

Learning Hawkes Processes Under Synchronization Noise



William Trouleau

Jalal Etesami

Negar Kiyavash

Matthias Grossglauser

Patrick Thiran



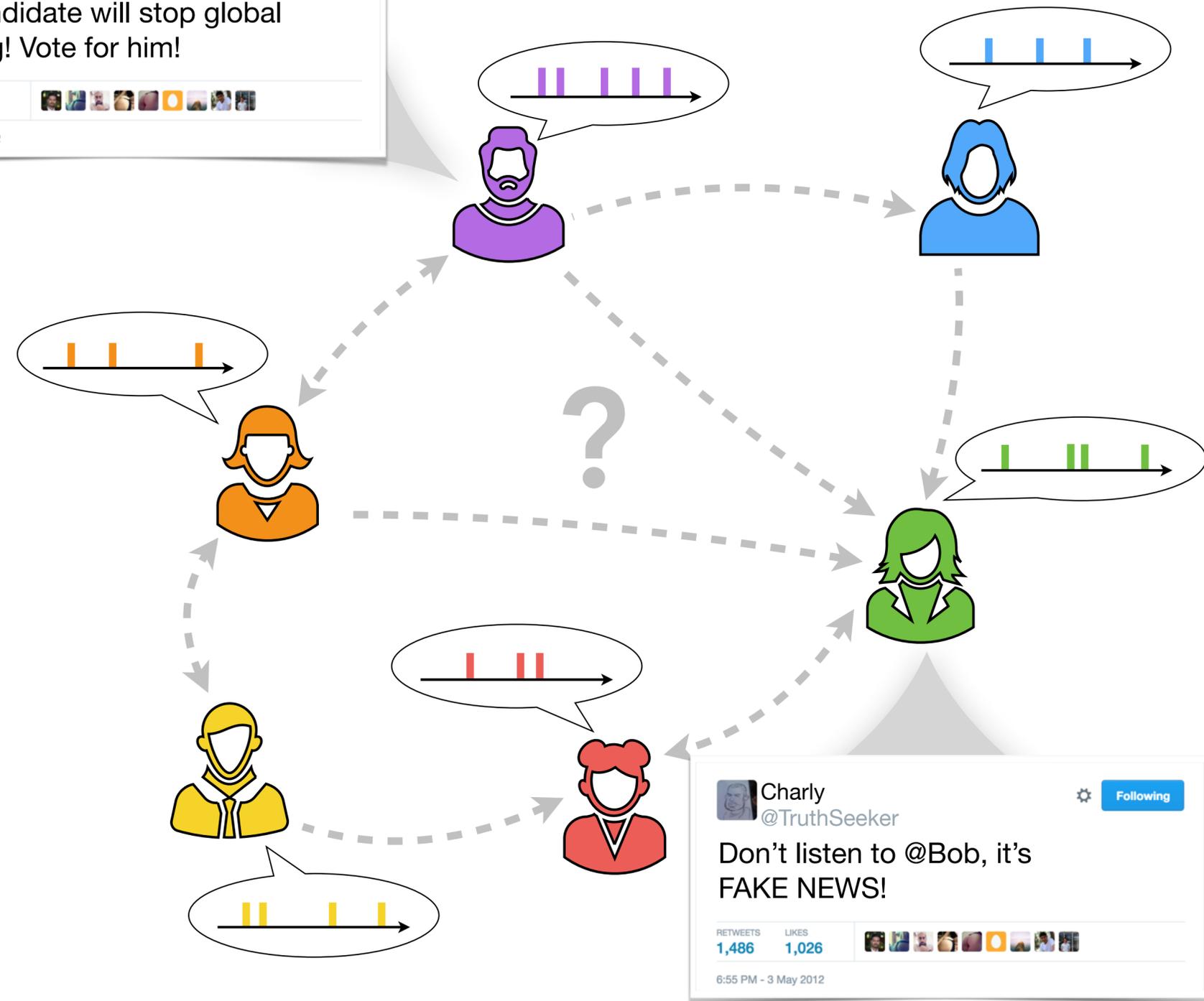
Presented at ICML'19
on Tue Jun 11th 2019

Question of interest

Learning the **causal structure** of
networks of multivariate time series
in continuous time

Example 1: Information Diffusion

- Consider a **network of users**
- We observe a **sequence of discrete events in continuous time: tweets, Facebook posts...**



Example 1: Information Diffusion

- Consider a **network of users**
- We observe a **sequence of discrete events** in **continuous time**: **tweets, Facebook posts...**
- Questions of interest:
Who **influences** whom?
How does **fake news spread**?



Fri 7 Jun 2019 08.53 EDT

The Guardian

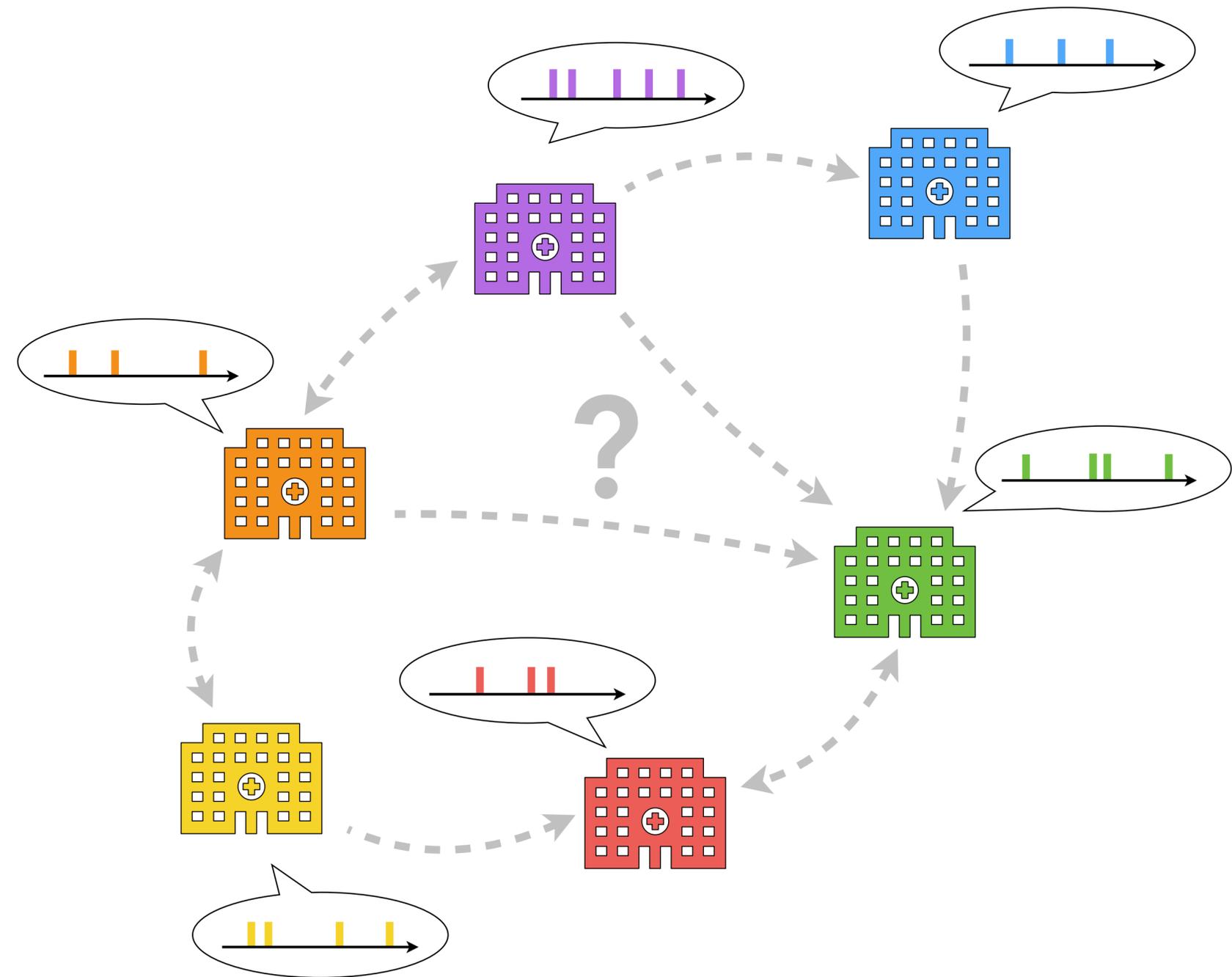
Half of Americans see fake news as bigger threat than terrorism, study finds

Almost 70% of Americans feel fake news has greatly affected their confidence in government institutions, a new study says



Example 2: Disease Dynamics

- Consider a **network of hospitals**
- We observe a **sequence of discrete events in continuous time: interactions, infections, recoveries...**
- Questions of interest:
Who **infected** whom?
How does the **disease spread**?
How to **control** it?



Example 2: Disease Dynamics

- Consider a **network of hospitals**
- We observe a a **sequence of discrete events in continuous time: interactions, infections, recoveries...**
- Questions of interest:
Who **infected** whom?
How does the **disease spread**?
How to **control** it?

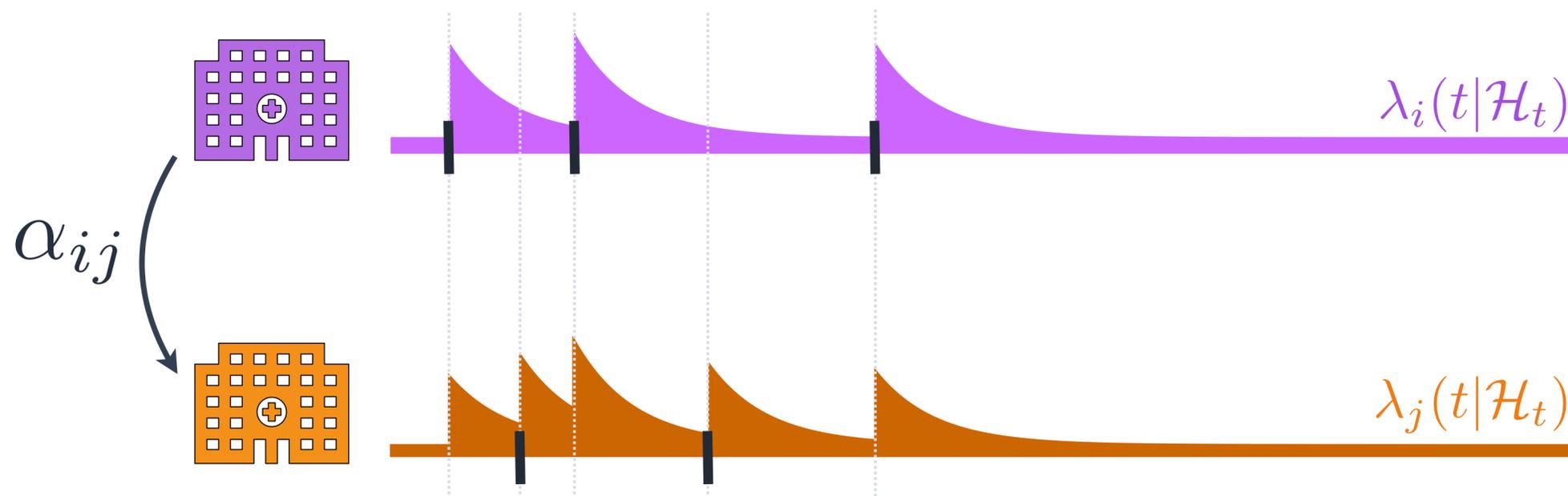


**How do we usually
solve it?**

Method: Multivariate Hawkes Process (MHP)

- Temporal Point Process
- Widely used model to learn **causal structure between time series**
- Captures **mutually exciting** patterns of influence between dimensions

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



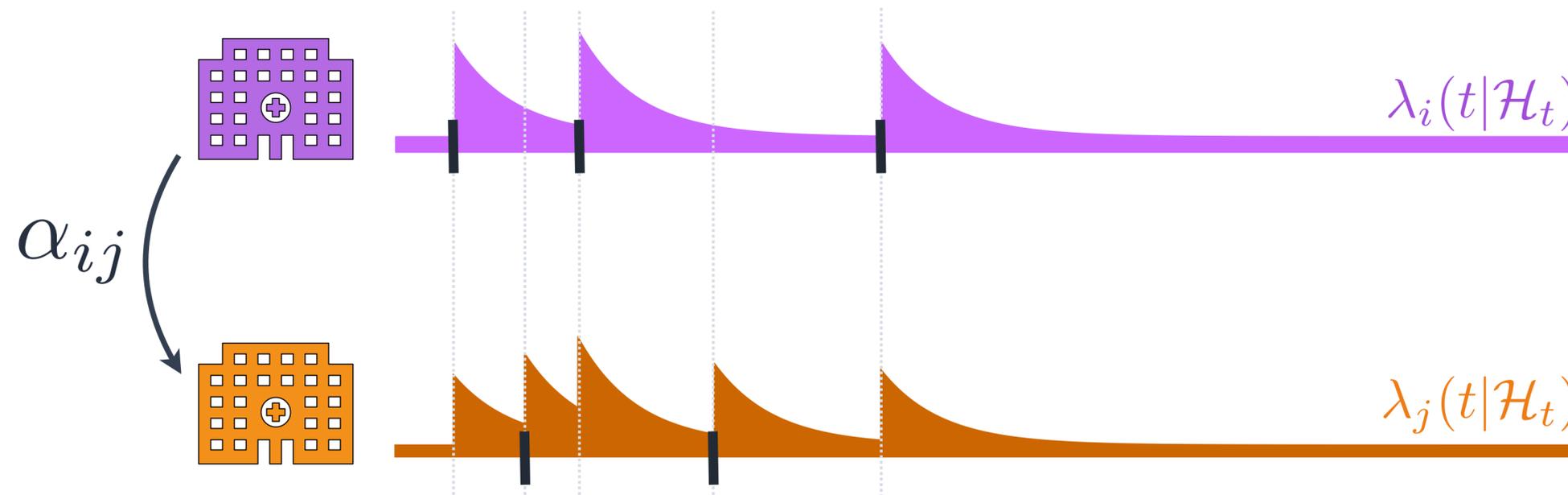
Method: Multivariate Hawkes Process (MHP)

- Temporal Point Process
- Widely used model to learn **causal structure between time series**
- Captures **mutually exciting** patterns of influence between dimensions

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

Exogenous intensity:
constant, independent
of the past

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



Method: Multivariate Hawkes Process (MHP)

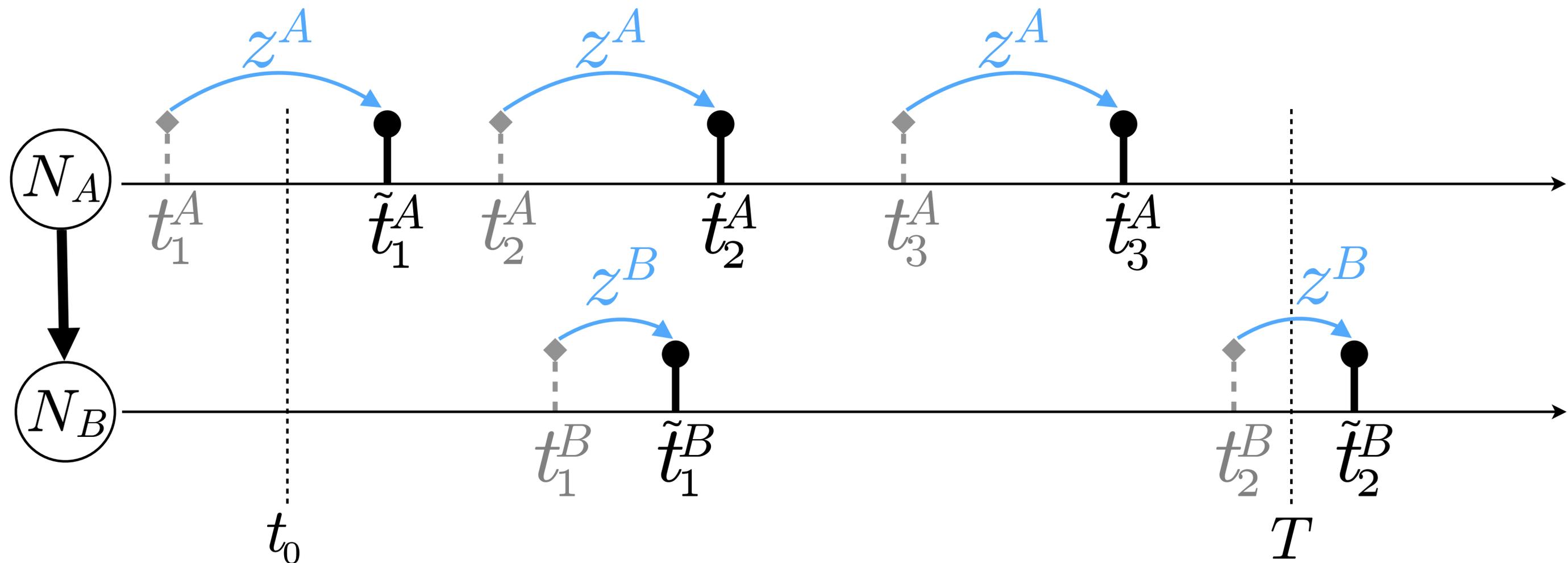


- Prior work assume **perfect traces without noise**
- What if the observed stream of events is subject to a **random and unknown time shift?**

**How to learn MHPs under
noisy observations?**

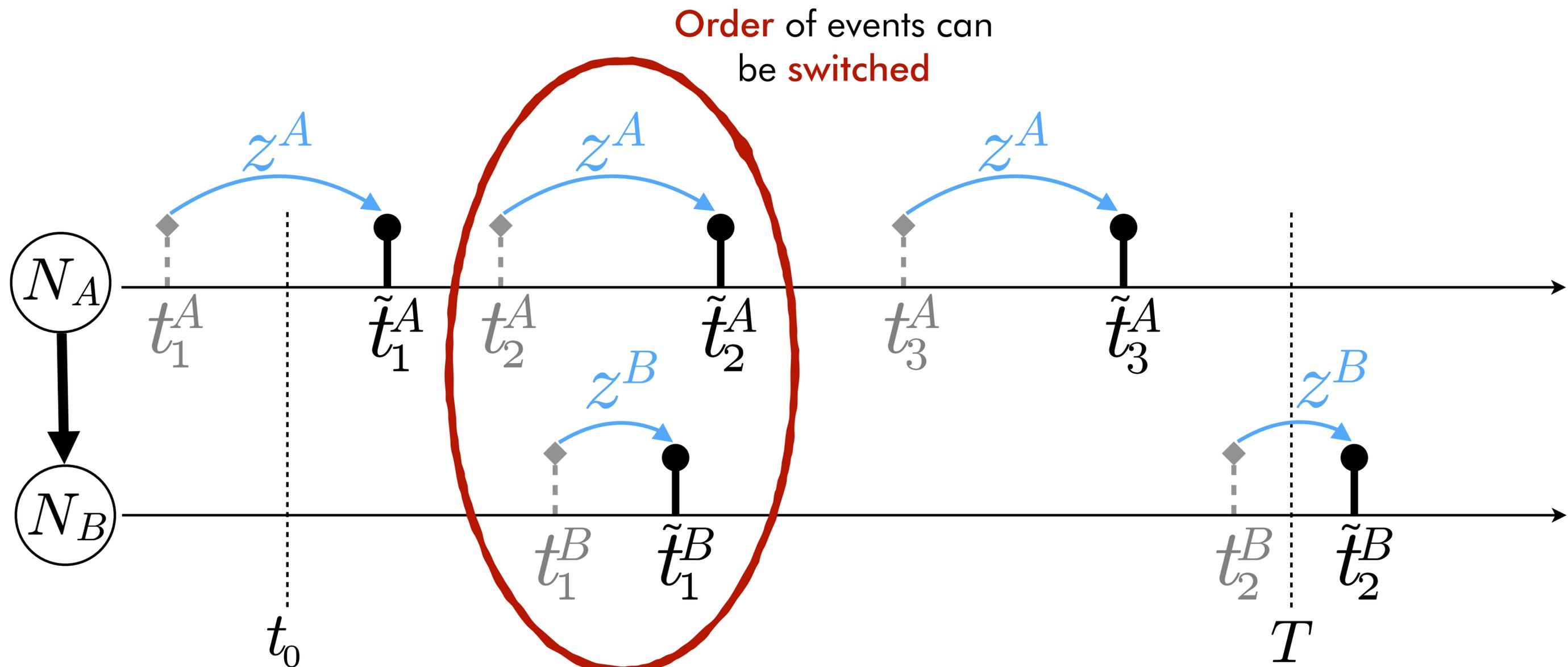
Multivariate Hawkes Process under Synchronization Noise

- What it events have **systematic measurement errors**?



Multivariate Hawkes Process under Synchronization Noise

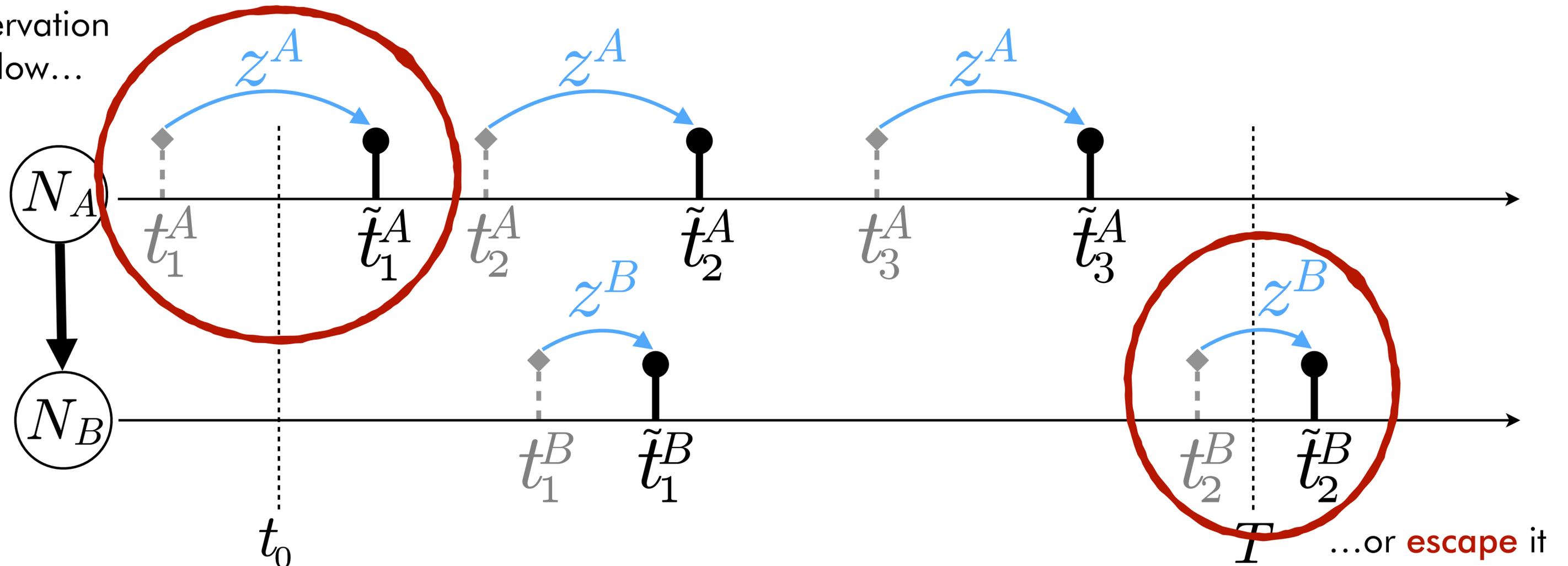
- What if events have **systematic measurement errors**?



Multivariate Hawkes Process under Synchronization Noise

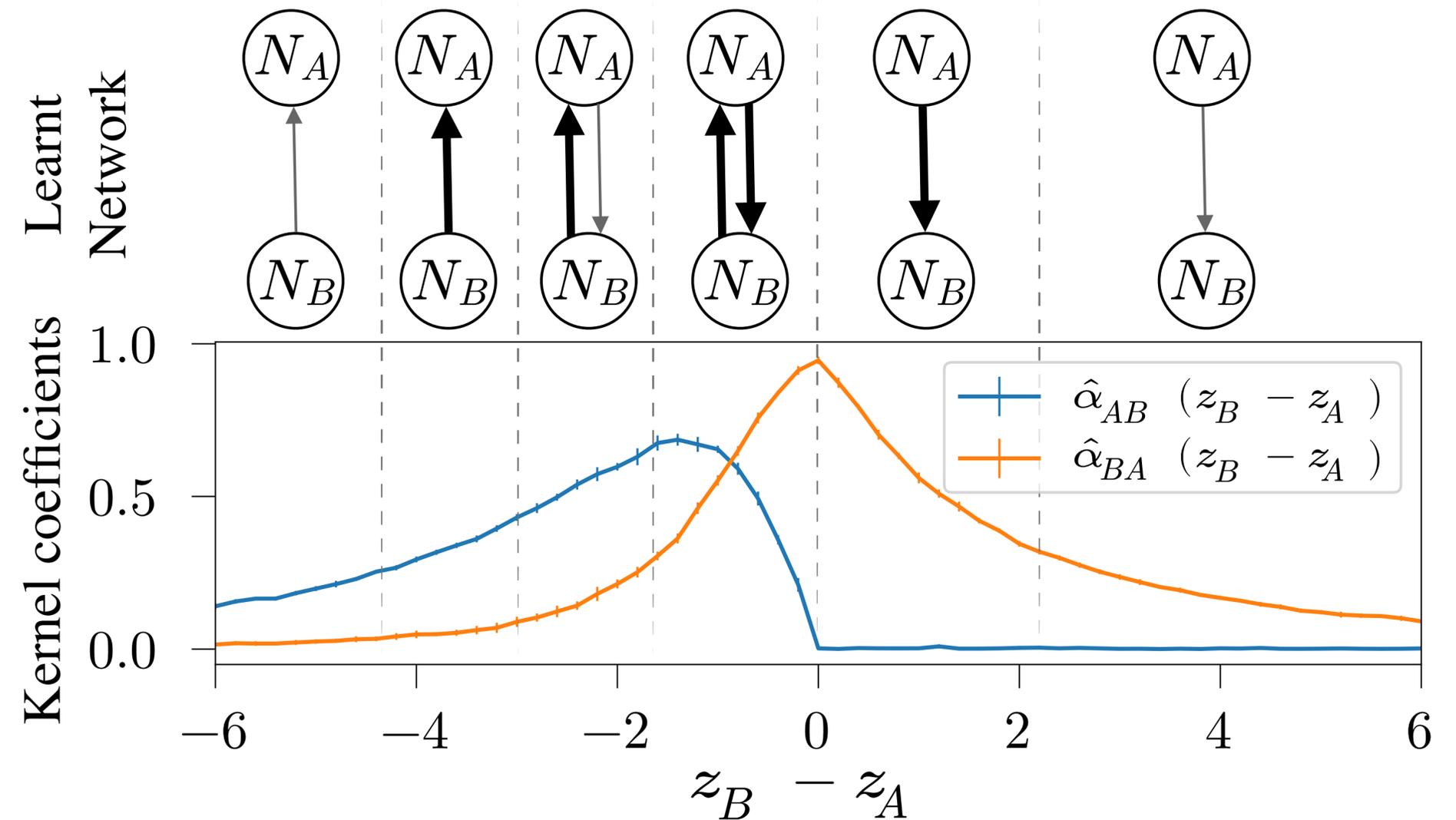
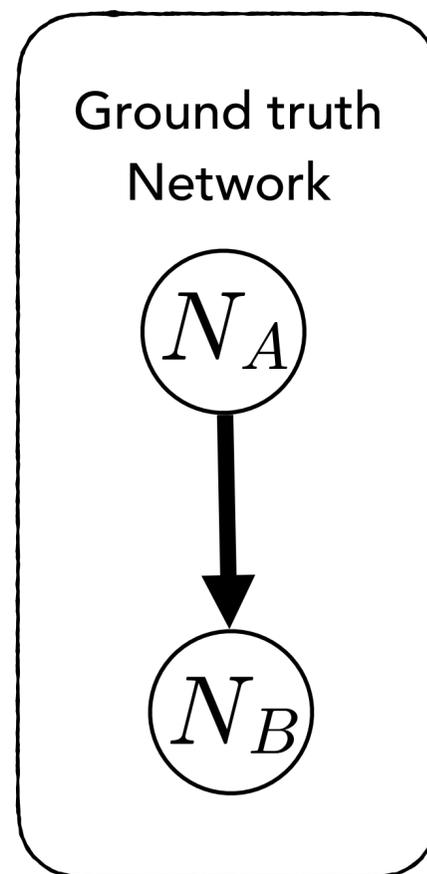
- What if events have **systematic measurement errors**?

Events can **enter** the observation window...



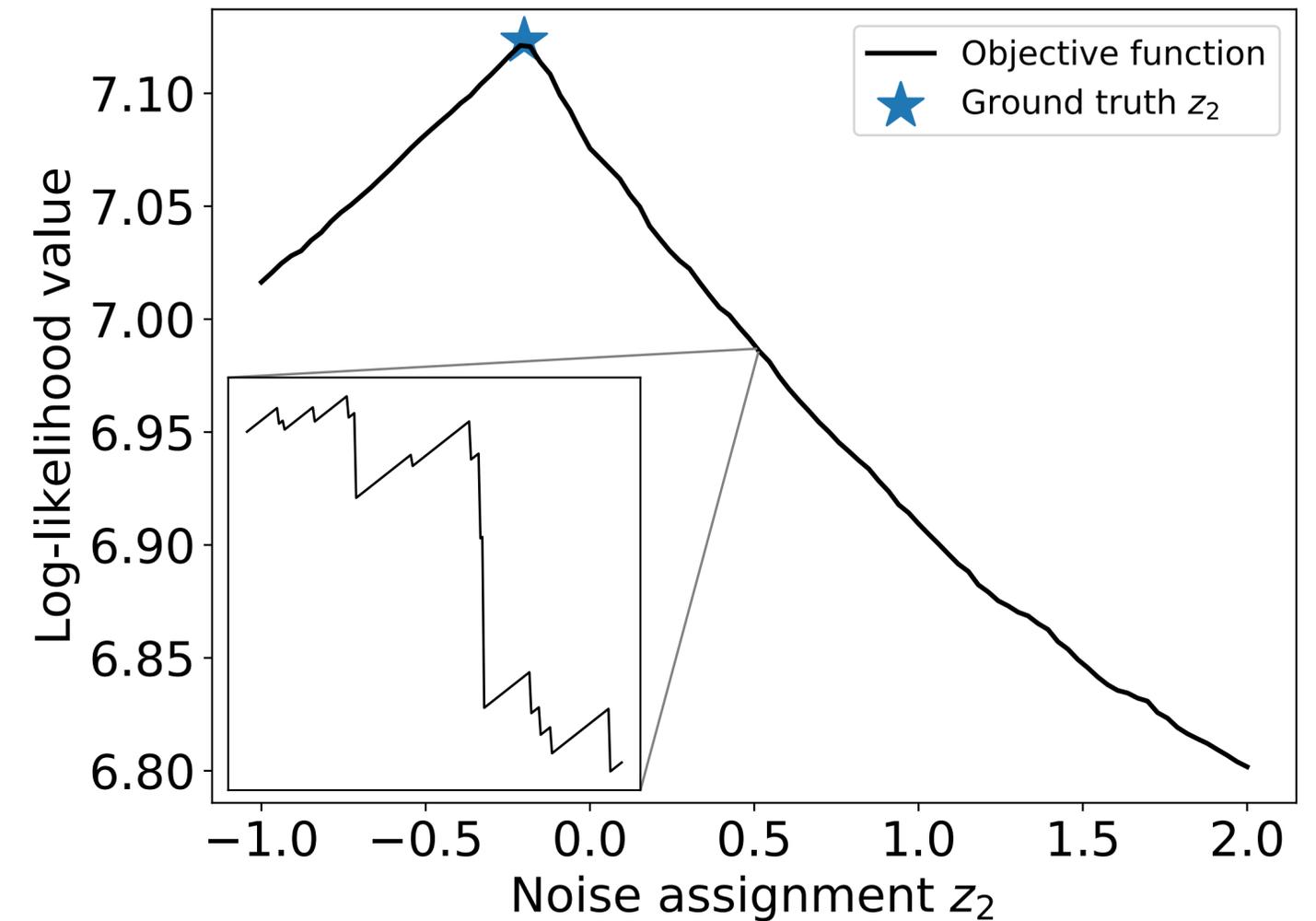
Multivariate Hawkes Process under Synchronization Noise

- What events have **systematic measurement errors**?
- Edges learnt by maximum likelihood estimation can be significantly affected by even small delays



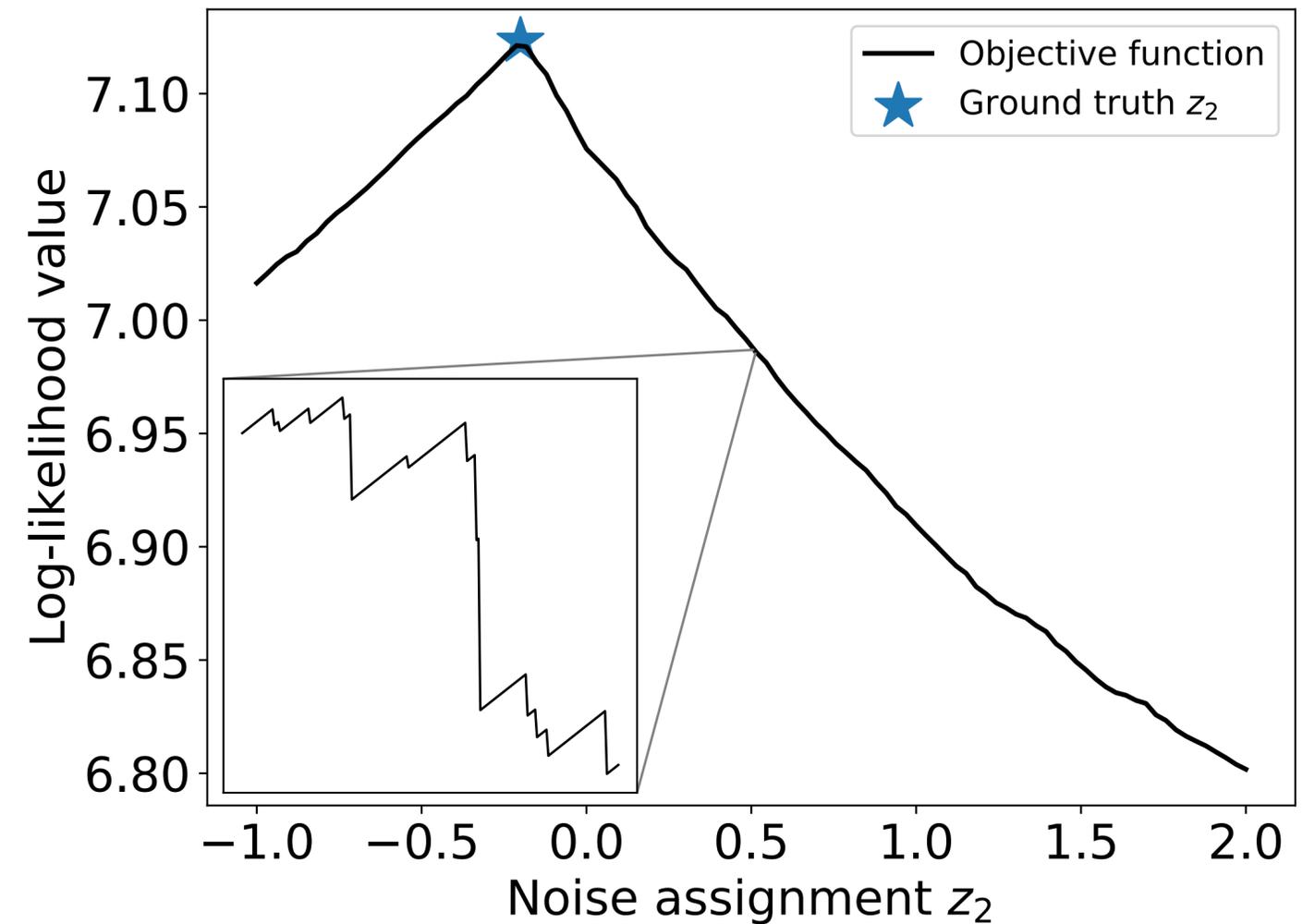
New approach DESYNC-MHP

- Idea:
 - Consider the noise as parameters
 - Maximize the joint log-likelihood over both MHP parameters and noise



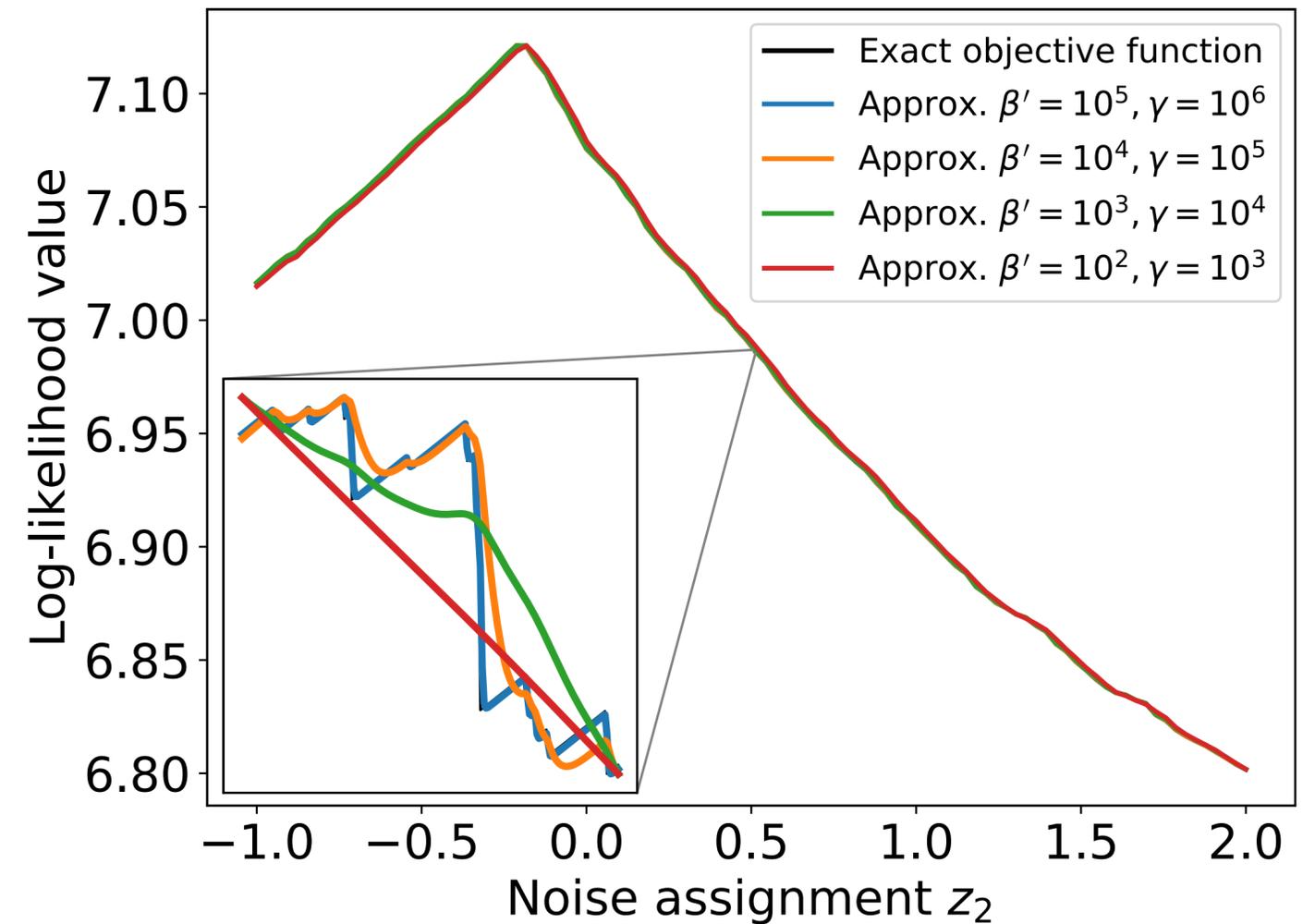
New approach DESYNC-MHP

- Idea:
 - Consider the noise as parameters
 - Maximize the joint log-likelihood over both MHP parameters and noise
- Challenges: resulting objective is
 - Non-smooth
 - Non-convex

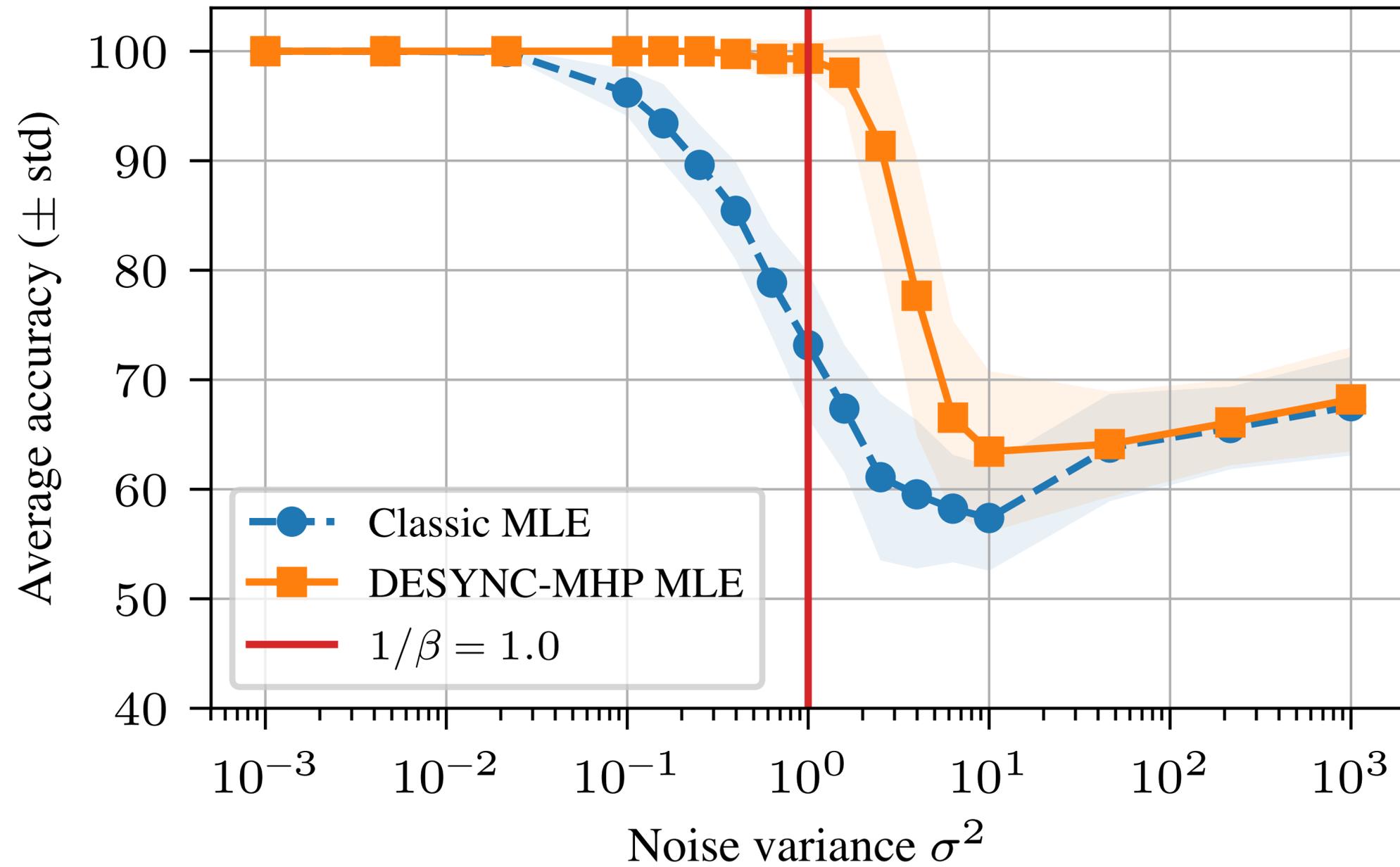


New approach DESYNC-MHP

- **Idea:**
 - Consider the noise as parameters
 - Maximize the joint log-likelihood over both MHP parameters and noise
- **Challenges:** resulting objective is
 - Non-smooth
 - Non-convex
- **Solution:**
 - Approximate the objective with a smooth approximation
 - Use SGD to escape local minima



Experimental Results



Learning Hawkes Processes Under Synchronization Noise

William Trouleau

Jalal Etesami

Negar Kiyavash

Matthias Grossglauser

Patrick Thiran

EPFL

**Georgia
Tech** 



Come check out our poster tonight !