Combining
Generative and Discriminative Models
for Hybrid Inference

Víctor Garcia Satorras

Zeynep Akata

Max Welling

UNIVERSITY OF AMSTERDAM

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

UvA - BOSCH DELTA LAB

Background | Inference

Given a set of observations $\quad$ $\mathbf{y}$

We want to **infer** the states of a process $\quad$ $\mathbf{x}$

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}}\, p(\mathbf{x}, \mathbf{y})$$
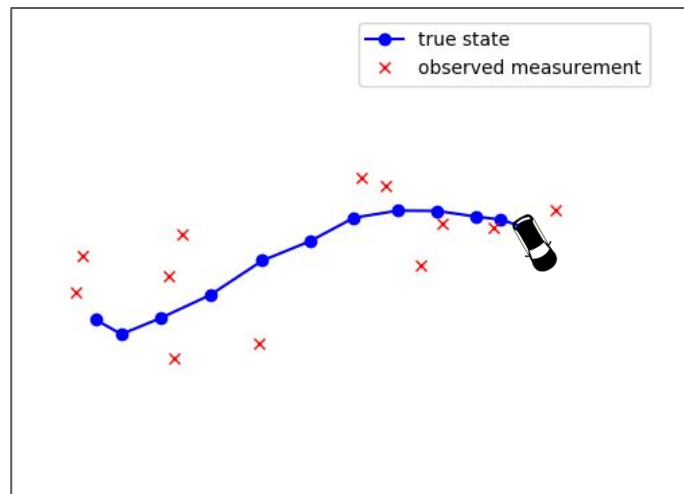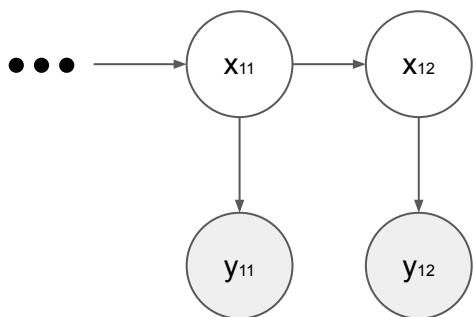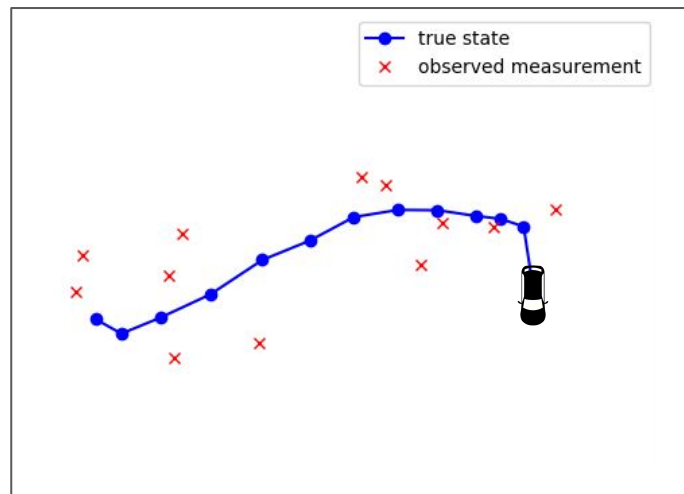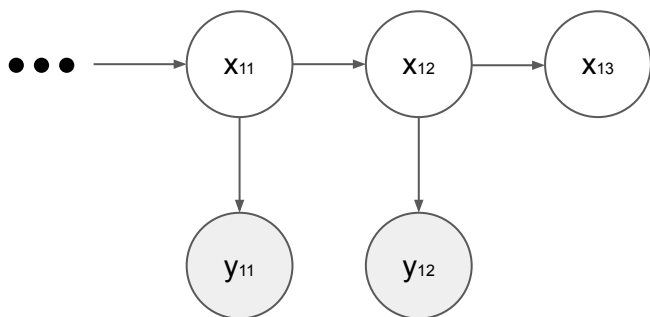
Background | Inference

Given a set of observations $\mathbf{y}$

We want to **infer** the states of a process $\mathbf{x}$

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\mathrm{argmax}}\, p(\mathbf{x}, \mathbf{y})$$

Background | Inference

Given a set of observations $\mathbf{y}$

We want to **infer** the states of a process $\mathbf{x}$

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}} \, p(\mathbf{x}, \mathbf{y})$$
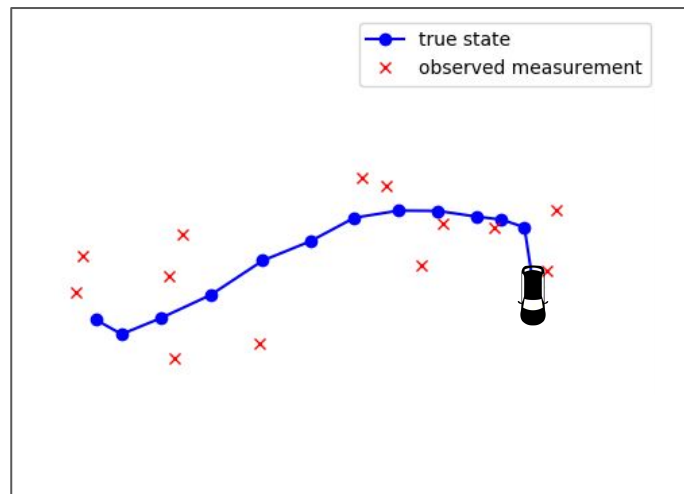
Background | Inference

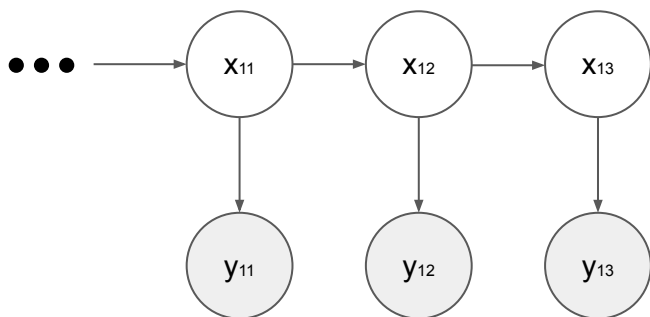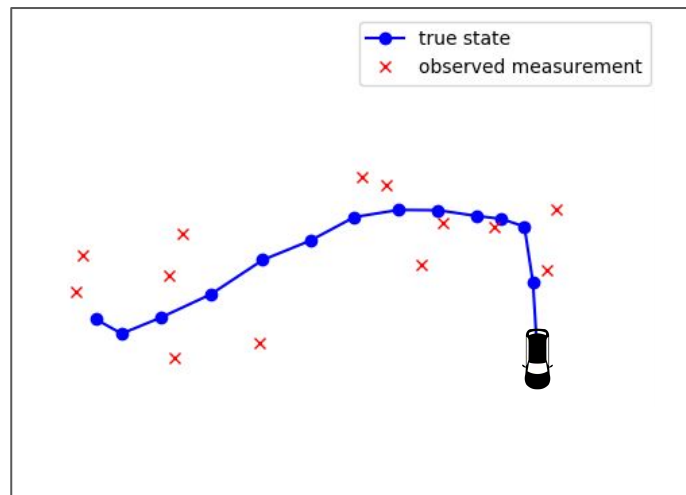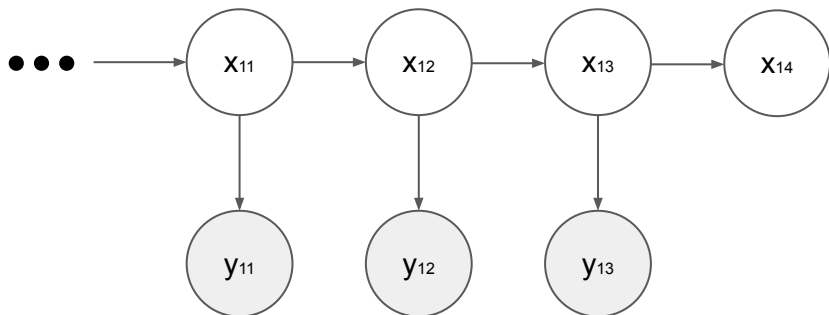**Given** a set of observations $\mathbf{y}$

We want to **infer** the states of a process $\mathbf{x}$

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}}\, p(\mathbf{x}, \mathbf{y})$$
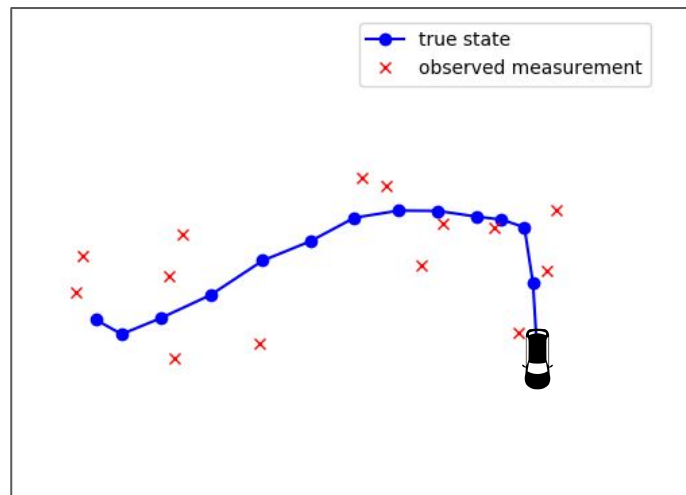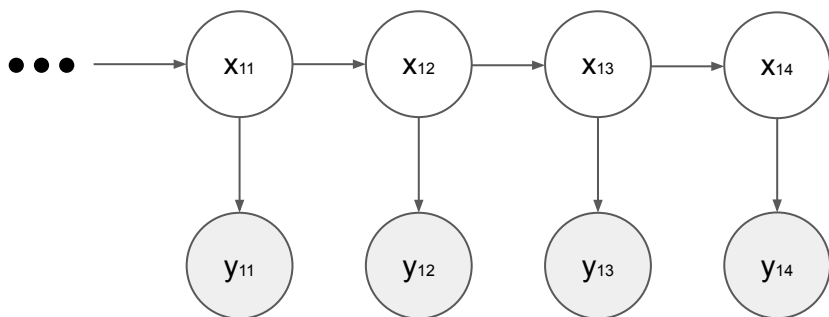
Background | Inference

Given a set of observations $\mathbf{y}$

We want to **infer** the states of a process $\mathbf{x}$

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\mathrm{argmax}} \, p(\mathbf{x}, \mathbf{y})$$
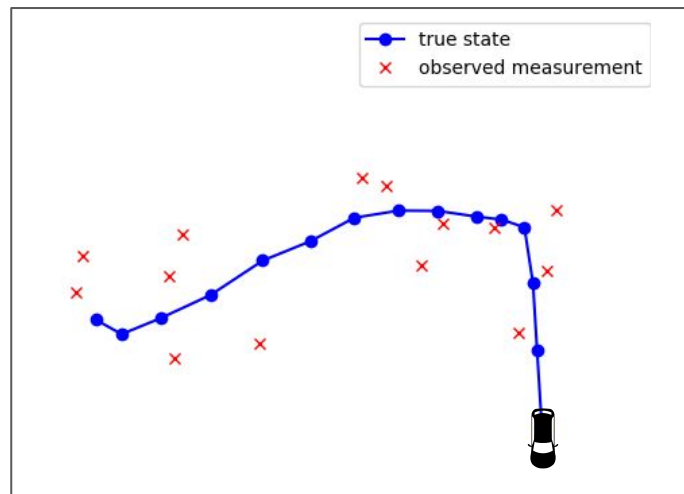
Background | Inference

Given a set of observations $\mathbf{y}$

We want to **infer** the states of a process $\mathbf{x}$

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\mathrm{argmax}}\, p(\mathbf{x}, \mathbf{y})$$
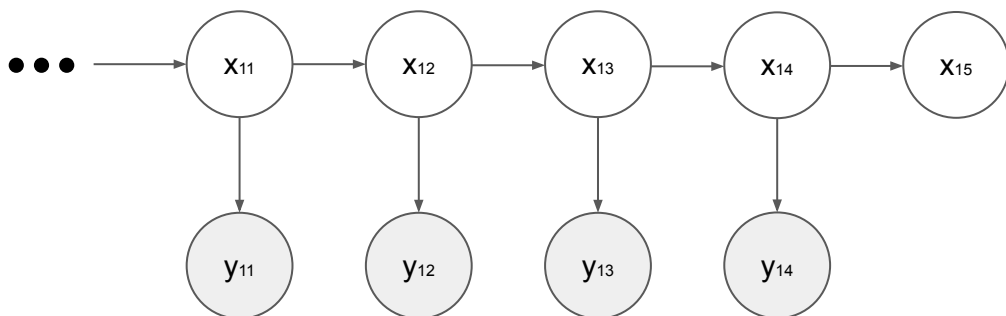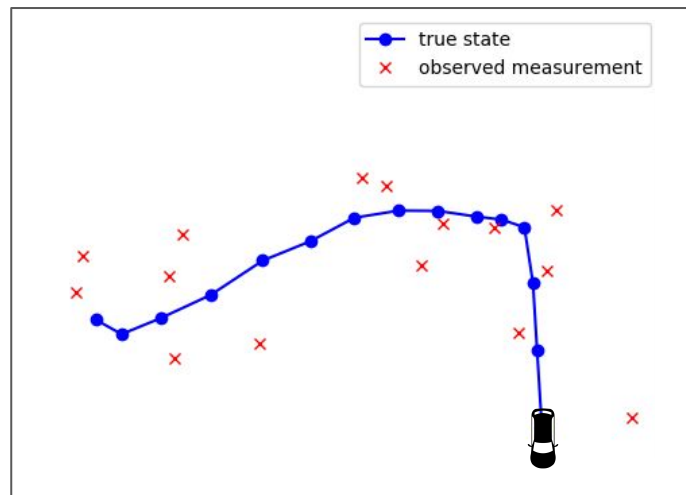
Background | Inference

Given a set of observations $\mathbf{y}$

We want to **infer** the states of a process $\mathbf{x}$

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}} \, p(\mathbf{x}, \mathbf{y})$$
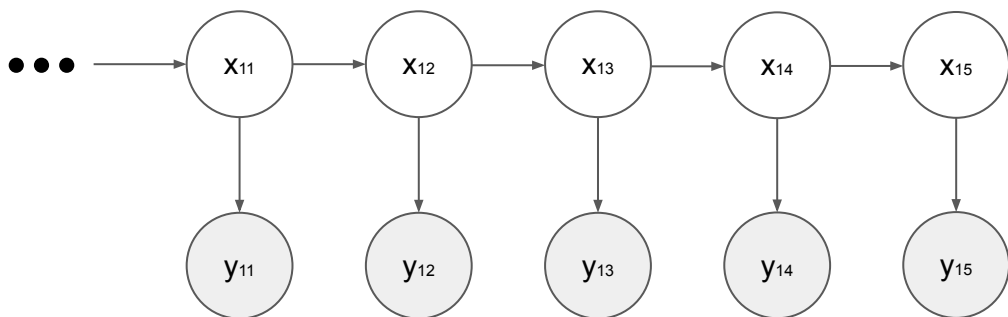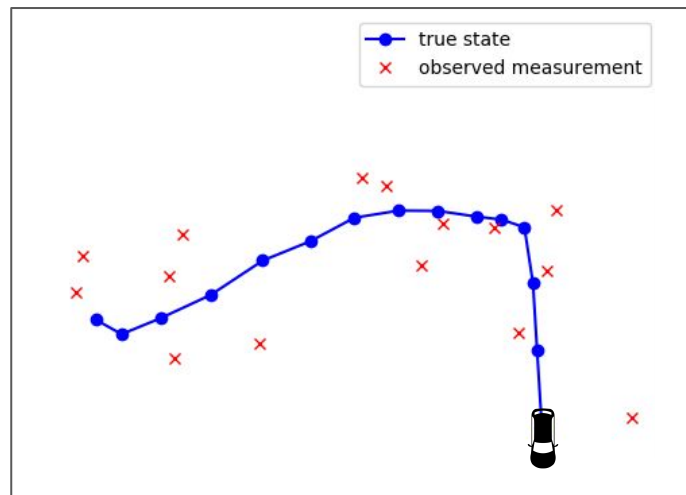
Background | Inference

Given a set of observations $\mathbf{y}$

We want to **infer** the states of a process $\mathbf{x}$

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\mathrm{argmax}}\, p(\mathbf{x}, \mathbf{y})$$

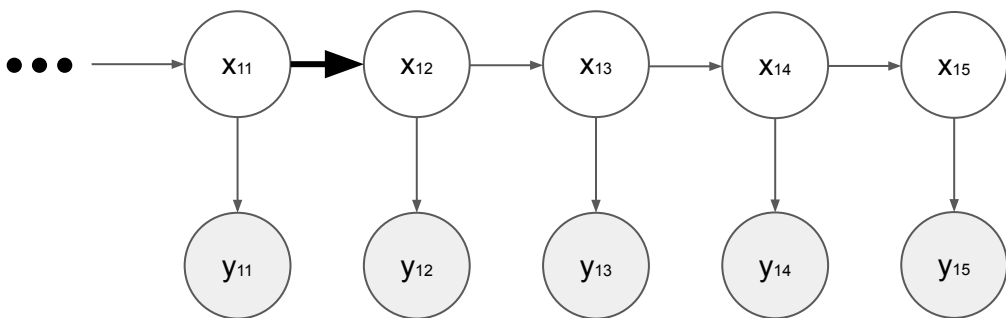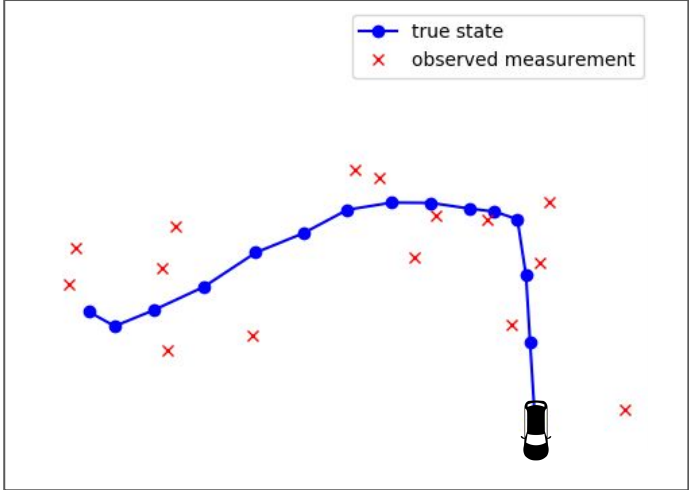# Background | Inference

Given a set of observations $\mathbf{y}$

We want to **infer** the states of a process $\mathbf{x}$

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}}\, p(\mathbf{x}, \mathbf{y})$$

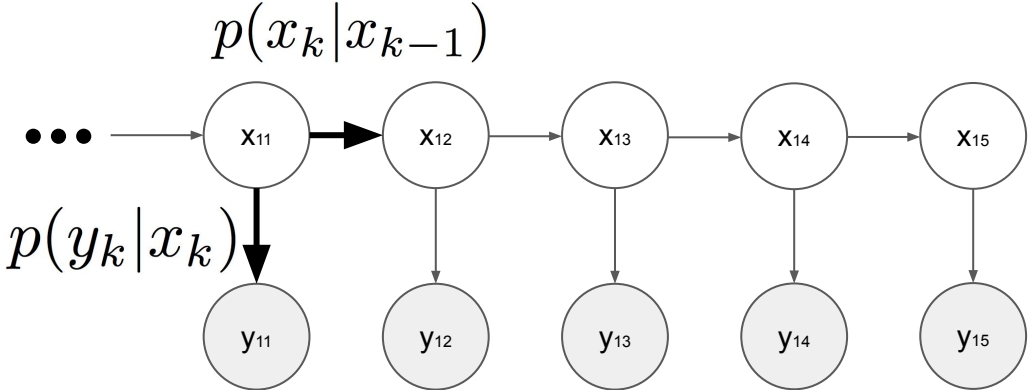$$p(x_k | x_{k-1})$$

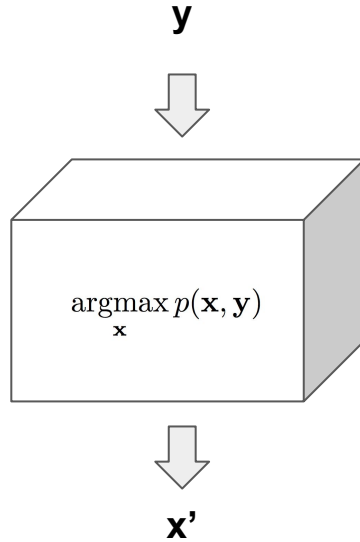# Background | Inference

Given a set of observations $\mathbf{y}$

We want to **infer** the states of a process $\mathbf{x}$

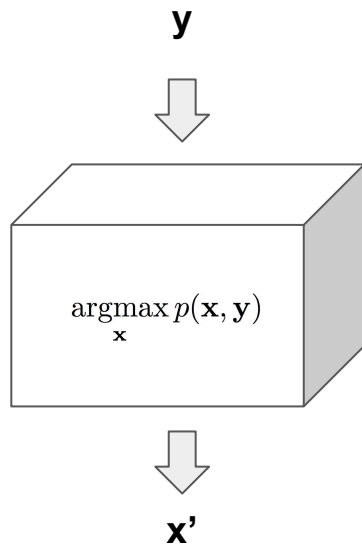$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\mathrm{argmax}}\, p(\mathbf{x}, \mathbf{y})$$
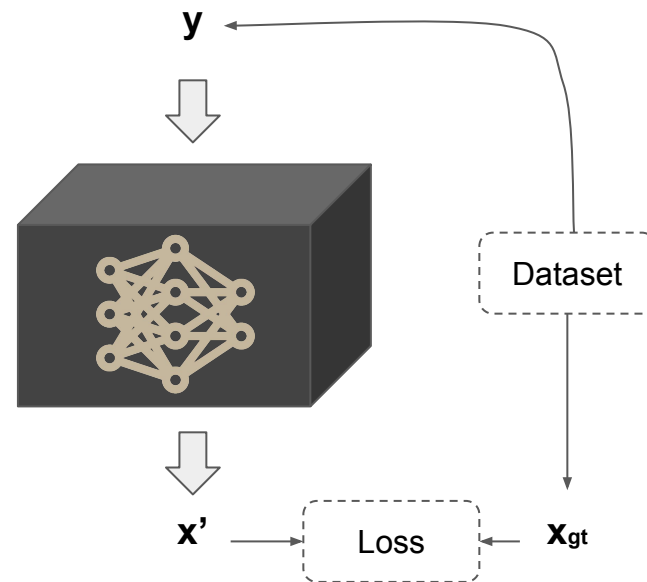
Background | Inference

y

$$\operatorname*{argmax}_{\mathbf{x}} p(\mathbf{x}, \mathbf{y})$$

x'

- Data efficient
- Interpretable
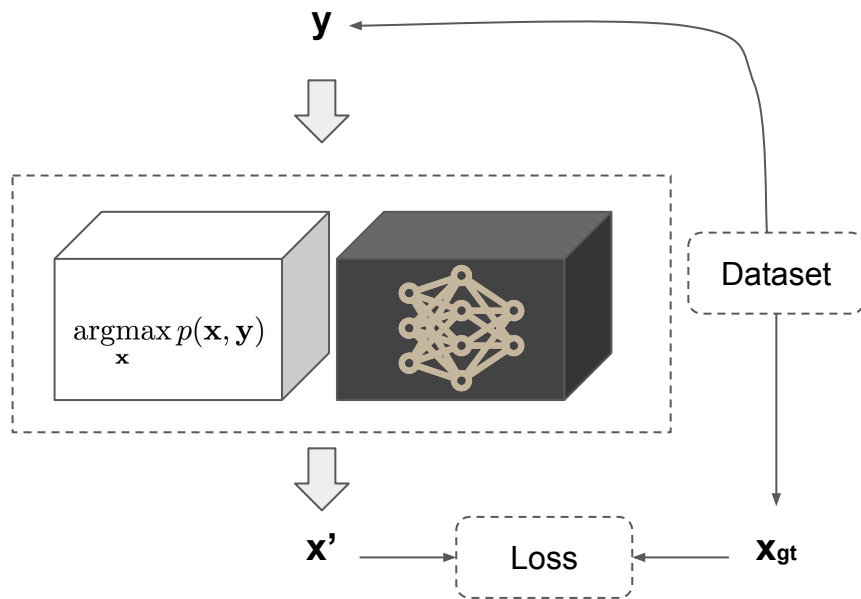- Better generalization (high bias)

Background | Inference

y

$$\underset{\mathbf{x}}{\operatorname{argmax}}\, p(\mathbf{x}, \mathbf{y})$$

x'

- Data efficient
- Interpretable
- Better generalization (high bias)

y

Dataset
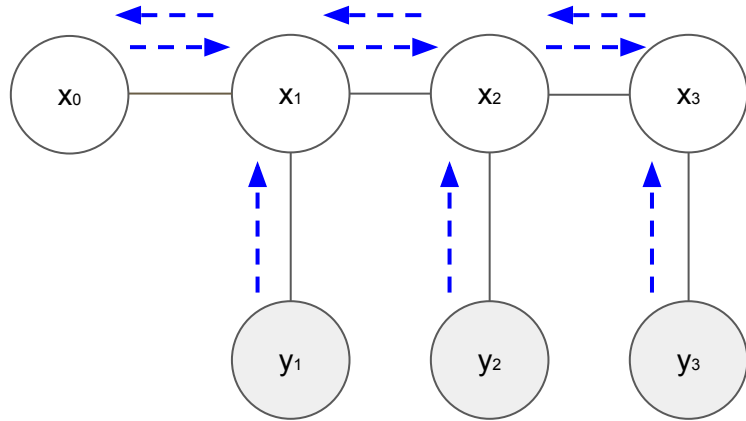
x' → Loss ← **x**$_{gt}$

- More flexible
- Accurate (large amounts of data)

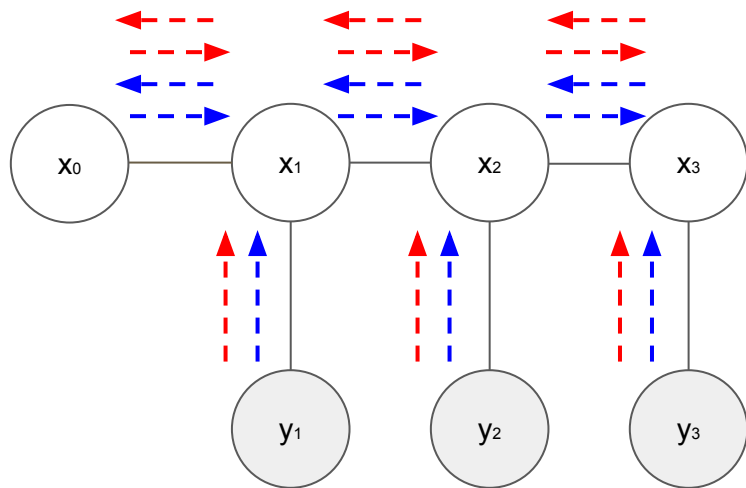# Model | Hybrid Inference

Model | Hybrid Inference

Graphical Inference (e.g. Kalman Filter)
-   **Based on prior knowledge**
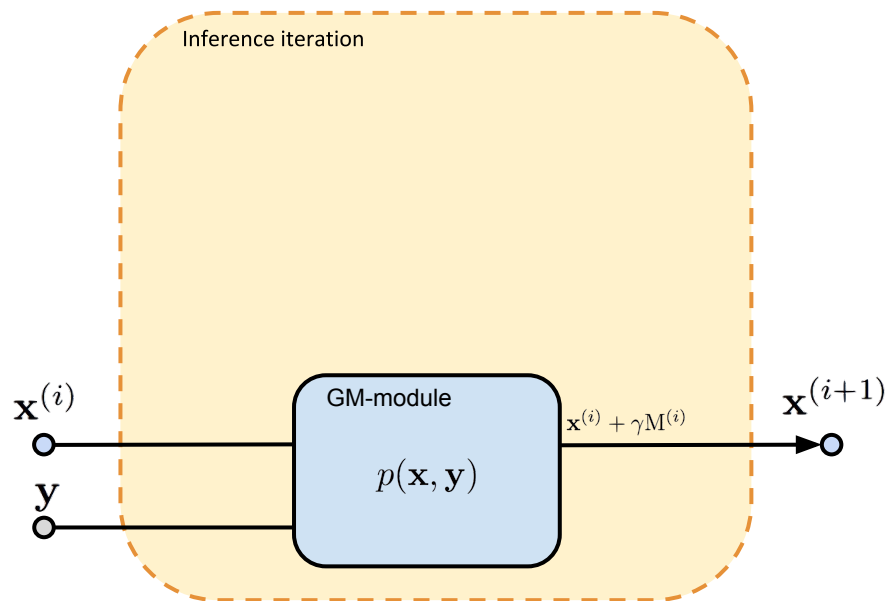
Model | Hybrid Inference

Graphical Inference (e.g. Kalman Filter)
- **Based on prior knowledge**

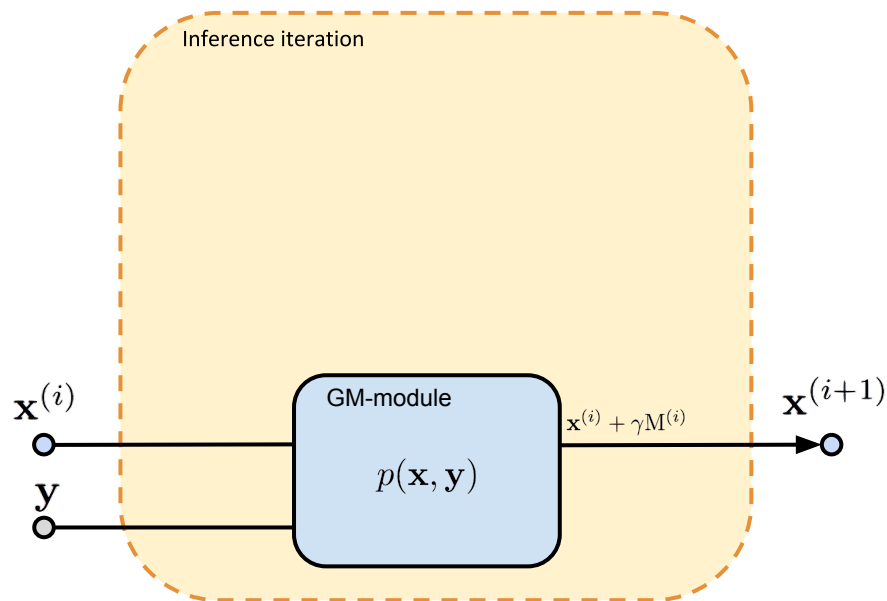**Graph Neural Network messages**
- **Learned from data**

Model | Hybrid Inference

$$x_k^{(i+1)} = x_k^{(i)} + \gamma(\mathrm{M}_k^{(i)})$$
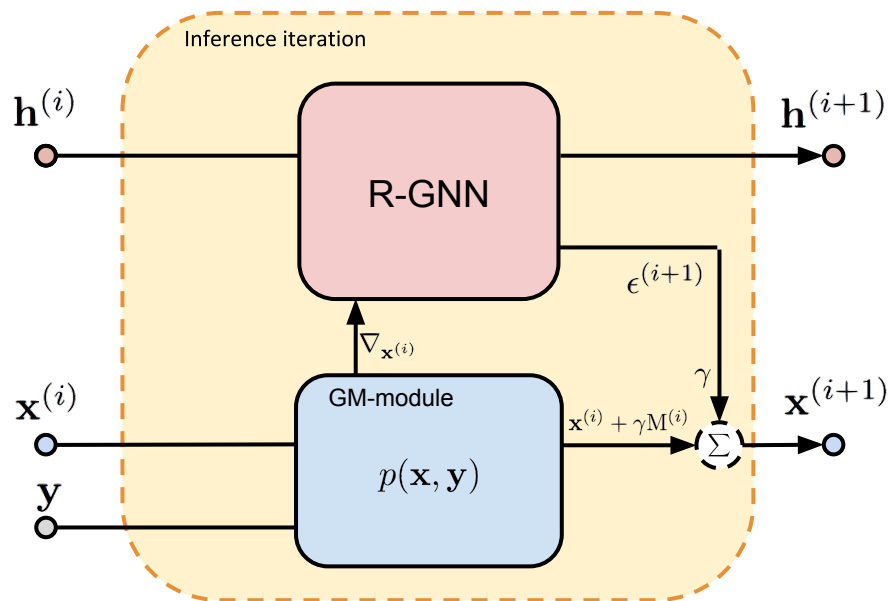
Hybrid model | Adding GNN messages

$$x_k^{(i+1)} = x_k^{(i)} + \gamma(\mathrm{M}_k^{(i)} + \epsilon_k^{(i+1)})$$

Hybrid model | Adding GNN messages

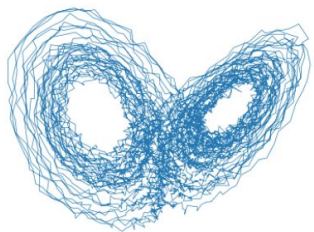$$x_k^{(i+1)} = x_k^{(i)} + \gamma(\mathrm{M}_k^{(i)} + \epsilon_k^{(i+1)})$$
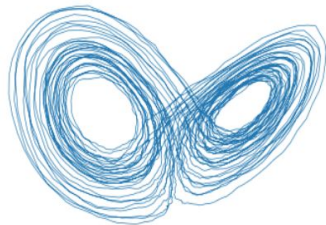
Recurrent
Graph Neural Network

Experiments | Lorenz attractor

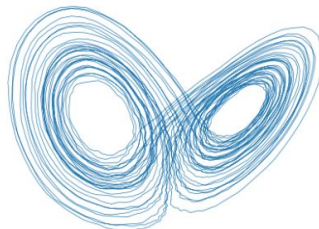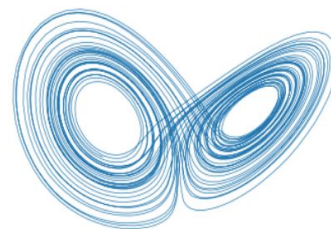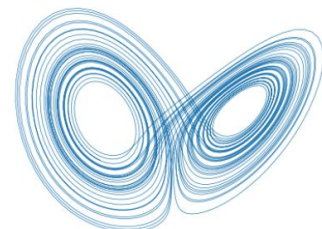0.2462 — Observations (baseline)
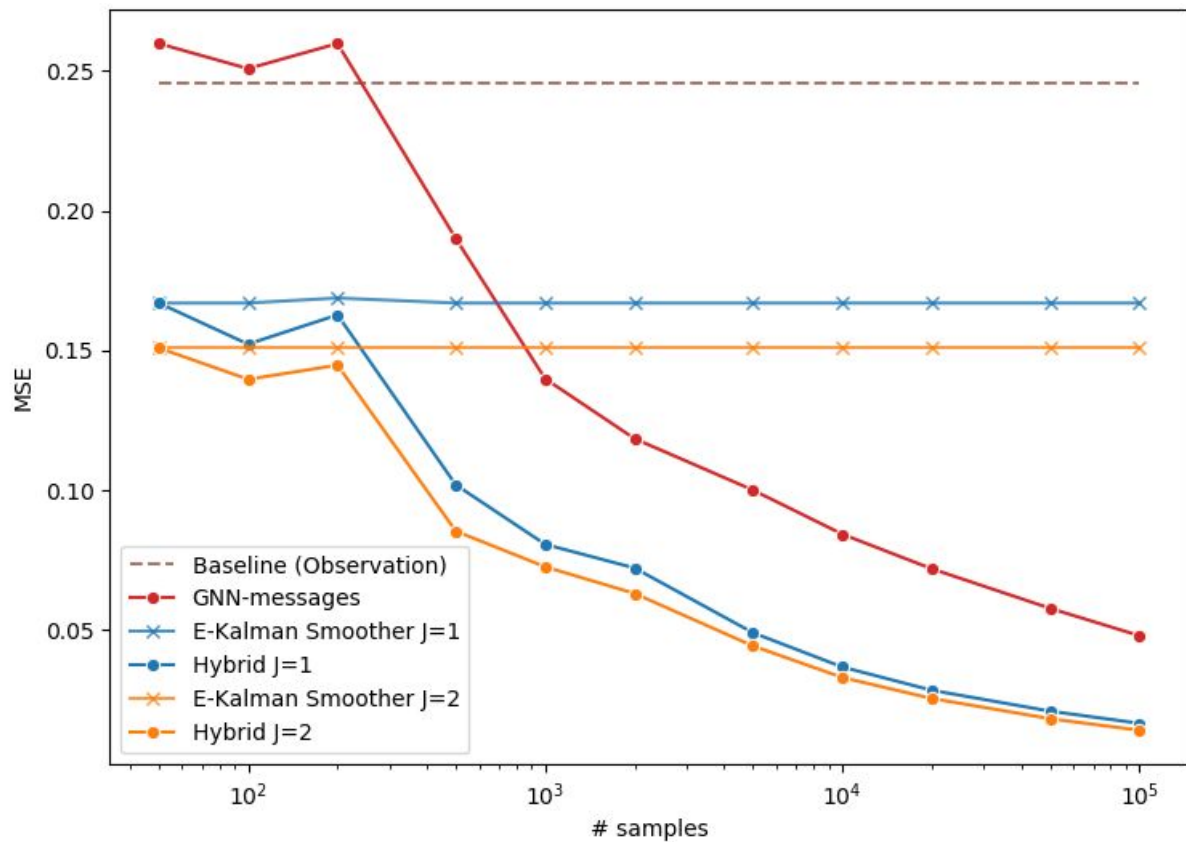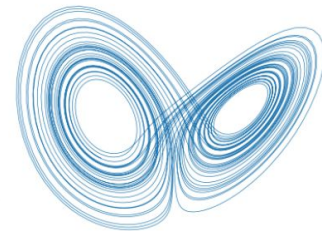
0.0613 — GNN-messages

0.0372 — E-Kalman Smoother

0.0169 — Hybrid model

0 — Ground truth

Experiments | Lorenz attractor

Combining Generative and Discriminative Models for Hybrid Inference

Víctor Garcia Satorras

Zeynep Akata

Max Welling

UNIVERSITY OF AMSTERDAM

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

UvA - BOSCH
DELTA LAB