levels of excitations are generated assuming Y follows i.i.d. Gamma, GBM, or Exponential Langevin.
– Performing experiments to recalibrate the parameters and subsequently sample the posterior Y gives the following interesting results:

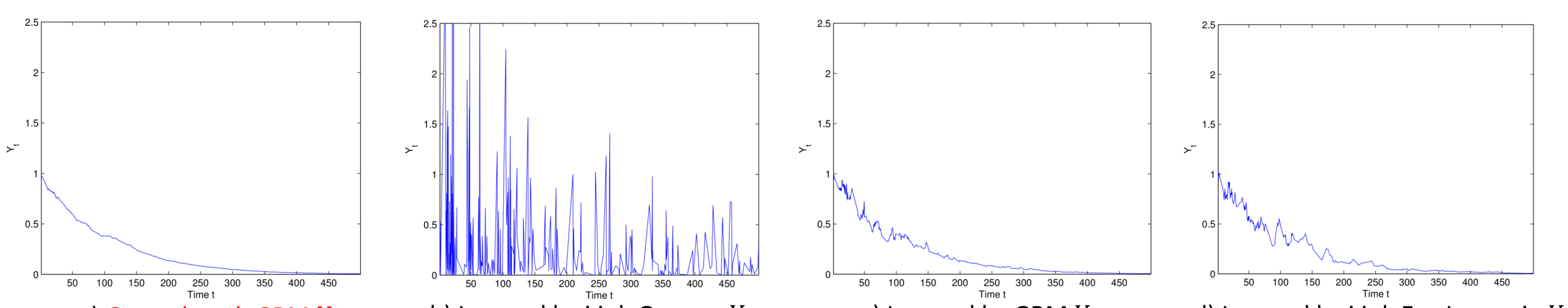
- **When ground truth Y is i.i.d. Gamma**



a) Ground truth i.i.d. Gamma Y b) Learned by i.i.d. Gamma Y c) Learned by GBM Y d) Learned by i.i.d. Exp Langevin Y

– Model (b) has the same model as (a), thus exhibiting same distribution (good result).
– Stochastic Hawkes (c and d) can also learn/imitate i.i.d. Y despite diff model. (good result)

- **When ground truth Y follows Geometric Brownian Motion**



a) Ground truth GBM Y b) Learned by i.i.d. Gamma Y c) Learned by GBM Y d) Learned by i.i.d. Exp Langevin Y

– Hawkes with i.i.d. Y (b) fails to learn excitations that follow SDE which exhibit correlation (a).
– While both Stochastic Hawkes (c and d) can learn back the Y.