Incorporating physical constraints in deep probabilistic models of coarse-grained dynamics

Introduction

Data-based discovery of effective, coarse-grained (CG) models of high-dimensional dynamical systems presents a unique challenge in the context of multiscale problems.

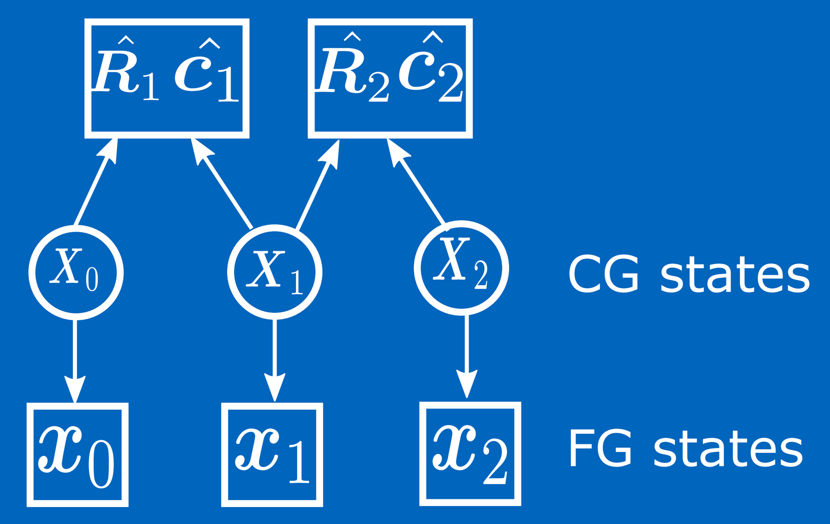
We present a generative, probabilistic Bayesian machine learning framework[1], which employs fine-grained (FG) simulation data in combination with virtual observables[2] to account for constraints. Our model simultanously identifies a probabilistic coarse-to-fine map as well as an evolution law for the CG dynamics. The former can be defined using a deep neural net to endow great expressiveness and flexibility[3].

By including physical constraints we are able to train our model in the Small Data regime and generate extrapolative predictions.

The Model

Training data for the model are the time dependent FG simulation data $\boldsymbol{x}$ together with residuals $\hat{\boldsymbol{R}}$ and constraints $\hat{\boldsymbol{c}}$.

During training we infer the CG coordinates $\boldsymbol{X}$ together with the parameters $\boldsymbol{\theta}$ for the coarse-to-fine map and the CG dynamics, whose form is parametrized.
The probabilistic graphical model is shown on the right.



Virtual Observables

Virtual Observables can be used to incorporate constraints or residuals into our model.
For instance for a residual $\boldsymbol{R}_l$ a new variable $\hat{\boldsymbol{R}}_l$ can be defined as:

$$\hat{R}_l = R_l(X) + \sigma_R \epsilon_R, \qquad \epsilon_R \sim \mathcal{N}(0, I)$$

We assume that $\hat{\boldsymbol{R}}_l$ has been virtually observed and $\hat{\boldsymbol{R}}_l = 0$ leads to an augmented version of our data.
This implies the virtual likelihood:

$$p(\hat{\boldsymbol{R}}_l = 0 | \boldsymbol{X}, \sigma_R) = \mathcal{N}(0 |\, \boldsymbol{R}_l(\boldsymbol{X}), \sigma_R^2 \boldsymbol{I})$$

The parameter $\boldsymbol{\sigma}_R$ determines the intensity of the enforcement.

Particle System

The training data $\mathcal{D}$ for our first example is generated by 500 000 identical particles moving in a bounded domain with stochastic dynamics corresponding to an inviscid Burgers-type-behaviour for the density.

As CG variables we use the particle density and the conservation of mass is incorporated with a virtual observable.

Images of oscillating nonlinear pendulum

The training data $\mathcal{D}$ for the second example is a series of images of a moving pendulum. Each image consists of 25x25 pixels taking binary values.

None of the only 16 training samples includes a full oscillation and the underlying CG system is chosen to be two-dimensional.

Model Training

- The training data $\mathcal{D}$ is augmented by virtual observables.
- We approximate the posterior by using Stochastic Variational Inference. The approximate posterior is decomposed as follows:

$$q_\phi(\boldsymbol{X}^{(1:n)}_{0:T\Delta t}, \ \boldsymbol{\theta}) = q_\phi(\boldsymbol{X}^{(1:n)}_{0:T\Delta t})\ q_\phi(\boldsymbol{\theta})$$

- We minimize the Kullback-Leibler divergence by maximing the Evidence-Lower-Bound $\mathcal{F}$:

$$\mathcal{F}(q_\phi(X^{(1:n)}_{0:T\Delta t}, \ \theta)) = \mathbb{E}_{q_\phi}\left[\log p(\mathcal{D} |\ X^{(1:n)}_{0:T\Delta t}, \ \theta)\right] + \mathbb{E}_{q_\phi}\left[\log \frac{p(X^{(1:n)}_{0:T\Delta t}, \ \theta)}{q_\phi(X^{(1:n)}_{0:T\Delta t}, \ \theta)}\right]$$

- Monte Carlo estimates in combination with stochastic gradient ascent are employed to find the necessary gradients.

Conclusion

We introduced a Bayesian machine learning framework for coarse-graining high-dimensional multiscale systems based on FG simulation data and a-priori available physical constraints.

Using only a small number of FG simulations as training data, we can learn a coarse-to-fine map as well as a coarse evolution law and can therefore produce probabilistic predictions based on the generated CG representation. We applied the framework to a Particle system and a series of images of a nonlinear pendulum and showed that the predictions are accurate under extrapolative conditions.

The proposed method is fully probabilistic and therefore leads to predictive distributions, i.e. it can quantify the inevitable uncertainty due to the information loss during coarse-graining.

Predictions

- We can analyze the learned CG representations and produce probabilistic predictions for the full FG system.
- Left: predictions for the Burgers' type particle are shown which even capture the location and propagation of a shock front.
- Right: results of the pendulum system are shown

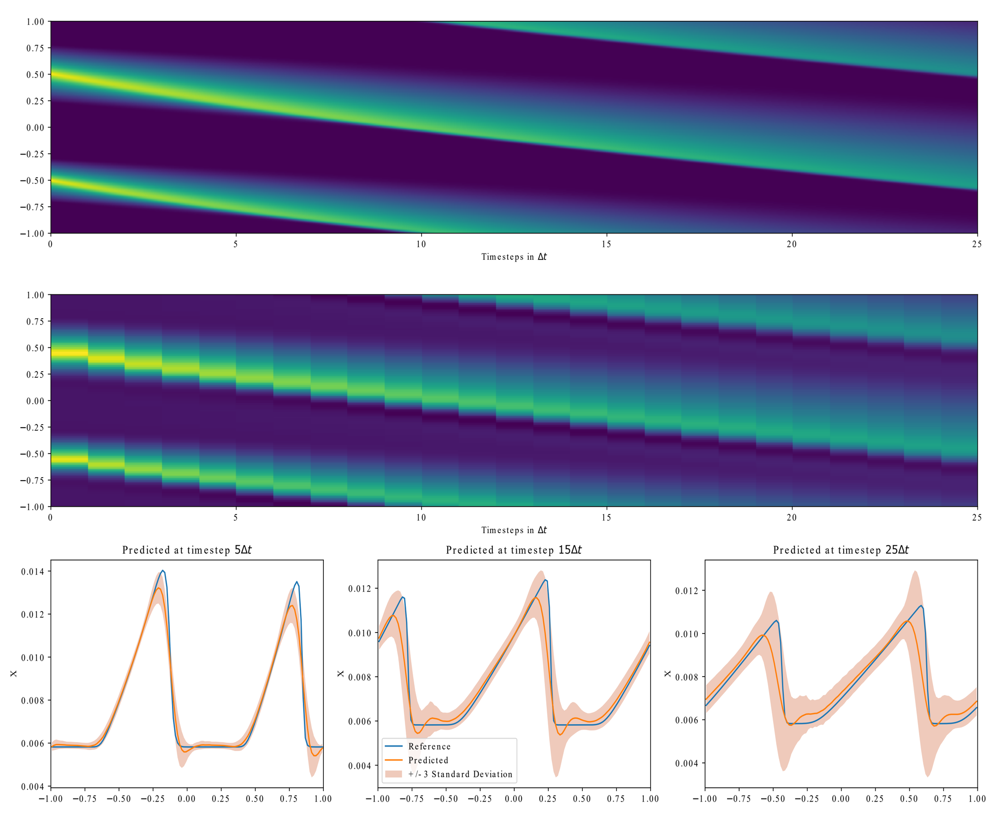
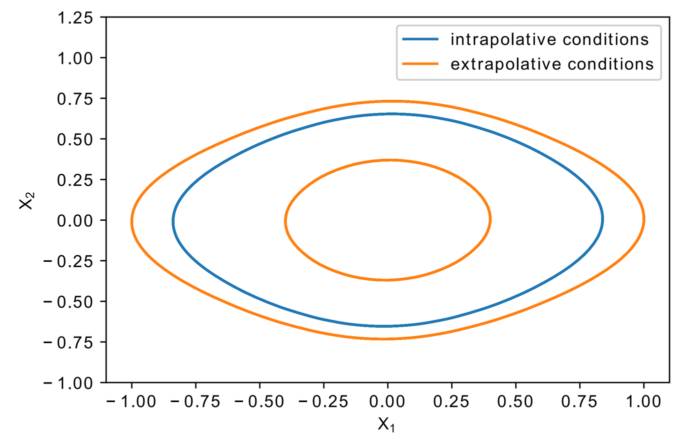
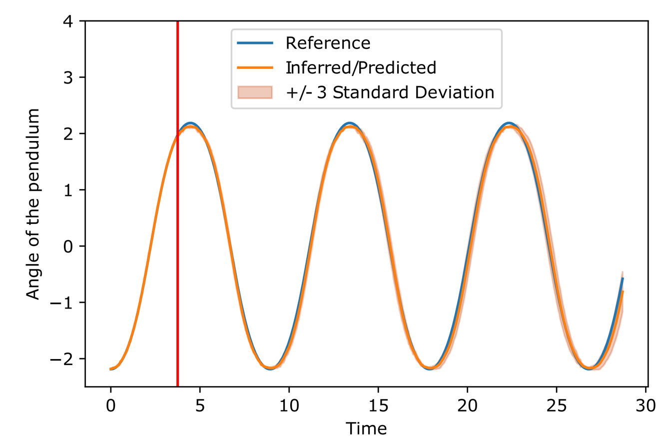


Figure 1: Top: Reference solution, Middle: Predictive posterior mean, Bottom: Snapshots at different time instances.





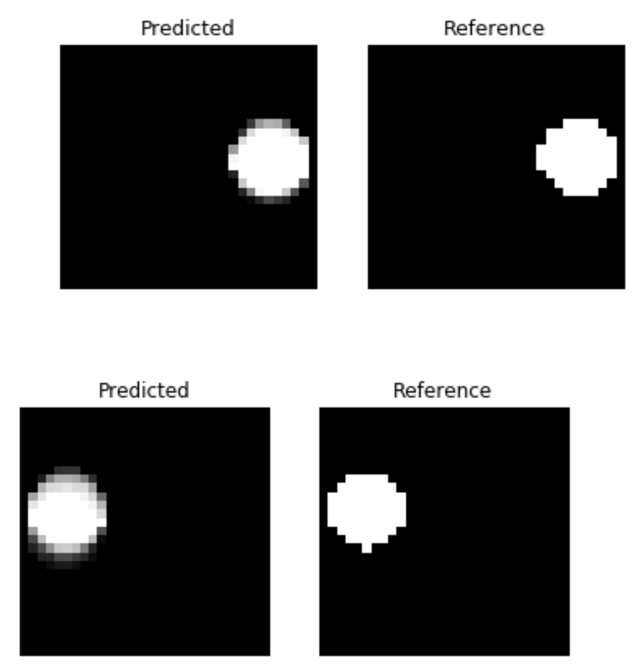


Figure 2,3,4,5: Left: Two-dimensional CG representation of the pendulum. Above: Probabilistic Predictions of the angle Right: Predictions and References for the full image

S. Kaltenbach*, P.-S. Koutsourelakis

Professorship of Continuum Mechanics,
Department of Mechanical Engineering,
Technical University of Munich
*sebastian.kaltenbach@tum.de

[1] P.-S. Koutsourelakis, E. Bilionis: Scalable Bayesian Reduced-Order Models for Simulating High-Dimensional Multiscale Dynamical Systems, Multiscale Modeling & Simulation 9 (2011)

[2] S.Kaltenbach, P.-S. Koutsourelakis: Incorporating physical constraints in a deep probabilistic machine learning framework for coarse-graining dynamical systems, arxiv preprint: 1912.12976

[3] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational Physics 378 (2019)