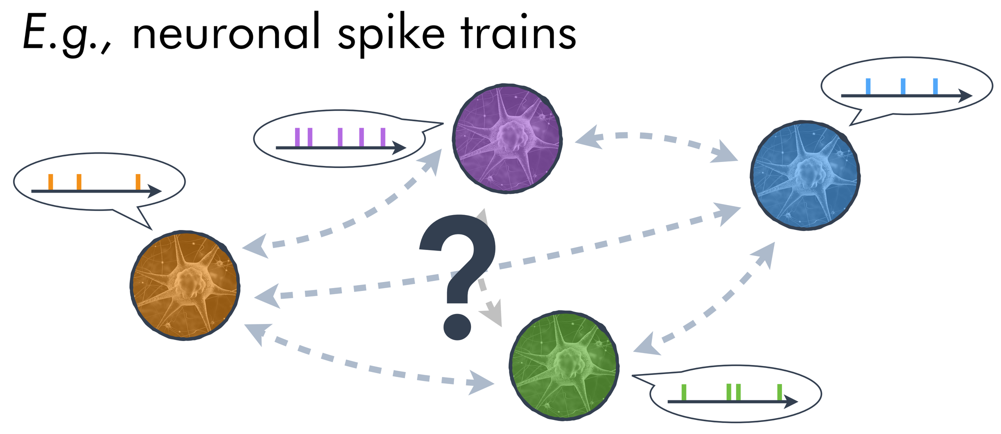
Learning Hawkes Processes Under Synchronization Noise

William Trouleau[1], Jalal Etesami[2], Matthias Grossglauser[1], Negar Kiyavash[3], Patrick Thiran[1]

[1] Information and Network Dynamics group, School of Computer and Communication Sciences, EPFL
[2] Bosch Center for Artificial Intelligence
[3] Department of Industrial and Systems Engineering (ISyE), Georgia Institute of Technology

Motivation

- We want to learn the **causal relationships** in a network of discrete events (time series).
  E.g., neuronal spike trains



- **Multivariate Hawkes Processes** (MHPs) are widely used to model mutually exciting patterns in discrete events.
- But experimental data are inherently noisy…

Question: How to identify the causal structure of the network when observations are noisy?

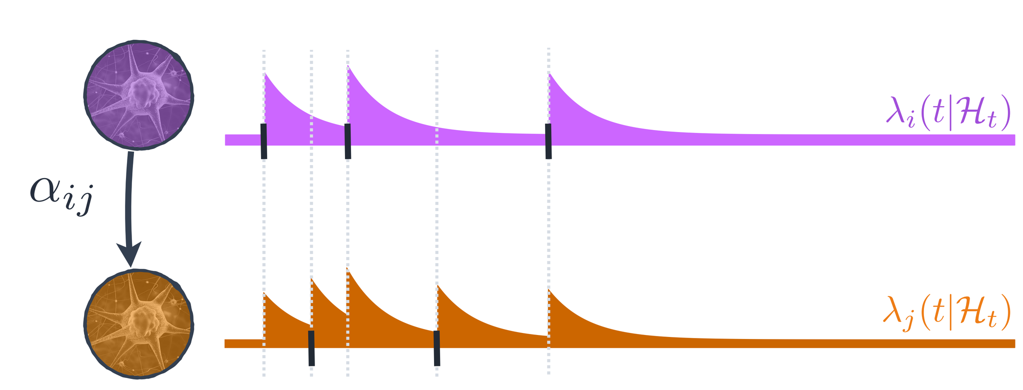
Model

- A Multivariate Hawkes Process is a point process with intensity

$$\lambda_i(t|\mathcal{H}_t) = \mu_i + \sum_{j=1}^{d} \sum_{\tau \in \mathcal{H}_t^j} \kappa_{ij}(t-\tau)$$

Exogenous intensity: constant, independent of the past

Endogenous intensity: due to excitation from past events, with excitation kernel
$\kappa_{ij}(t) = \alpha_{ij} e^{-\beta t} \mathbb{1}\{t > 0\}$



Contributions

1 We introduce the so-called **synchronization noise** model, where the stream of events generated by each dimension is subject to a random and unknown time shift.

2 We characterize the robustness of the classic maximum likelihood estimator of MHPs to synchronization noise.

3 We introduce, DESYNC-MHP, a new approach for learning the causal structure of a MHP in the presence of noise.

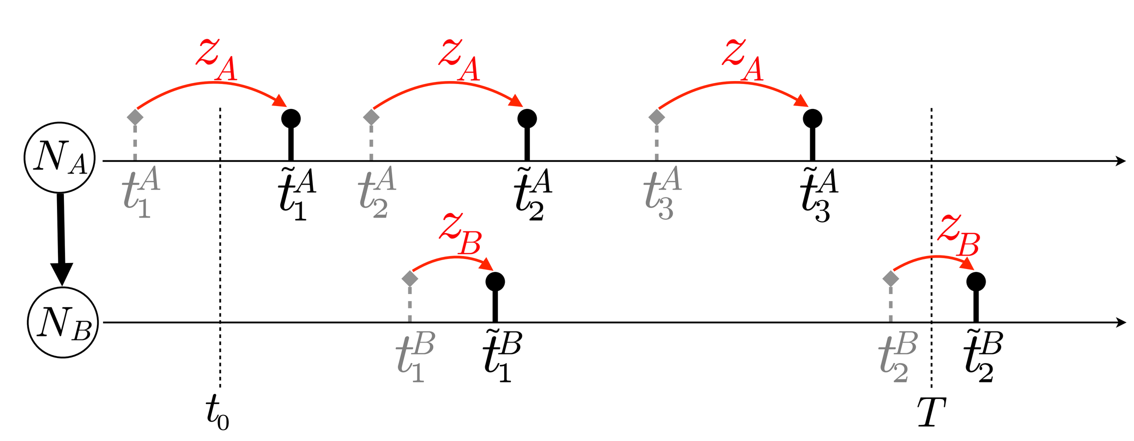
Noisy Observation Framework

- With **synchronization noise**, we assume that all the events within a dimension are shifted equally by an unknown offset:

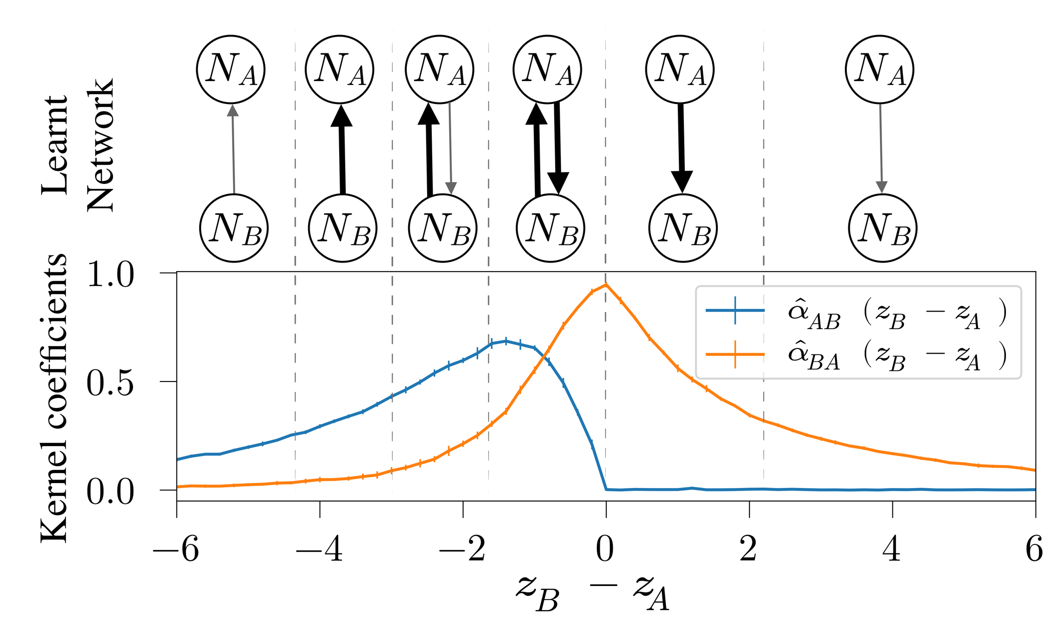
$$\tilde{t}_k^i - t_k^i = z_i \text{ for all events in dimension } i.$$

k^{th} noiseless event in dimension i
corresponding observed event

- Synchronization noise do not change the relative order of the arrivals within a dimension, but it **affects the relative order of the arrivals between different dimensions.**



Example of noisy observations from a process N_A➞N_B in d=2 dimensions.



Maximum likelihood estimate on the example N_A➞N_B.

Inference Under Synchronization Noise

- We define a new multivariate point process called **DESYNC-MHP**, that is a MHP with synchronization noise and we derive its maximum likelihood estimator.

Challenge

Noise introduces discontinuities in the log-likelihood function!

$$\kappa_{21}(\tilde{t}_2 - z_2 - \tilde{t}_1 + z_1)$$
$$= \alpha_{21}\, e^{-\beta(\tilde{t}_2 - z_2 - \tilde{t}_1 + z_1)}\, \mathbb{1}\{\tilde{t}_2 - z_2 - \tilde{t}_1 + z_1 > 0\}$$

$$\lim_{\tau \to 0^+} \kappa_{21}(\tau) = \alpha_{21} \neq 0 = \lim_{\tau \to 0^-} \kappa_{21}(\tau)$$

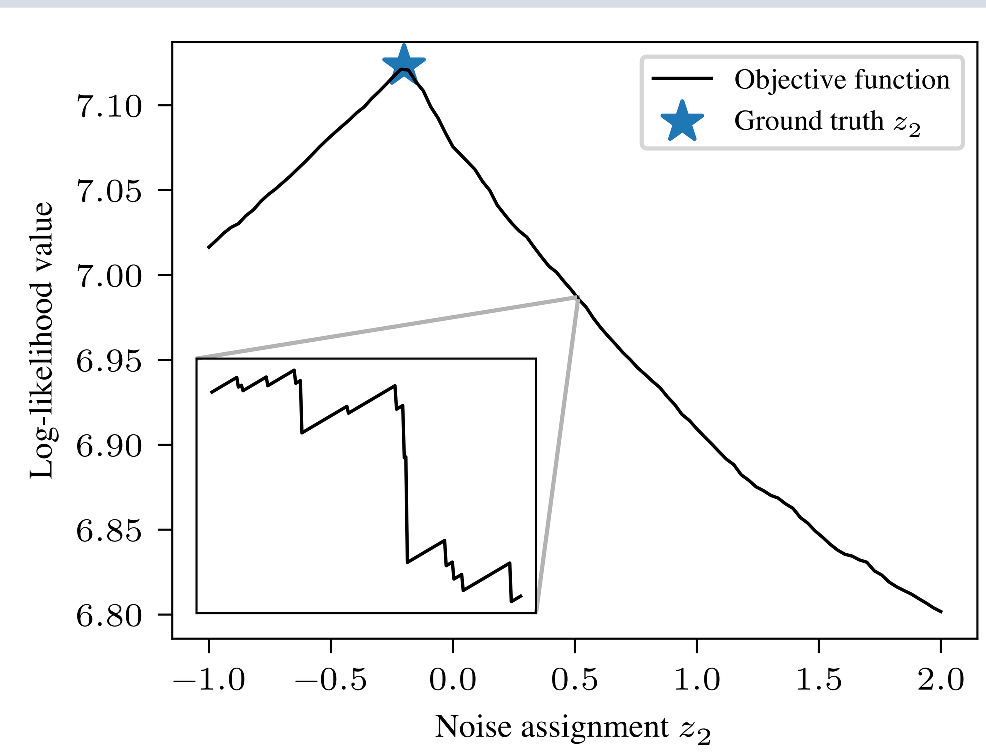


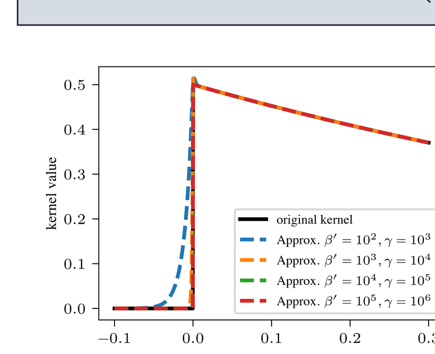
Illustration of the discontinuities of the log-likelihood function for a two-dimensional DESYNC-MHP projected on dimension z_2.

Solution

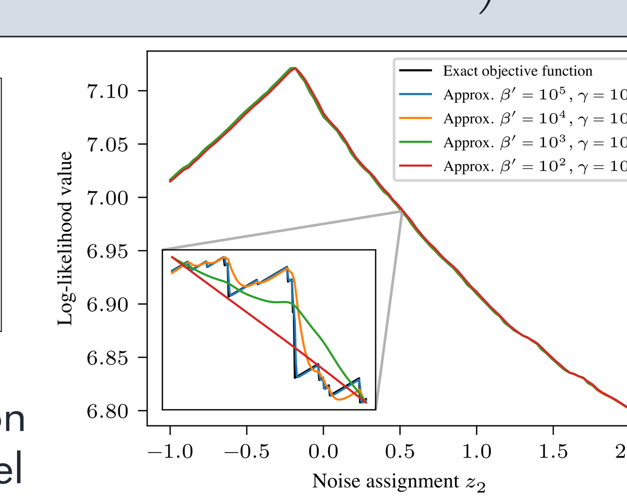
Discontinuities come from the non-smoothness of the excitation kernel in t=0.

➜ Use a smooth approx. of the excitation kernel:

$$\tilde{\kappa}_{ij}(t) \triangleq \alpha_{ij} \left(\sigma(\gamma t) e^{-\beta t} + (1 - \sigma(\gamma t)) e^{\beta' t} \right)$$

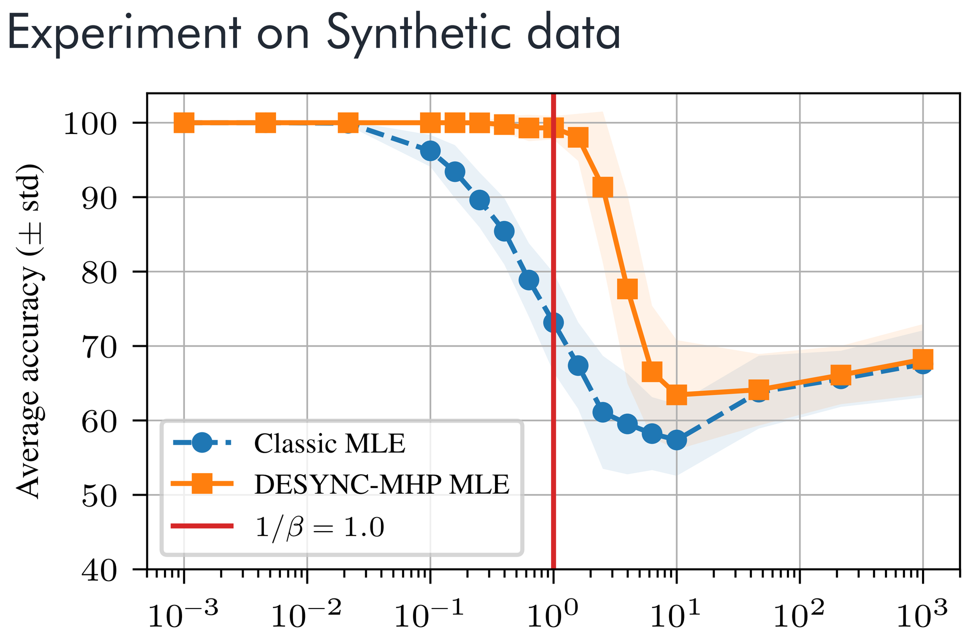


Smooth approximation of the excitation kernel



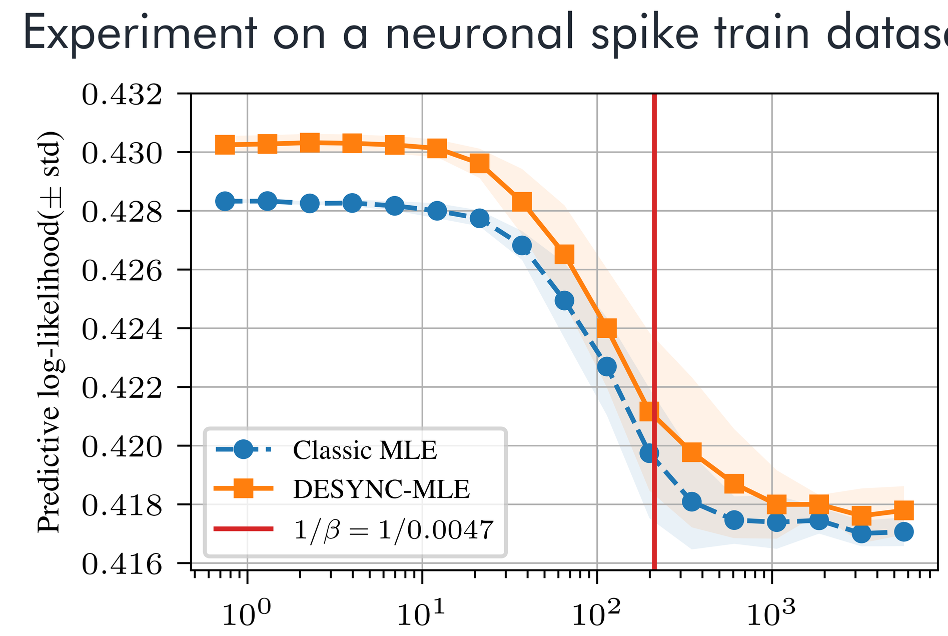
Experimental Results

- Experiment on Synthetic data

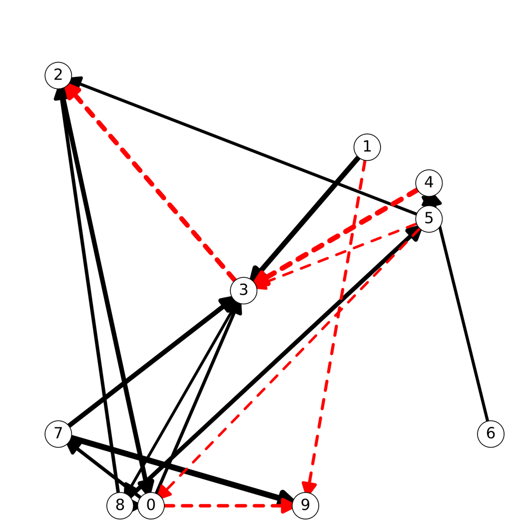


Analysis of the sensitivity to the noise scale (averaged over 10 simulated observations over 10 random graphs, with d=10 fixed)

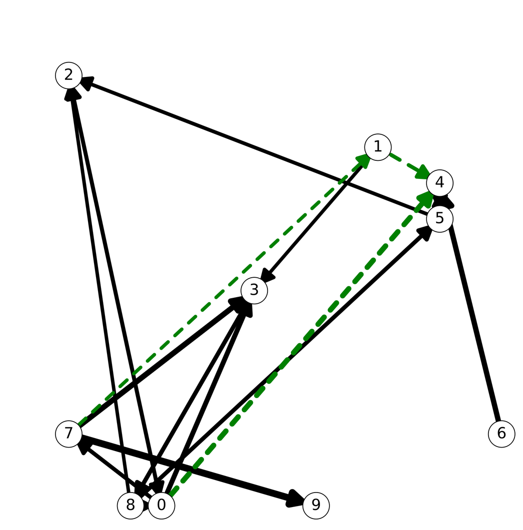
- Experiment on a neuronal spike train dataset



Analysis of the robustness of the DESYNC-MHP maximum likelihood estimator to various noise levels



Causal graph learnt by the classic MLE.



Causal graph learnt by the DESYNC-MHP MLE.

For more information, see: *W. Trouleau, J. Etesami, M. Grossglauser, N. Kiyavash, and P. Thiran.* **Learning Hawkes Processes Under Synchronization Noise**. *Proceedings of the 36th International Conference on Machine Learning. (ICML 2019)*