A Nonparametric Spatio-temporal SDE Model

Aalto University
School of Science

Cagatay Yildiz[1], Markus Heinonen[1 2], Harri Lähdesmäki[1]
[1]Aalto University, Finland [2]Helsinki Institute for Information Technology, Finland

Nonparametric SDEs [1]

- In [1] we considered SDEs with vector valued drift $\mathbf{f}(\mathbf{x}) : \mathbb{R}^D \to \mathbb{R}^D$ and scalar valued diffusion $\sigma(\mathbf{x}) : \mathbb{R}^D \to \mathbb{R}$ functions:

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t)dt + \sigma(\mathbf{x}_t)dW_t$$

- State solutions are given by the Itô integral[2]

$$\mathbf{x}_t = \mathbf{x}_0 + \int_0^t \mathbf{f}(\mathbf{x}_\tau)d\tau + \int_0^t \sigma(\mathbf{x}_\tau)dW_\tau,$$

- Assume $\mathbf{f}(\cdot)$ and $\sigma(\cdot)$ are completely unknown

- Gaussian process priors over drift and diffusion functions

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}(\mathbf{0}, K_{\mathbf{f}}(\mathbf{x},\mathbf{x}'))$$
$$\sigma(\mathbf{x}) \sim \mathcal{GP}(0, k_\sigma(\mathbf{x},\mathbf{x}'))$$

- Inducing points and variables

$$Z = (\mathbf{z}_1, \dots, \mathbf{z}_M)^T \in \mathbb{R}^{M\times D}$$
$$U_{\mathbf{f}} = (\mathbf{f}(\mathbf{z}_1), \dots, \mathbf{f}(\mathbf{z}_M))^T \in \mathbb{R}^{M\times D}$$
$$\mathbf{u}_\sigma = (\sigma(\mathbf{z}_1), \dots, \sigma(\mathbf{z}_M))^T \in \mathbb{R}^{M}$$

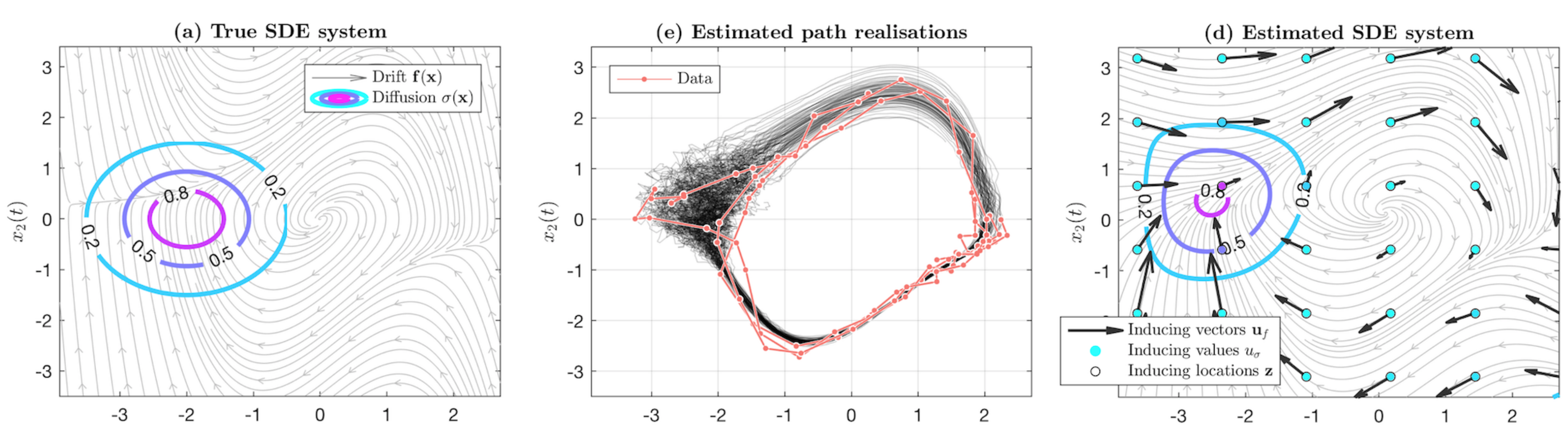
- Kernel interpolation

$$\mathbf{f}(\mathbf{x}) \triangleq \mathbf{K}_{\mathbf{f}}(\mathbf{x}, Z)\mathbf{K}_{\mathbf{f}}(Z,Z)^{-1}\mathrm{vec}\,U_{\mathbf{f}}$$
$$\sigma(\mathbf{x}) \triangleq K_\sigma(\mathbf{x}, Z)K_\sigma(Z,Z)^{-1}\mathbf{u}_\sigma$$

☺ **Nonparametric** estimation of **arbitrary** drift and diffusion functions from the data.

☺ **Full forward simulation** of the system.

☹ Not capable of modeling **time dependent** drift and diffusion functions.



Nonparametric Spatio-Temporal SDEs

Here, we augment the input space of the drift function to include a temporal component:

$$\mathbf{f}(\mathbf{x}, t) \; : \; (\mathbb{R}^D \times \mathbb{R}^+) \to \mathbb{R}^D$$
$$\mathbf{f}(\mathbf{x}, t) \sim \mathcal{GP}(\mathbf{0}, K_{\mathbf{f}}((\mathbf{x},t),(\mathbf{x}',t')))$$

Identity decomposable kernel $K_{\mathbf{f}}((\mathbf{x},t),(\mathbf{x}',t')) = k((\mathbf{x},t),(\mathbf{x}',t')) \cdot I_D$, where

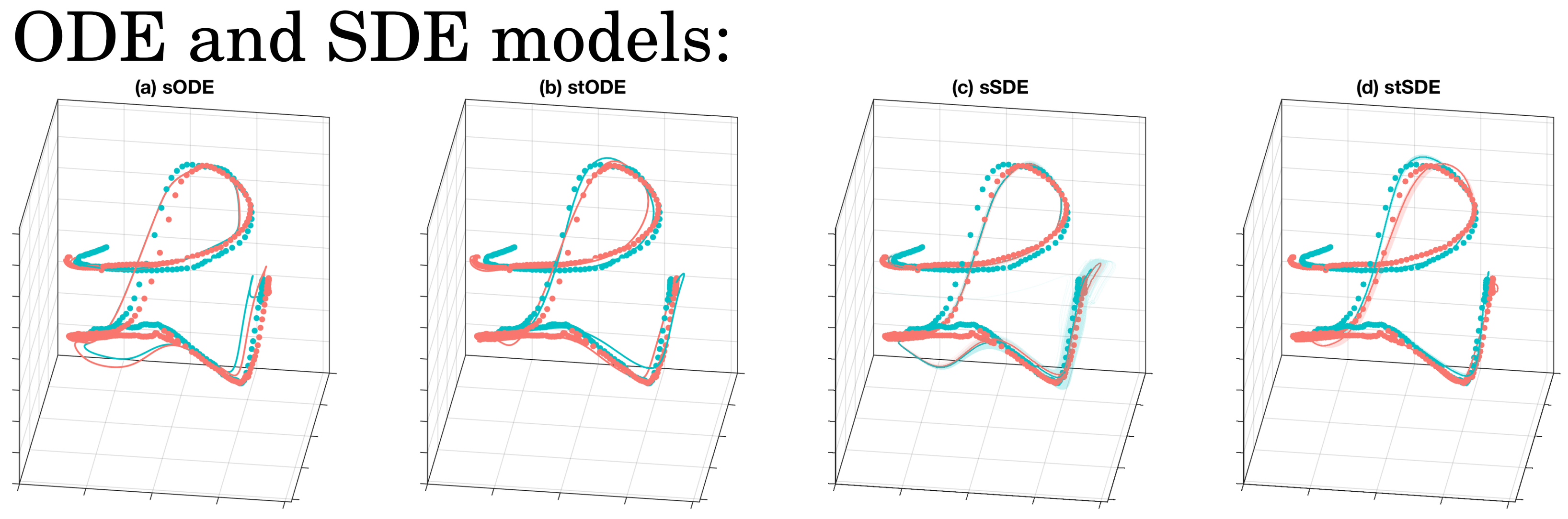
$$k((\mathbf{x},t),(\mathbf{x}',t')) = \exp\left(-\frac{1}{2}\sum_{d=1}^{D}\frac{(x_d - x_d')^2}{\ell_{\mathbf{f}d}^2} - \frac{(t-t')^2}{2\ell_t^2}\right)$$

MAP estimates of the unknowns using approximated posterior; end-to-end differentiable

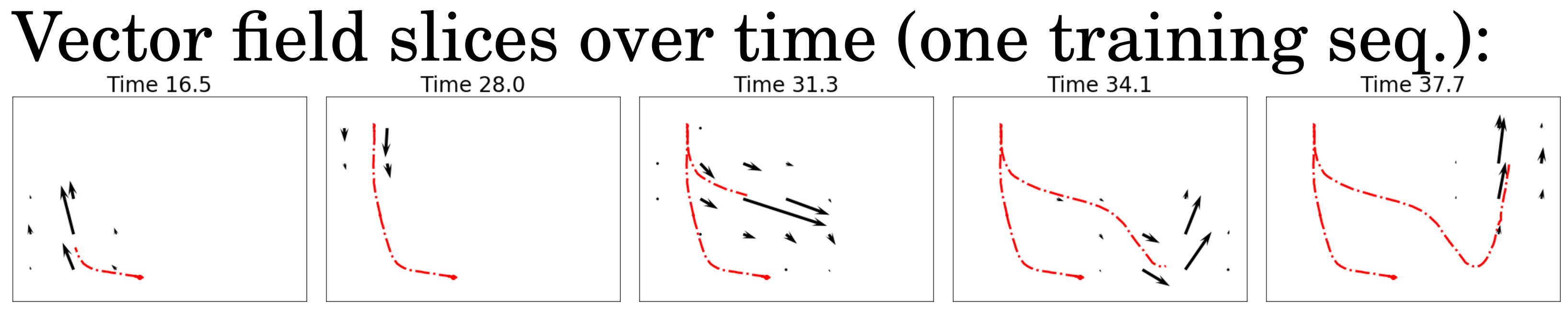
$$\underbrace{p(\mathbf{u}_{\mathbf{f}}, \mathbf{u}_\sigma)}_{\text{GP priors}} p(\Omega) \underbrace{\prod_{i=1}^{N} \mathbb{E}_{p(\mathbf{x}_i|t_i;\mathbf{f},\sigma)}[\mathcal{N}(\mathbf{y}_i|\mathbf{x}_i, \Omega)]}_{\text{likelihood}}$$

Results

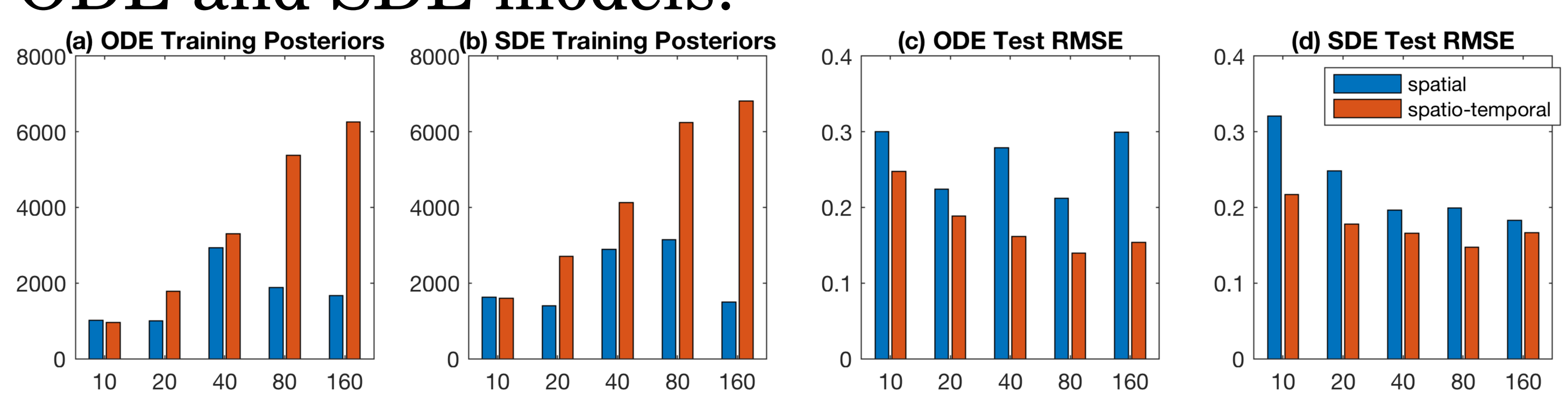
Paths estimated by spatial and spatio-temporal ODE and SDE models:



Vector field slices over time (one training seq.):



Training posteriors and prediction errors for ODE and SDE models:



☺ **Nonparametric** estimation of arbitrary **time-varying** SDE from data

☹ Model has high capacity: more data needed, inference is sensitive

- TO DO: variational inf. to avoid overfitting

References

[1] Yildiz, C., Heinonen, M., Intosalmi, J., Mannerstrom, H. and Lahdesmaki, H., "Learning Stochastic Differential Equations With Gaussian Processes Without Gradient Matching". *IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2018.

[2] B. Øksendal. "Stochastic Differential Equations: An Introduction with Applications". *Springer, 6th edition*, 2014.

Contact: Cagatay Yildiz - cagatay.yildiz@aalto.fi