Modeling Irregular Time Series with Continuous Recurrent Units

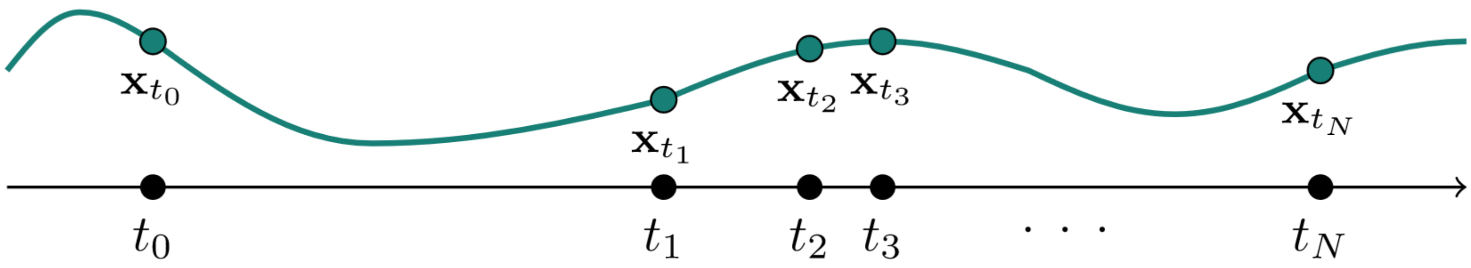
Mona Schirmer[1,†], Mazin Eltayeb[2], Stefan Lessman[1], Maja Rudolph[3]

1: Humboldt Universität zu Berlin, Germany
2: Bosch Center for AI, Germany
3: Bosch Center for AI, USA
†: Work done during an internship at Bosch Center for AI

BOSCH

Modeling Irregularly-Sampled Time Series
Motivation

Goal: Model a time-series $\mathbf{x}_\mathcal{T} = [\mathbf{x}_t | t \in \mathcal{T} = \{t_0, t_1, \cdots t_N\}]$ whose observation times $\mathcal{T} = \{t_0, t_1, \cdots t_N\}$ can occur at irregular time intervals.

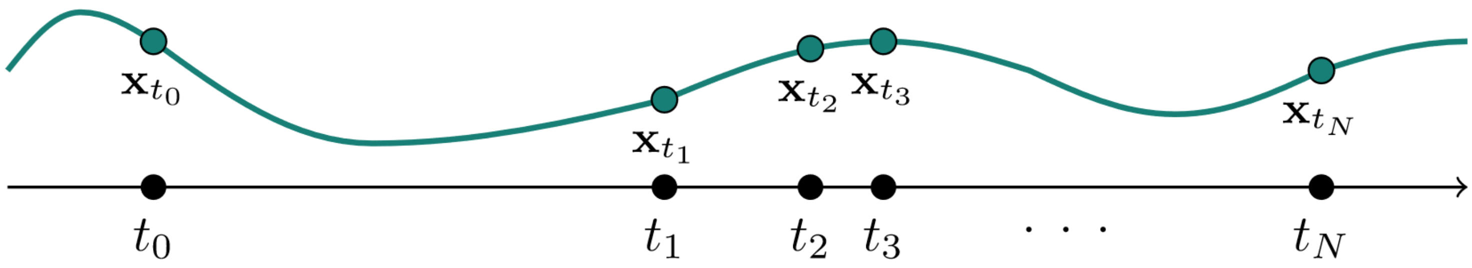


Modeling Irregularly-Sampled Time Series
Motivation

Goal: Model a time-series $\mathbf{x}_{\mathcal{T}} = [\mathbf{x}_t | t \in \mathcal{T} = \{t_0, t_1, \cdots t_N\}]$ whose observation times $\mathcal{T} = \{t_0, t_1, \cdots t_N\}$ can occur at irregular time intervals.

Challenges:

- Data from continuous processes
- Noisy and partially observed inputs
- Non-linear dynamics

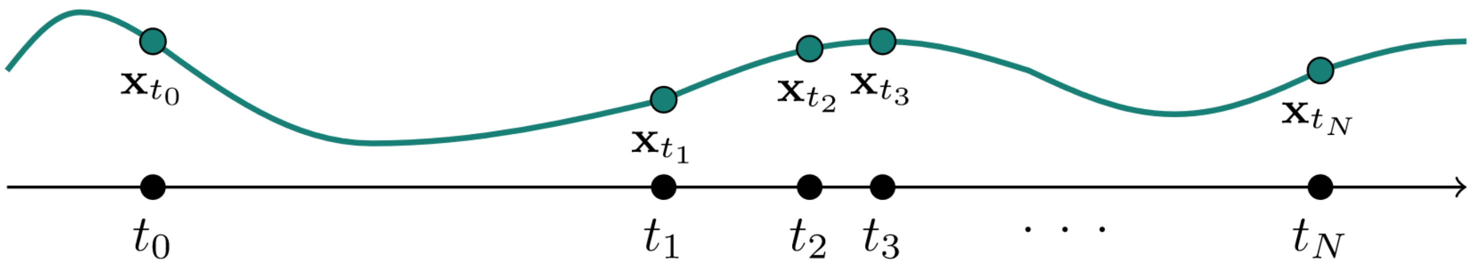


Modeling Irregularly-Sampled Time Series
Motivation

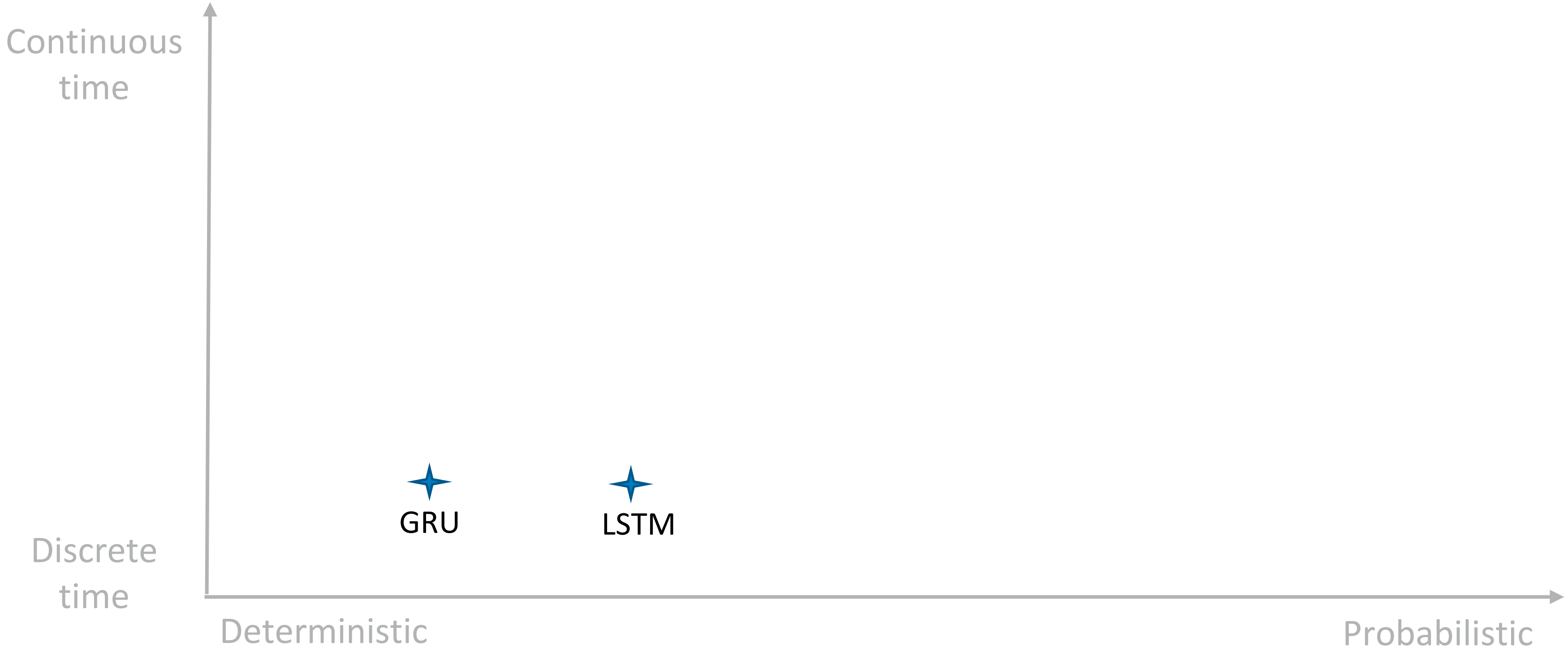
Goal: Model a time-series $\mathbf{x}_{\mathcal{T}} = [\mathbf{x}_t | t \in \mathcal{T} = \{t_0, t_1, \cdots t_N\}]$ whose observation times $\mathcal{T} = \{t_0, t_1, \cdots t_N\}$ can occur at irregular time intervals.

Challenges:

- Data from continuous processes → Continuous state dynamics
- Noisy and partially observed inputs → Uncertainty handling
- Non-linear dynamics → Expressive and flexible functions

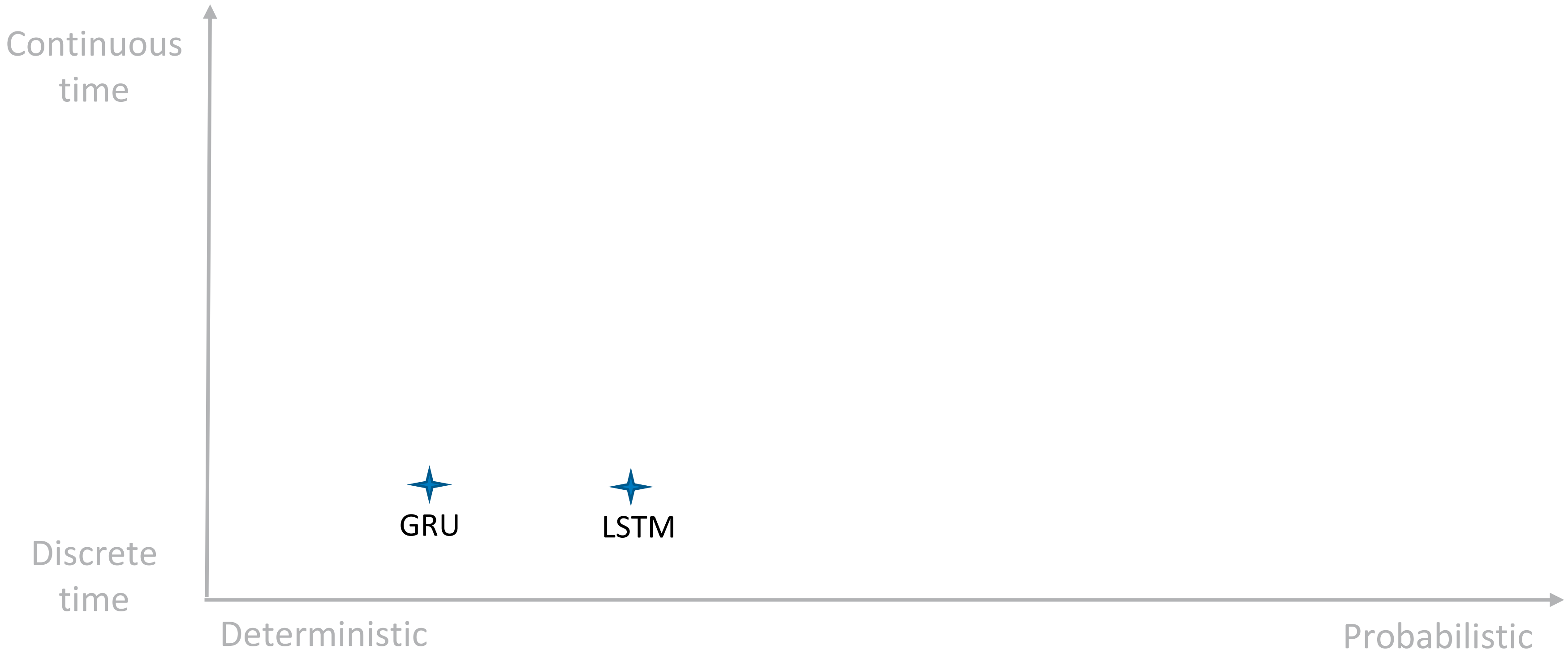


Modeling Irregularly-Sampled Time Series
Related Work



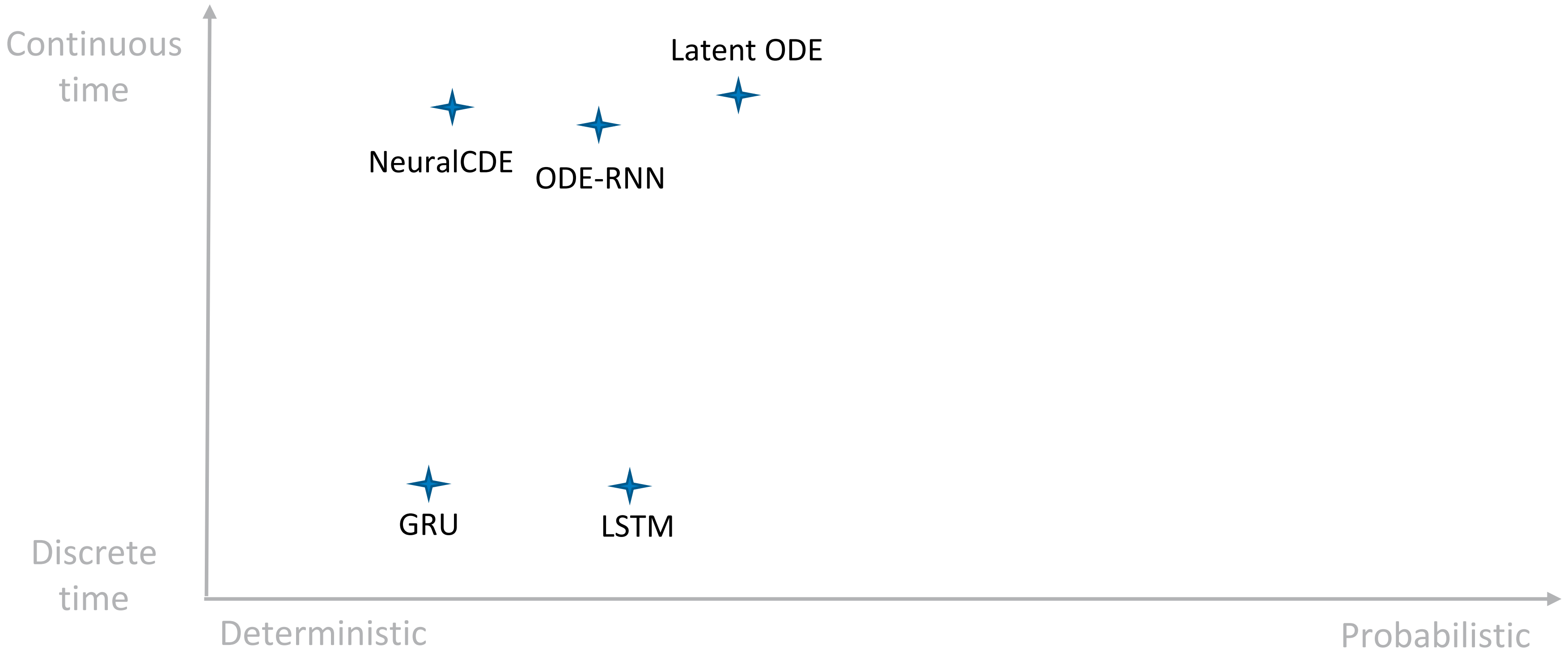
BOSCH

Modeling Irregularly-Sampled Time Series
Related Work



lack continuous dynamics

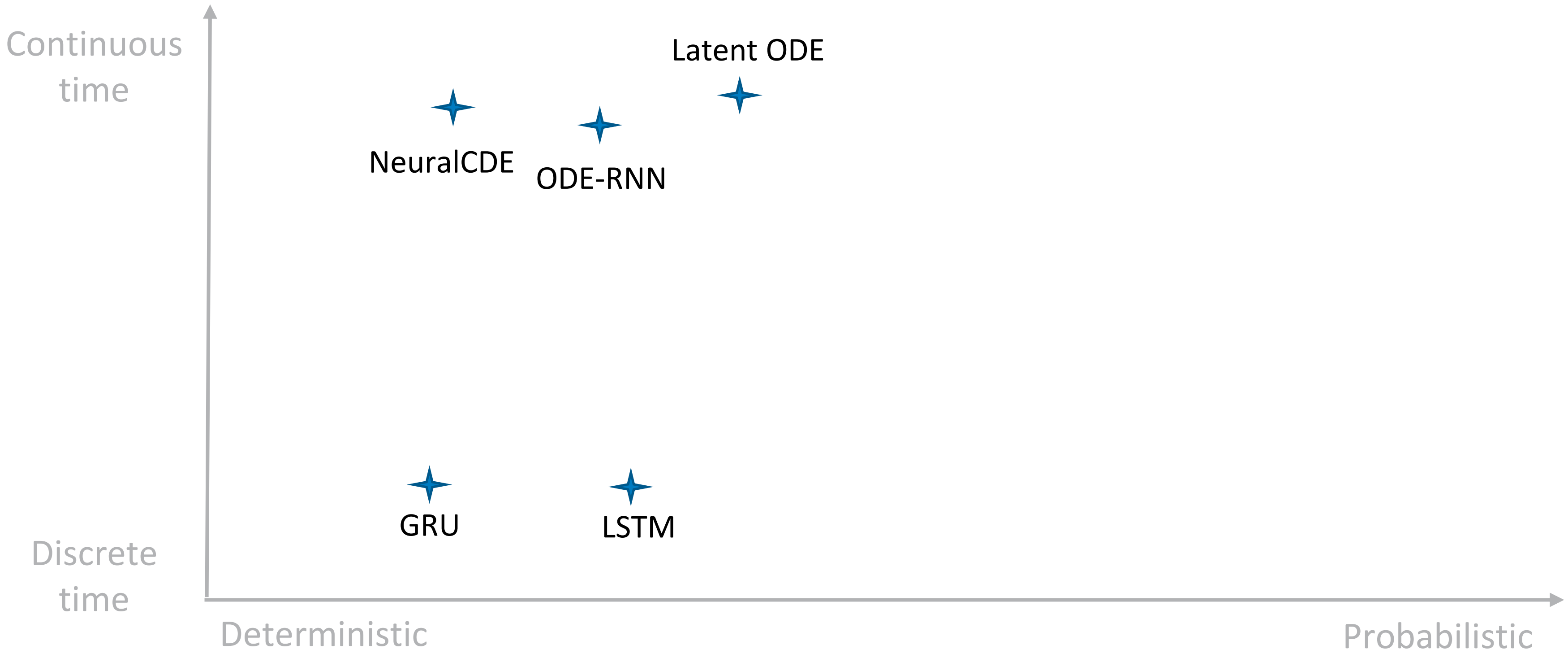
Modeling Irregularly-Sampled Time Series
Related Work



lack continuous dynamics

BOSCH

Modeling Irregularly-Sampled Time Series
Related Work

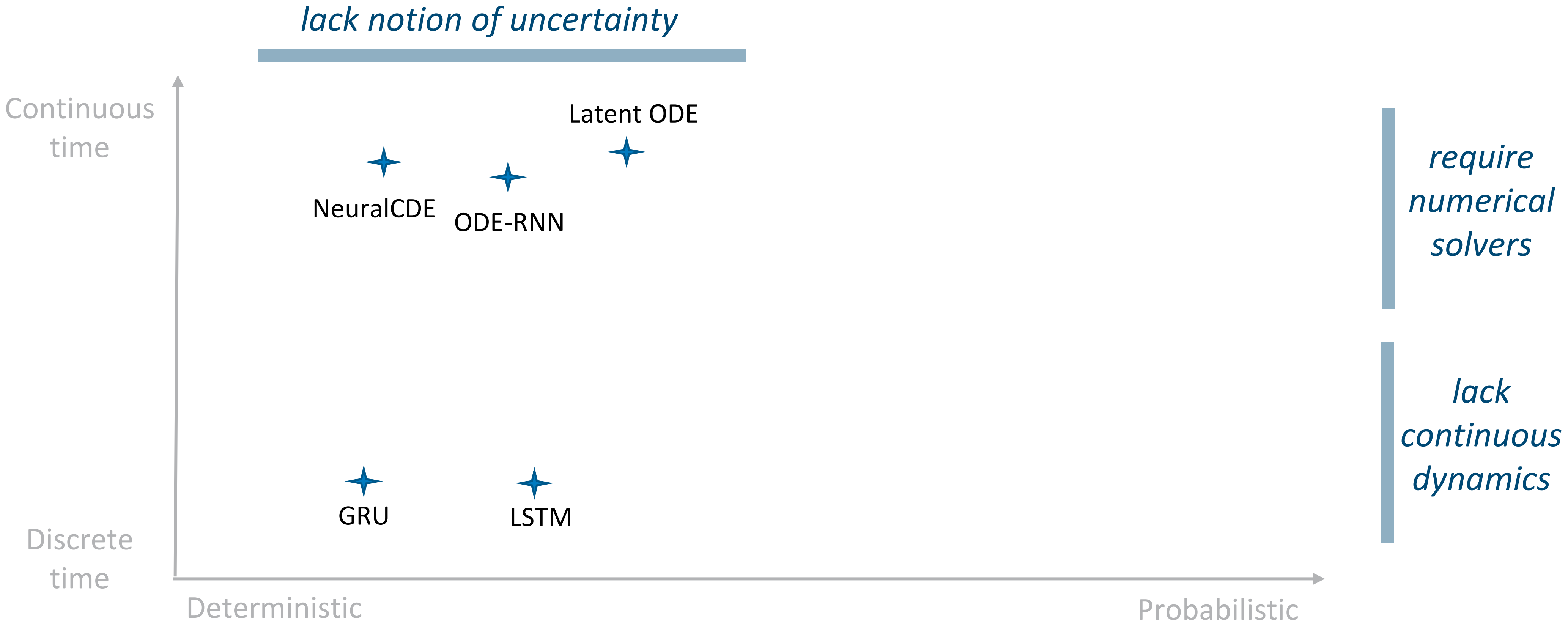


require numerical solvers

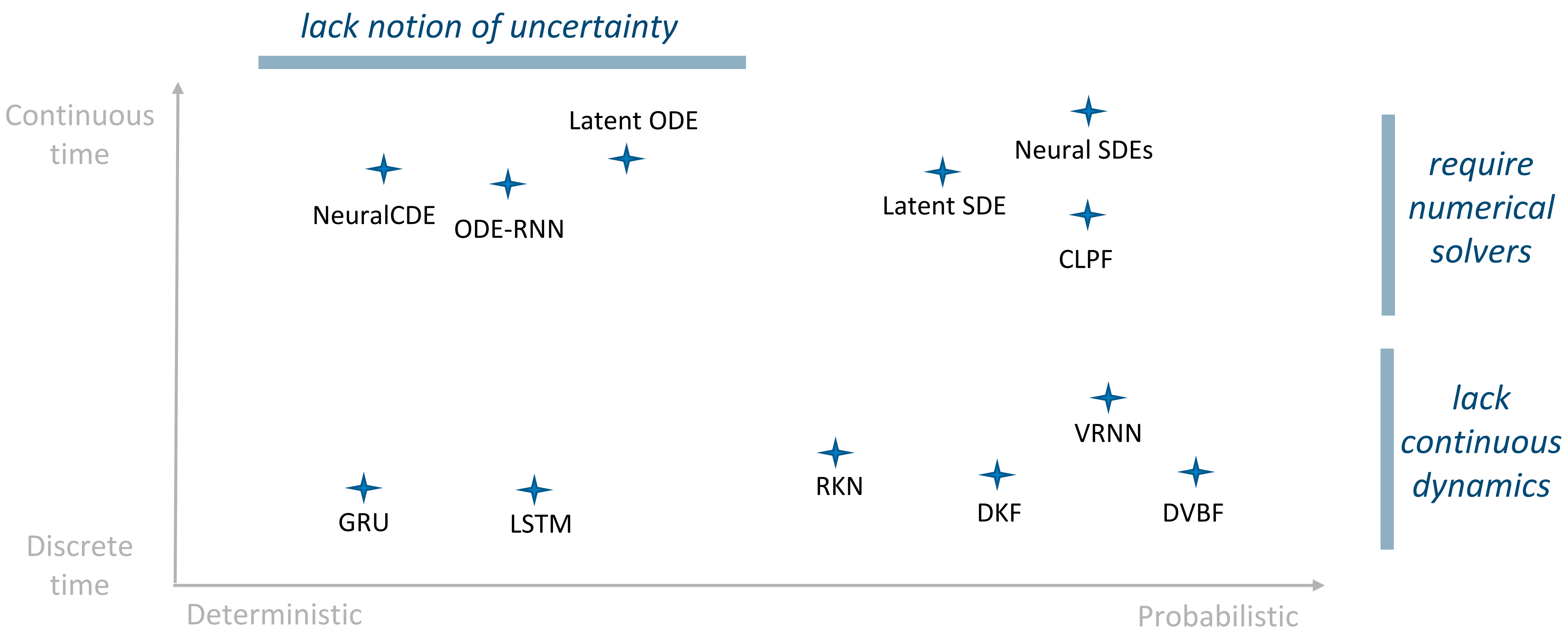
lack continuous dynamics

BOSCH

Modeling Irregularly-Sampled Time Series
Related Work

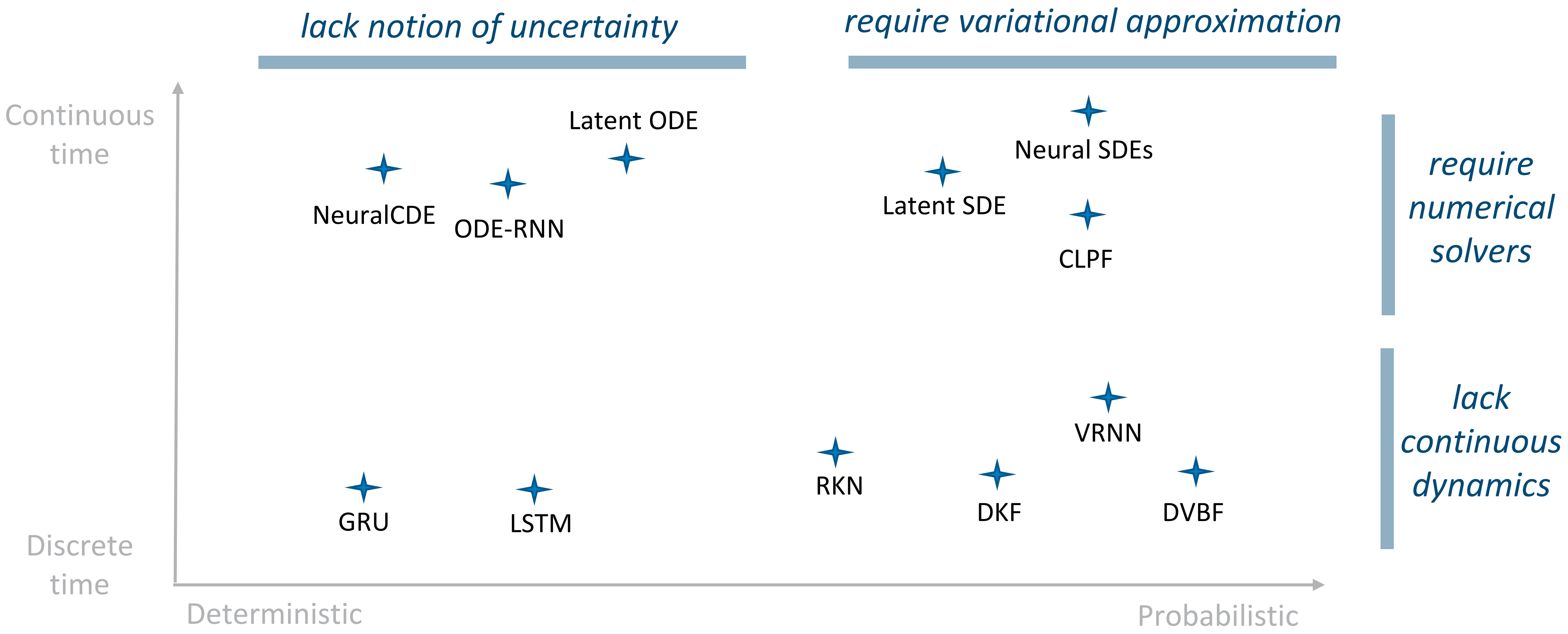


Modeling Irregularly-Sampled Time Series
Related Work

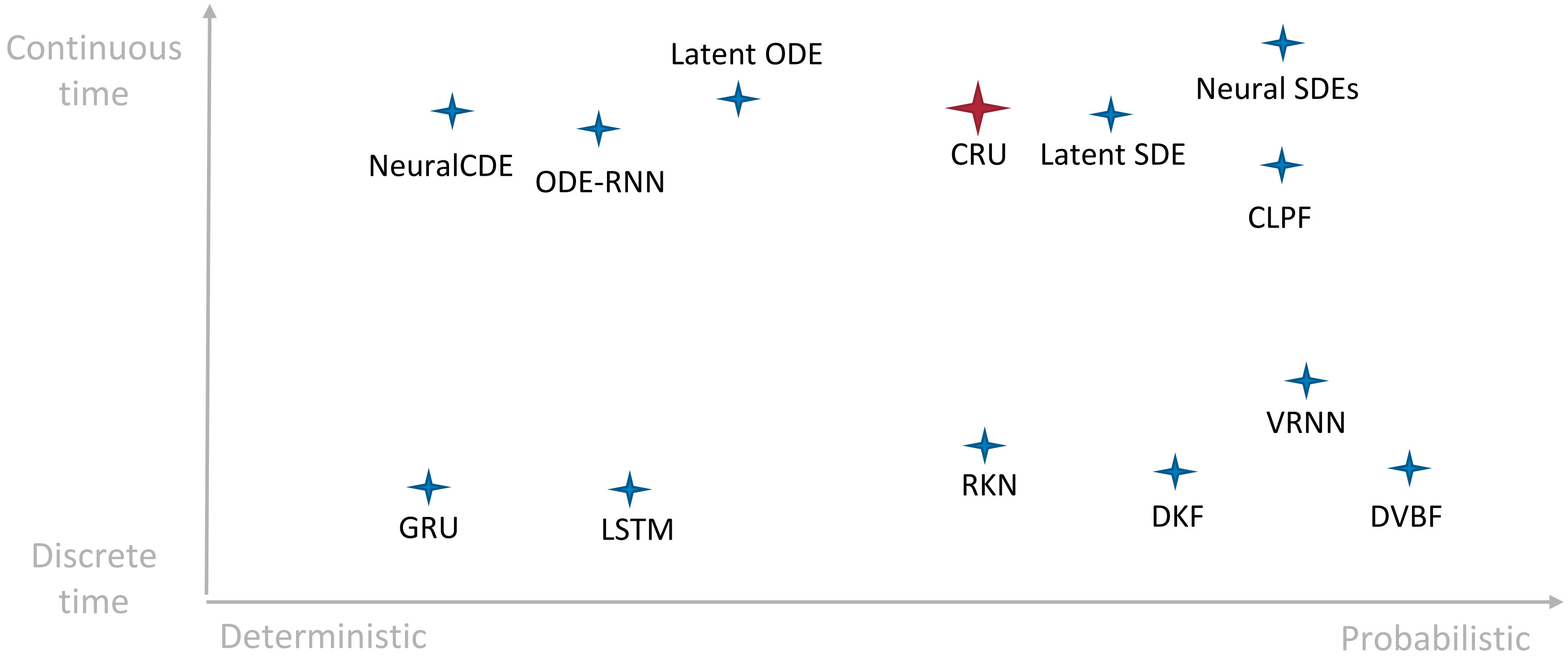


BOSCH

Modeling Irregularly-Sampled Time Series
Related Work



Modeling Irregularly-Sampled Time Series
Related Work



BOSCH

Continuous Recurrent Unit
Overview

Continuous Recurrent Unit (CRU) is a probabilistic recurrent network

Continuous Recurrent Unit
Overview

Continuous Recurrent Unit (CRU) is a probabilistic recurrent network with **continuous** state evolutions

Continuous Recurrent Unit
Overview

Continuous Recurrent Unit (CRU) is a probabilistic recurrent network with **continuous** state evolutions propagated in **closed form**

BOSCH

Continuous Recurrent Unit
Overview

Continuous Recurrent Unit (CRU) is a probabilistic recurrent network with **continuous** state evolutions propagated in **closed form**, an **uncertainty-driven gating mechanism**

Continuous Recurrent Unit

Overview

Continuous Recurrent Unit (CRU) is a probabilistic recurrent network with **continuous** state evolutions propagated in **closed form**, an **uncertainty-driven gating mechanism** and an architecture that is trainable **end-to-end.**

Continuous Recurrent Unit
Overview

Continuous Recurrent Unit (CRU) is a probabilistic recurrent network with **continuous** state evolutions propagated in **closed form**, an **uncertainty-driven gating mechanism** and an architecture that is trainable **end-to-end.**

CRU assumes a continuous latent state $\mathbf{z}$ that evolves according to a linear SDE

$$d\mathbf{z} = \mathbf{A}\mathbf{z}dt + \mathbf{G}d\boldsymbol{\beta}$$

with transition matrix $\mathbf{A}$ and diffusion coefficient $\mathbf{G}$

Continuous Recurrent Unit
Overview

Continuous Recurrent Unit (CRU) is a probabilistic recurrent network with **continuous** state evolutions propagated in **closed form**, an **uncertainty-driven gating mechanism** and an architecture that is trainable **end-to-end.**

CRU assumes a continuous latent state $\mathbf{z}$ that evolves according to a linear SDE

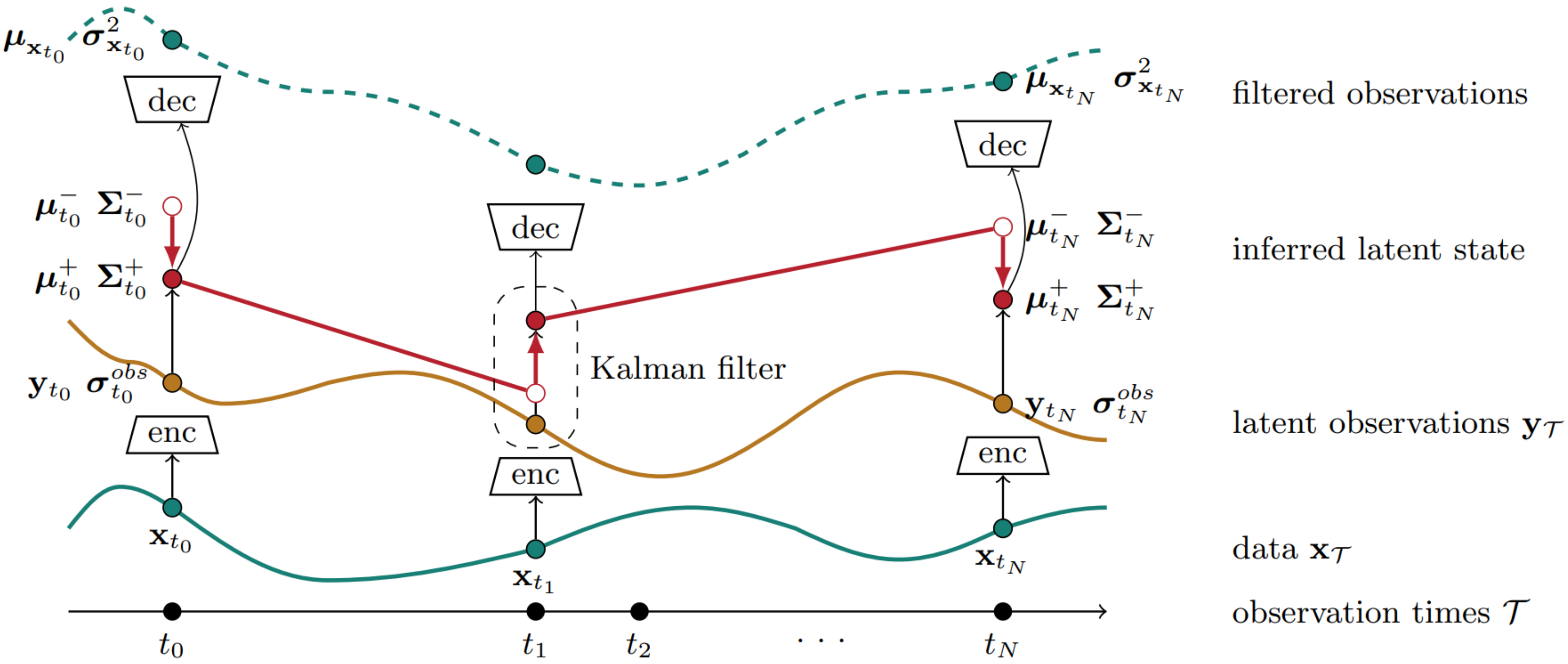
$$d\mathbf{z} = \mathbf{A}\mathbf{z}dt + \mathbf{G}d\boldsymbol{\beta}$$

with transition matrix $\mathbf{A}$ and diffusion coefficient $\mathbf{G}$ and discrete Gaussian latent observations

$$\mathbf{y}_t \sim \mathcal{N}(\mathbf{H}\mathbf{z}_t, (\boldsymbol{\sigma}_t^{\text{obs}})^2\mathbf{I})$$

BOSCH

Continuous Recurrent Unit
Architecture



Continuous Recurrent Unit

Modeling flexibility and complexity

- Increased flexibility with locally linear state transitions

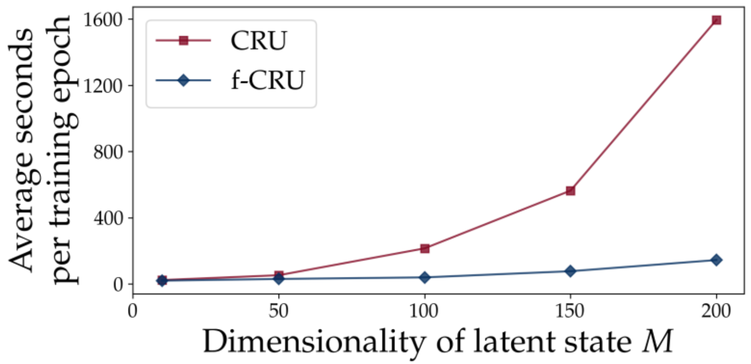
$$\mathbf{A}_t = \sum_{k=1}^{K} \alpha_t^{(k)} \mathbf{A}^{(k)}, \qquad \text{with } \boldsymbol{\alpha}_t = w_\psi(\boldsymbol{\mu}_t^+).$$

Continuous Recurrent Unit
Modeling flexibility and complexity

- Increased flexibility with locally linear state transitions

$$\mathbf{A}_t = \sum_{k=1}^{K} \alpha_t^{(k)} \mathbf{A}^{(k)}, \qquad \text{with } \boldsymbol{\alpha}_t = w_\psi(\boldsymbol{\mu}_t^+).$$

- Reduced complexity with a fast implementation f-CRU



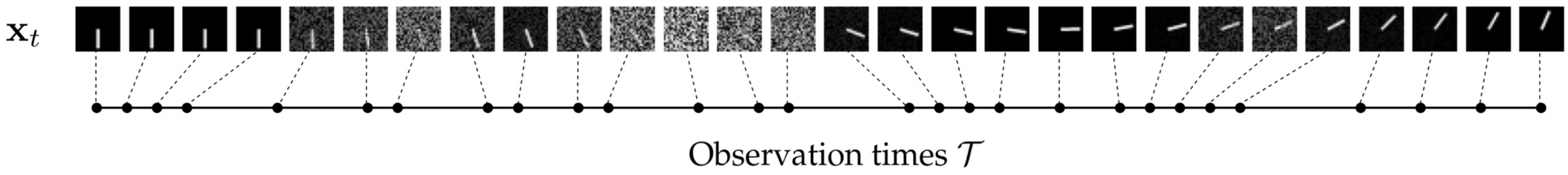
BOSCH

Continuous Recurrent Unit
Experimental study

- Best performance in 4 out of 6 comparisons
- Gating mechanism weights noisy and partially observed inputs accurately

Continuous Recurrent Unit
Experimental study

- Best performance in 4 out of 6 comparisons
- Gating mechanism weights noisy and partially observed inputs accurately

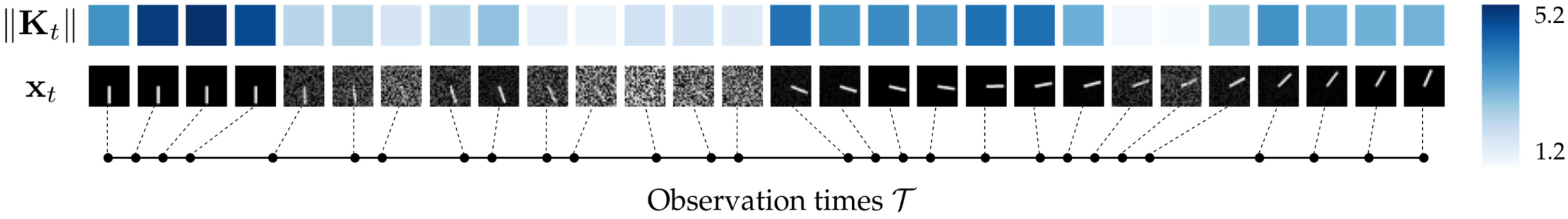


BOSCH

Continuous Recurrent Unit

Experimental study

- Best performance in 4 out of 6 comparisons
- Gating mechanism weights noisy and partially observed inputs accurately



BOSCH

THANK YOU

Paper

Correspondence:

mona.schirmer@ensae.fr

maja.rudolph@us.bosch.com

BOSCH