Georgia Institute of Technology

# Switched Hawkes Processes

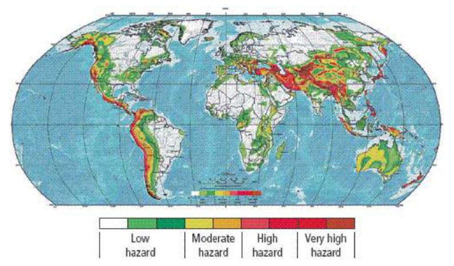
Namrata Nadagouda and Mark A. Davenport

4877

## Hawkes Processes

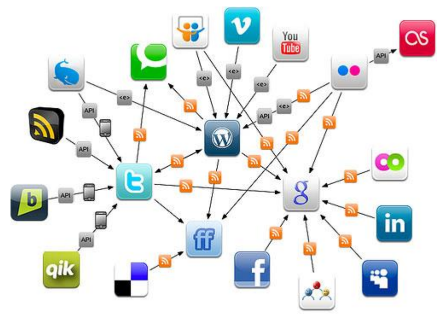
A class of auto-regressive point processes used to model data in which events tend to cluster and influence the likelihood of future events.

**Applications**



Earthquakes

Finance



Social media

**Conditional Intensity Function**

$\lambda(t|\mathcal{H}_t) = \mu + a\sum_{\tau_k \in \mathcal{H}_t} \gamma(t-\tau_k)$

$\mu$ - base intensity
$a$ - excitation co-efficient
$\mathcal{H}_t = \{\tau_k : \tau_k < t\}$ - process history
$\gamma(t)$ - excitation kernel

**Multivariate process**

$\lambda_i(t|\mathcal{H}_t) = \mu_i + \sum_{j=1}^{N} a_{i,j} \sum_{\tau_k \in \mathcal{H}_{t,j}} \gamma(t-\tau_k)$

$\mathcal{H}_{t,j} = \{\tau_k : \tau_k < t, \theta_k = j\}$
$\boldsymbol{\mu} = [\mu_i]$ - base intensity
$\boldsymbol{A} = [a_{i,j}]$ - excitation co-efficients

Here, $a_{i,j}$ characterize the inter-process influences

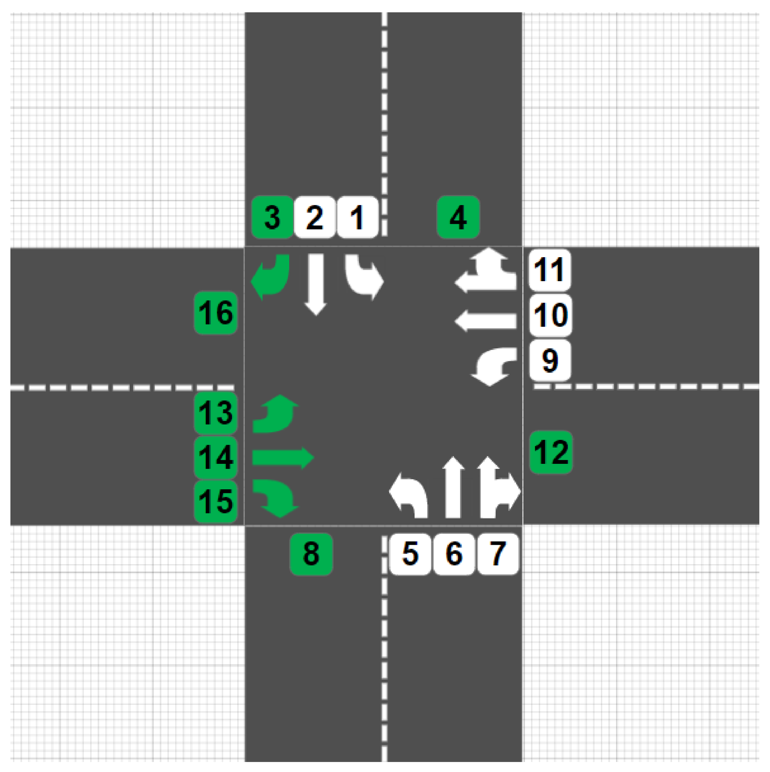
## Motivation

Majority of the works on Hawkes Processes assume that the parameters determining the intensity function remain constant.

But, in most of the practical scenarios, the underlying dynamics which influence the nature of the process can change over time.

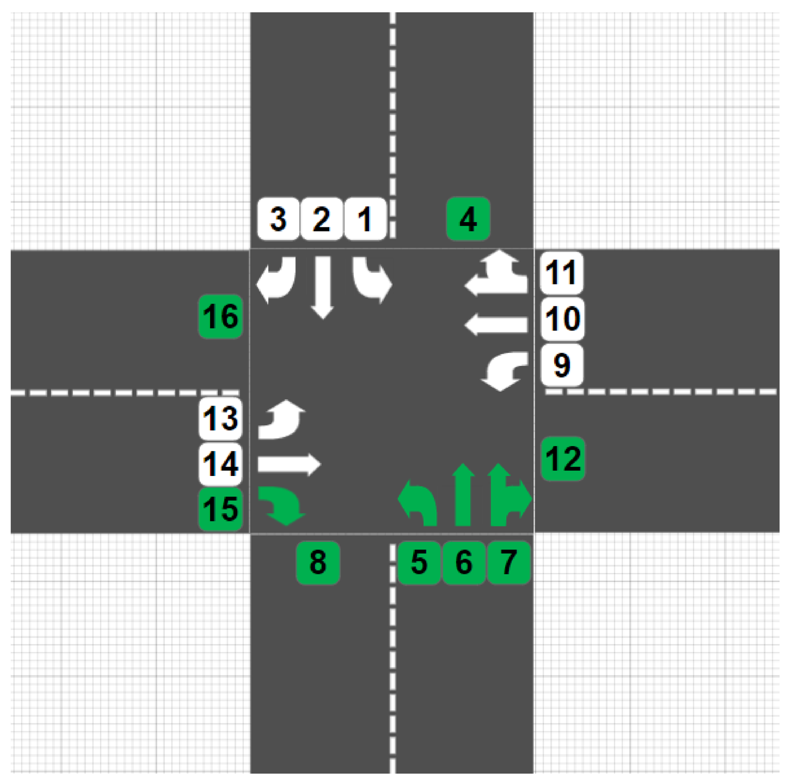
**An example - traffic flow analysis**

Two representative states of traffic flow at an intersection



(Green arrows indicate active directions of traffic flow)

State 1



State 2

- Squares numbered 1 - 16 represent sensors: they record an event whenever a vehicle passes over them
- State 1 - events recorded by sensors 13, 3, ... and correlations between events of sensors 13 and 4, 3 and 16, ... but that does not hold good for State 2
- The nature of events observed depends on the underlying dynamics - in this case being the traffic lights
- A static model would fail to explain such observations

## Switched Model

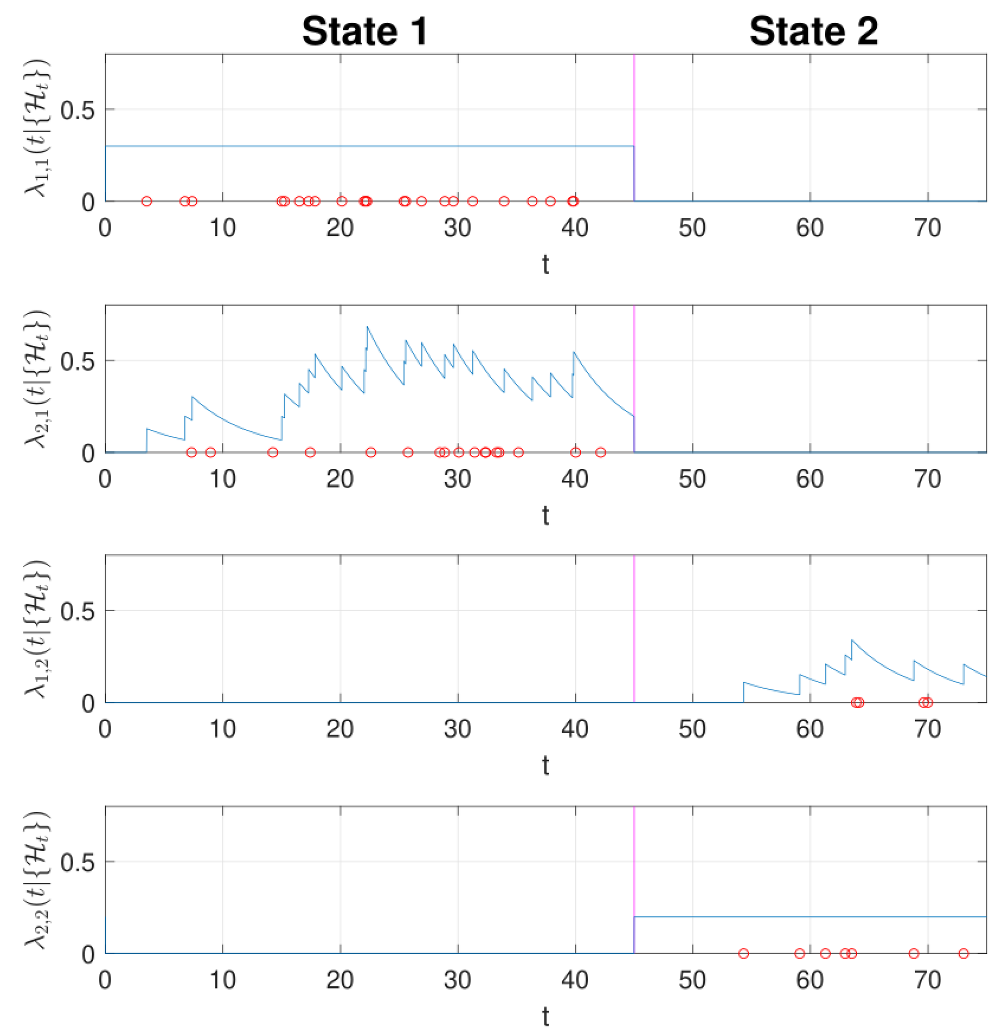
We propose a switched model - the overall process switches among a known finite number of states and every state is characterized by an individual process

**Conditional Intensity Function for an $N$-dimensional process of order $M$**

$$\lambda_{i,s}(t|\mathcal{H}_t) = \mu_{i,s} + \sum_{j=1}^{N}\sum_{s'=1}^{M} (a_{i,s})_{j,s'} \sum_{\tau_k \in \mathcal{H}_{t,j,s'}} \gamma(t-\tau_k)$$

where $\mathcal{H}_{t,j,s'} = \{\tau_k : \tau_k < t, \theta_k = j, s_k = s'\}$

$\boldsymbol{\mu} = [\mu_{i,s}]$ - base intensity, $\boldsymbol{A} = [(a_{i,s})_{j,s'}]$ - excitation co-efficients




CIF for the process characetrizing the data simulated according to the model described on the right. The vertical line at t = 45s represents a state transition. The red circles indicate the event occurrence times.

## Maximum Likelihood Parameter Estimation

Estimation of $\boldsymbol{\mu}, \boldsymbol{A}$

Suppose that we observe the sequence of events $\{(\tau_1, \theta_1, s_1), \ldots, (\tau_K, \theta_K, s_K)\}$ on $[0, T)$. The likelihood of these events is given by

$$L(\boldsymbol{\mu}, \boldsymbol{A}|\mathcal{H}_T) = \left(\prod_{k=1}^{K} \lambda_{\theta_k, s_k}(\tau_k)\right) \exp\left(-\int_0^T \lambda(t)dt\right)$$

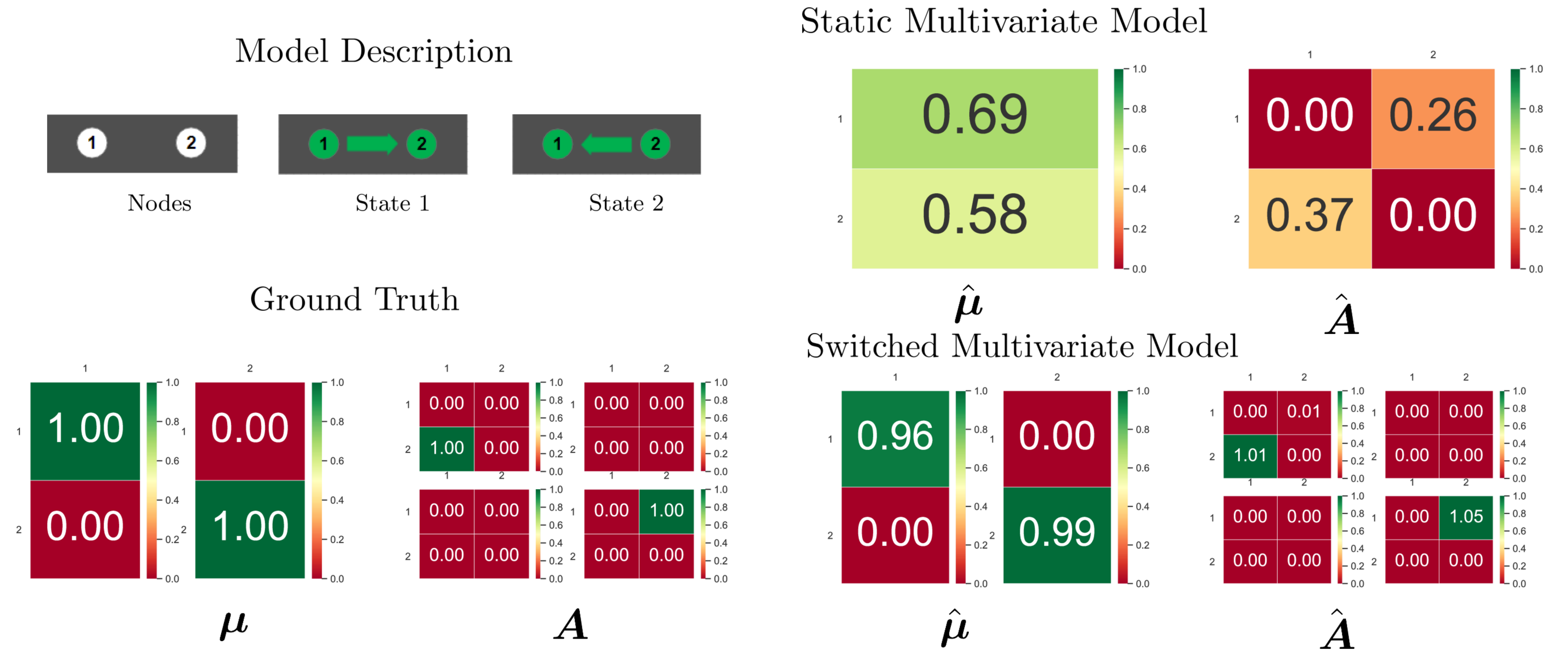
$$\text{where} \quad \lambda(t) = \sum_{i=1}^{N}\sum_{s=1}^{M} \lambda_{i,s}(t)$$

Separable form

$$\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{A}|\mathcal{H}_T) = \sum_{i=1}^{N}\sum_{s=1}^{M} \mathcal{L}_{i,s}(\boldsymbol{\mu}, \boldsymbol{A}|\mathcal{H}_T)$$

$$\text{where} \quad \mathcal{L}_{i,s}(\boldsymbol{\mu}, \boldsymbol{A}|\mathcal{H}_T) = \int_0^T \lambda_{i,s}(t)dt - \sum_{\tau_k \in \mathcal{H}_{T,i,s}} \log \lambda_{i,s}(\tau_k)$$

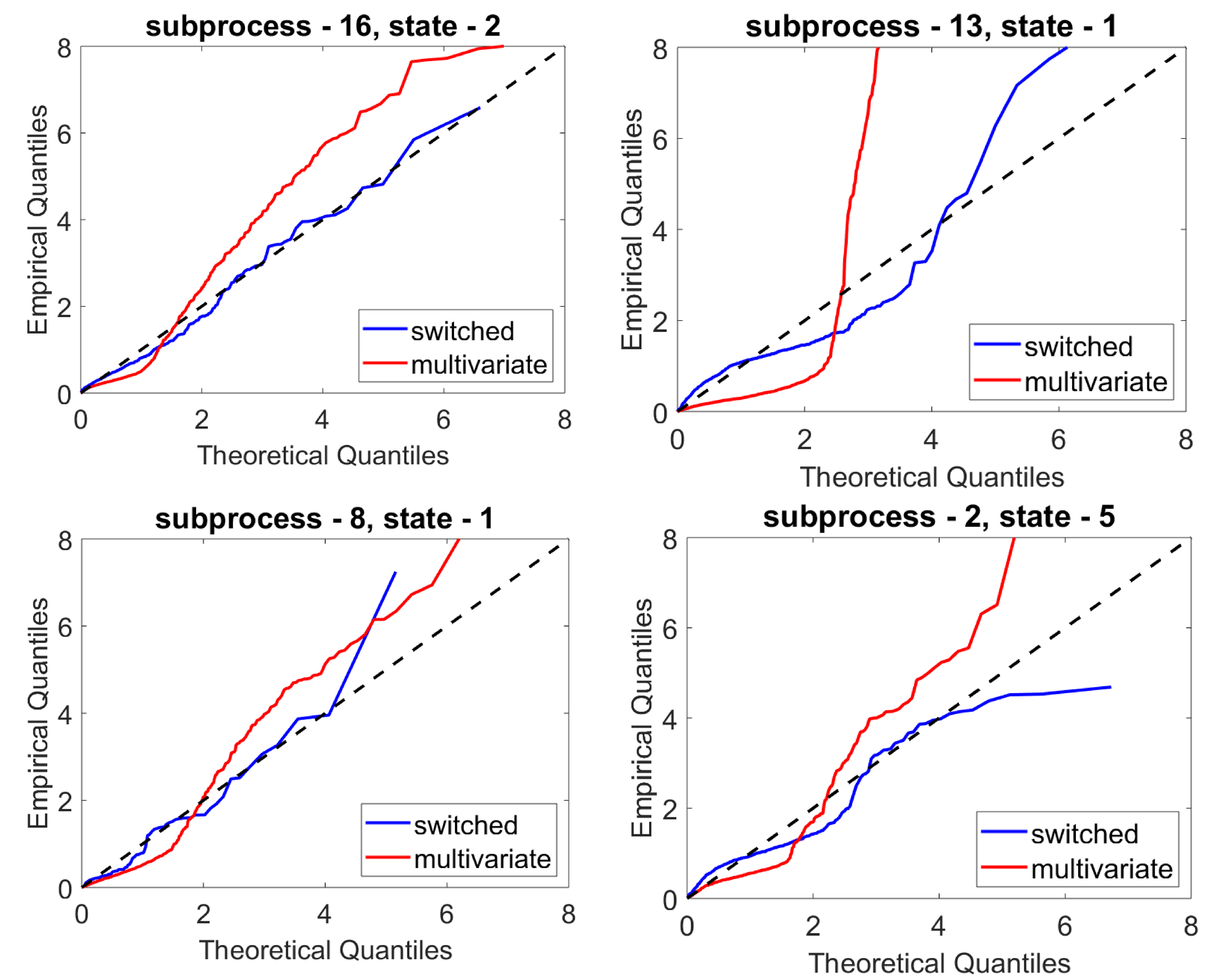
## Simulated Data – Estimated Parameters




- Parameters estimated for both switched and multivariate models on the data simulated according to a switched model as described above
- The multivariate model learns parameters that capture the average performance of the system in both the states.
- The switched model more accurately picks up the base and excitation intensities corresponding to the underlying state of the system.

## Traffic Data – QQ plots

- We consider traffic data at an intersection and investigate which of the two models better explains the data.
- We estimate a 16-dimensional switched Hawkes model of order 6 and a 16-dimensional static Hawkes model.




Representative QQ plots for the estimated models

- QQ-plots are used to determine how well a model fits the given data - the closer the plot is to the black dashed reference line, the better the fit is.
- The curve for the switched model is closer to the reference than that of the multivariate model, thus confirming that the former is a better fit to the data.