REAL-TIME ALGORITHMS FOR DIGITAL TWINS

Outlook

UKACM Autumn School 2026
Prof. Dirk Hartmann

TECHNISCHE UNIVERSITÄT DARMSTADT

1 “Advanced” Machine Learning for Dynamical Systems

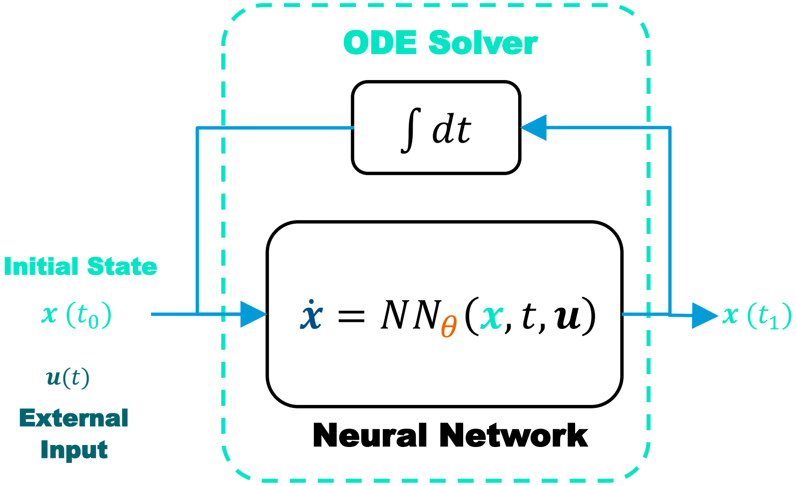
2 Nonlinear Dimension Reduction

3 Uncertainty Quantification

Machine Learning for Dynamical Systems
(Hybrid) Architectures

NeuralODE

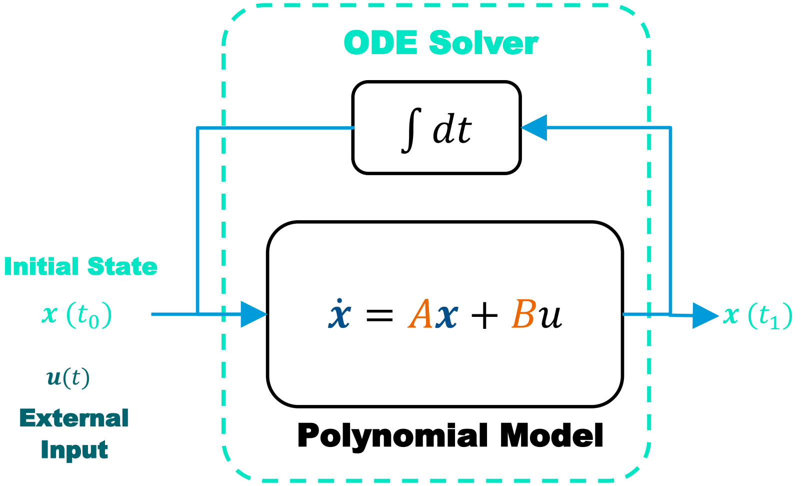
$$\frac{d}{dt}\boldsymbol{x} = NN_\theta\left(\boldsymbol{x}, t, \boldsymbol{u}\right)$$



Differential Operator Inference

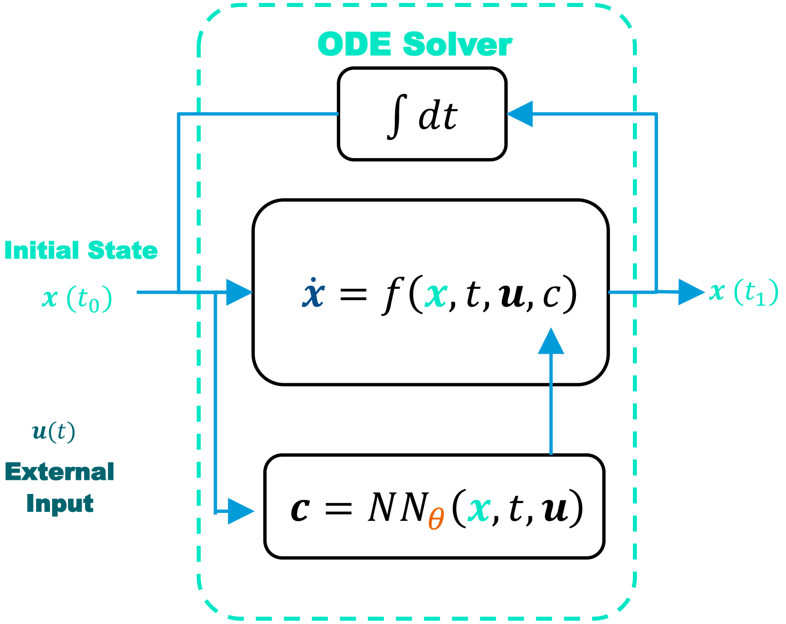
$$\frac{d}{dt}\boldsymbol{x} = Ax + Bu$$

Linear Operators for Simplicity



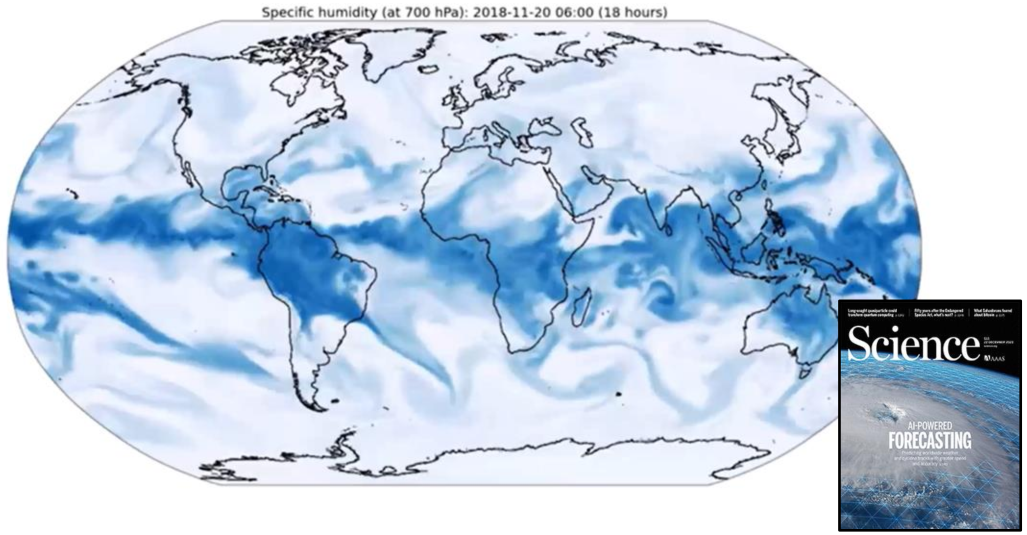
ML-augmented solver

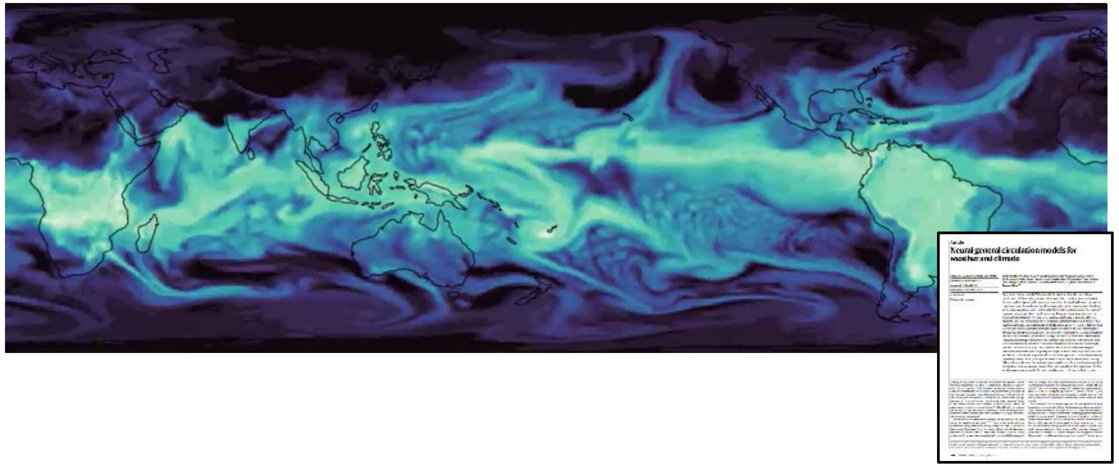
$$\frac{d}{dt}\boldsymbol{x} = f(x, u, NN_\theta(x, u))$$



Machine Learning for Dynamical Systems
(Hybrid) Architectures – Google GraphCast vs Neural GCM

TECHNISCHE UNIVERSITÄT DARMSTADT





GraphCast ([Dec 2022, DeepMind])
- ▶ Graph-based Surrogate Model
- ▶ Autoregressive Training
- ▶ Accurate 10-day Forecast

NeuralGCM ([Nov 2023, Google Research])
- ▶ Hybrid Model Architecture
- ▶ Autoregressive Training
- ▶ Outperforms state-of-the-art forecast accuracy between 5 and 15 days

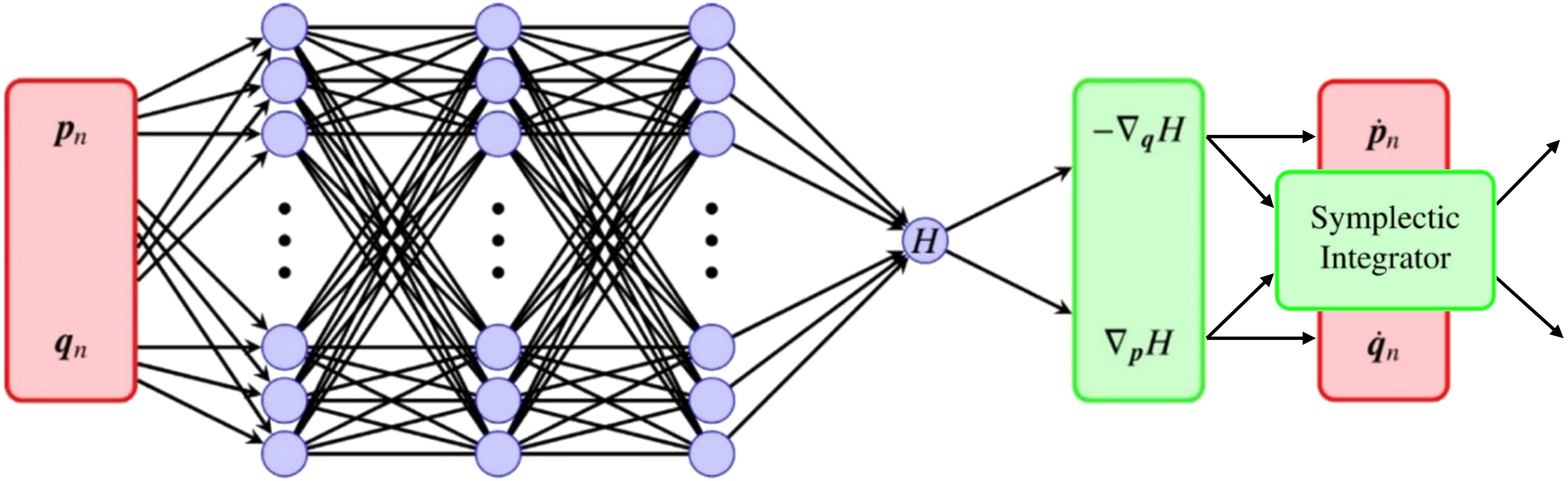
Source: R Lam, A Sanchez-Gonzalez, M Willson, P Wirnsberger, M Fortunato, F Alet, ... , P Battaglia (2023):.Learning skillful medium-range global weather forecasting. Science, 382(6677), 1416-1421.
D Kochkov, J Yuval, I Langmore, P Norgaard, J Smith, G Mooers ... & S Hoyer (2024): Neural general circulation models for weather and climate. *Nature*, *632*(8027), 1060-1066.

Machine Learning for Dynamical Systems

Model Biases – Hamiltonian Neural Networks

Idea: [Hamiltonian Networks]:

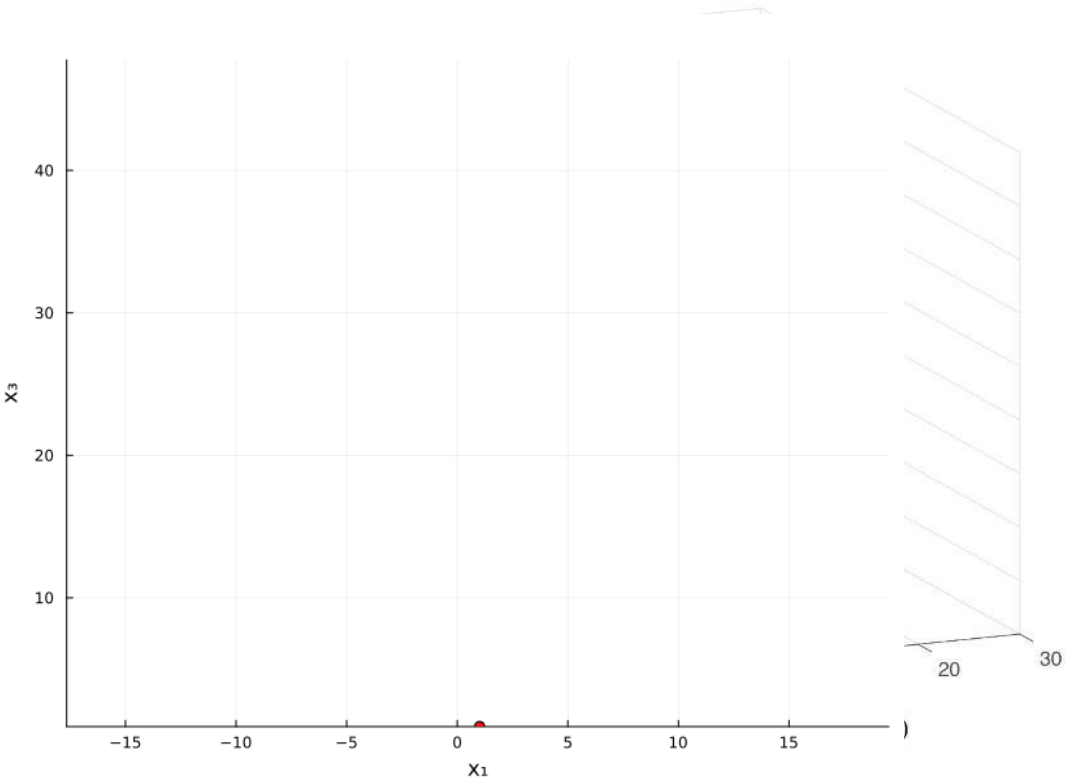
Under the assumption of the existence of an appropriate Hamoltonian H, i.e., $\frac{d}{dt}\binom{p}{q} = J \cdot \nabla H(p,q)$, we parameterize H with a Neural network H_{NN} and train the network to minimize $\left\| \frac{d}{dt}\binom{p}{q} = J \cdot \nabla H_{NN}(p,q) \right\|$. After training symplectic Integration ensures energy conservation.



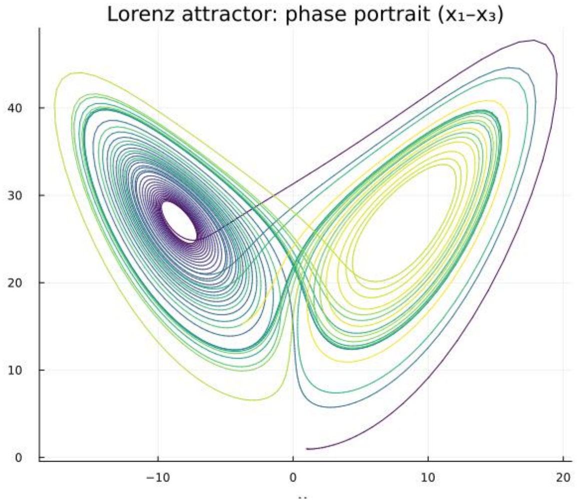
Source: P Horn, VS Ulibarrena, B Koren, SP Zwart (2025): A generalized framework of neural networks for Hamiltonian systems. *Journal of Computational Physics*, *521*, 113536.

Machine Learning for Dynamical Systems
Observability

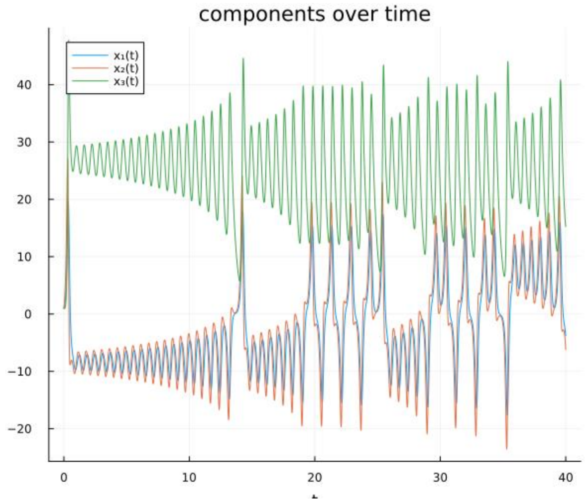
Lorenz Attractor



Partial Measurements

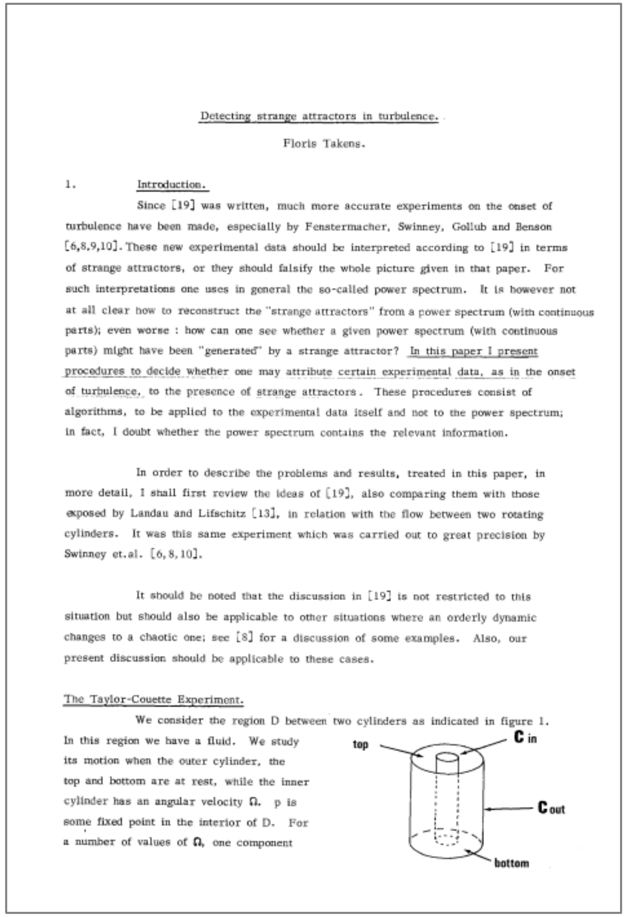


2D observations



1D observation

- ▶ If we do not spend sufficient time to understand the system in all details, we might miss out details when we identify the underlying system by data-based means



Detecting strange attractors in turbulence.

Floris Takens.

1. Introduction.

Since [19] was written, much more accurate experiments on the onset of turbulence have been made, especially by Fenstermacher, Swinney, Gollub and Benson [6,8,9,10]. These new experimental data should be interpreted according to [19] in terms of strange attractors, or they should falsify the whole picture given in that paper. For such interpretations one uses in general the so-called power spectrum. It is however not at all clear how to reconstruct the "strange attractors" from a power spectrum (with continuous parts); even worse : how can one see whether a given power spectrum (with continuous parts) might have been "generated" by a strange attractor? In this paper I present procedures to decide whether one may attribute certain experimental data, as in the onset of turbulence, to the presence of strange attractors. These procedures consist of algorithms, to be applied to the experimental data itself and not to the power spectrum; in fact, I doubt whether the power spectrum contains the relevant information.

In order to describe the problems and results, treated in this paper, in more detail, I shall first review the ideas of [19], also comparing them with those exposed by Landau and Lifschitz [13], in relation with the flow between two rotating cylinders. It was this same experiment which was carried out to great precision by Swinney et.al. [6, 8, 10].

It should be noted that the discussion in [19] is not restricted to this situation but should also be applicable to other situations where an orderly dynamic changes to a chaotic one; see [8] for a discussion of some examples. Also, our present discussion should be applicable to these cases.

The Taylor-Couette Experiment.

We consider the region D between two cylinders as indicated in figure 1. In this region we have a fluid. We study its motion when the outer cylinder, the top and bottom are at rest, while the inner cylinder has an angular velocity Ω. p is some fixed point in the interior of D. For a number of values of Ω, one component

What happens if we can not obtain the underlying true state of the system via observations? That is the evolution does not only depend on the currently observed state and input (Non-Markocian Dynamics)

Takens' (1981): We can get a "hidden" state that is predictive for the observations by using delays of the observation, i.e., create new coordinates by lagging that same data (Delay Embedding)

For the reconstruction to be mathematically "faithful, we require

- **Deterministic System:** The system must follow clear rules rather than being purely random.
- **Generic Observable:** The measurement must be "typical" and not a specialized value missing key dynamics.
- **Embedding Dimension m:** the number of coordinates m must $m > 2d$ where d is the dimension of the original system

Source: F Takens (2006): Detecting strange attractors in turbulence. In *Dynamical Systems and Turbulence, Warwick 1980: proceedings of a symposium held at the University of Warwick 1979/80* (pp. 366-381). Berlin, Heidelberg: Springer Berlin Heidelberg.

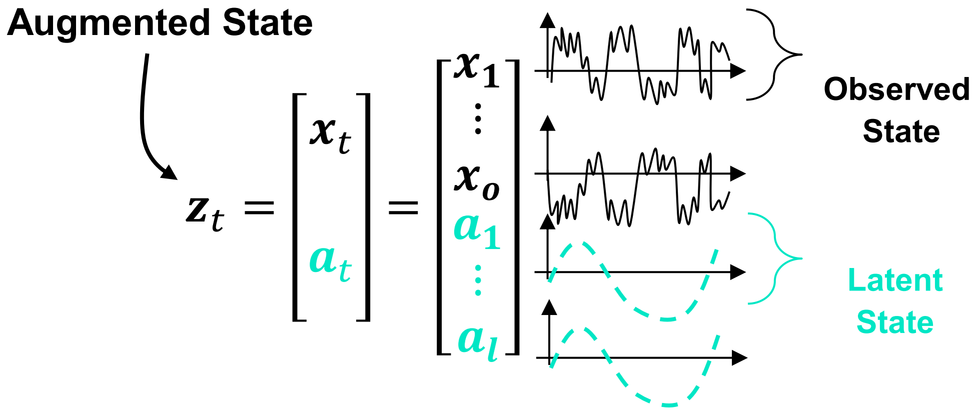
Machine Learning for Dynamical Systems

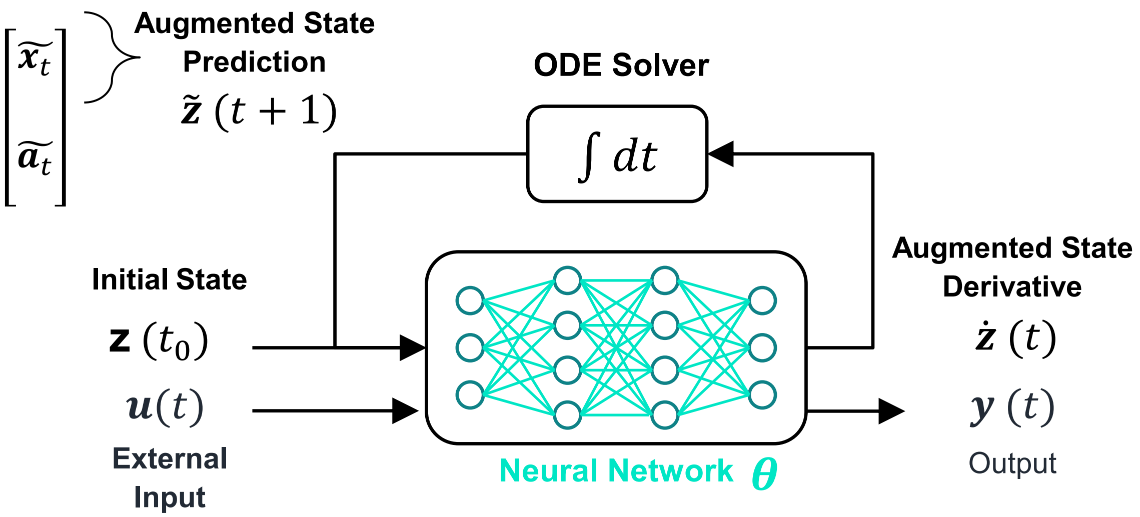
Observability – Takens' Theorem (ii)

- In a typical realization, delay embeddings are directly added to the "observed" state, i.e. the observed state does not only depend on the observed variable y(t) and the input u(t) but also on past time steps variables.

$$\tilde{y}(t) \coloneqq \left(y(t), y\left(t - \Delta t^{sampling}\right), \dots, y\left(t - \mathrm{q}\, \Delta t^{sampling}\right)\right)$$

Idea [Neural Embedding for Dynamical Systems]: Encoding of time delays by additional latent states (similar to recurrent neural networks)





Source: S Ouala, D Nguyen, L Drumetz, B Chapron, A Pascual, F Collard,... R Fablet (2020): Learning latent dynamics for partially observed chaotic systems. Chaos: An Interdisciplinary Journal of Nonlinear Science, 30(10).

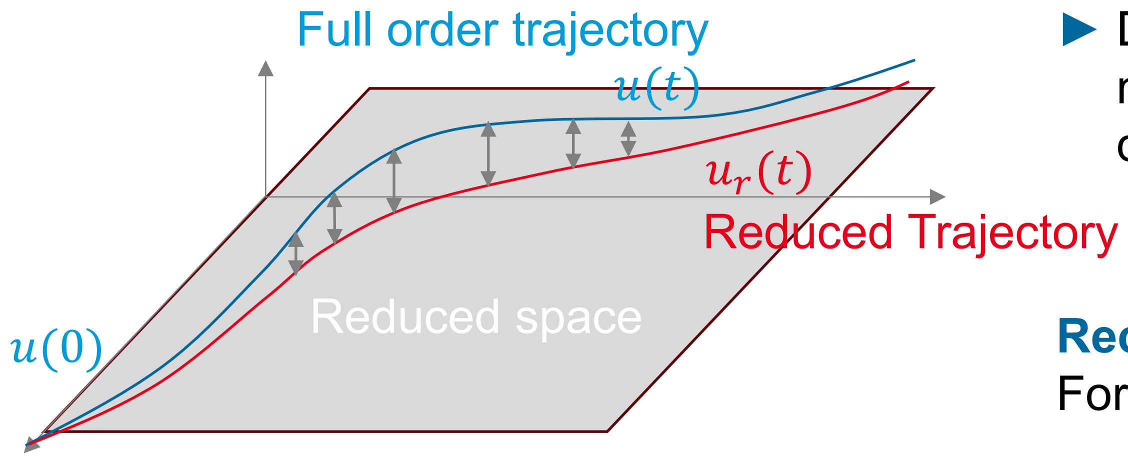
TECHNISCHE UNIVERSITÄT DARMSTADT

1 “Advanced” Machine Learning for Dynamical Systems

2 **Nonlinear Dimension Reduction**

3 Uncertainty Quantification

Non-linear Dimension Reduction



- Despite potentially “high” dimensional observations, many dynamical systems are (effectively) “low” dimensional

Recap [Projection based methods]:
For a generic dynamical system:

$$\frac{d}{dt}s(t) = W^T f\big(Ws(t), u(t)\big) t \in (0, T)$$

$$s(0) = W^T y_0$$

with reduced variables $s = W^T y$ and state variables $y = Ws$.

- However, many systems are not suitable to be compactly described through a linear projection

TECHNISCHE UNIVERSITÄT DARMSTADT

- The Kolmogorov n-width is a useful metric to quantify the suitability of a linear projection.

Definition[The Kolmogorov n-width]:

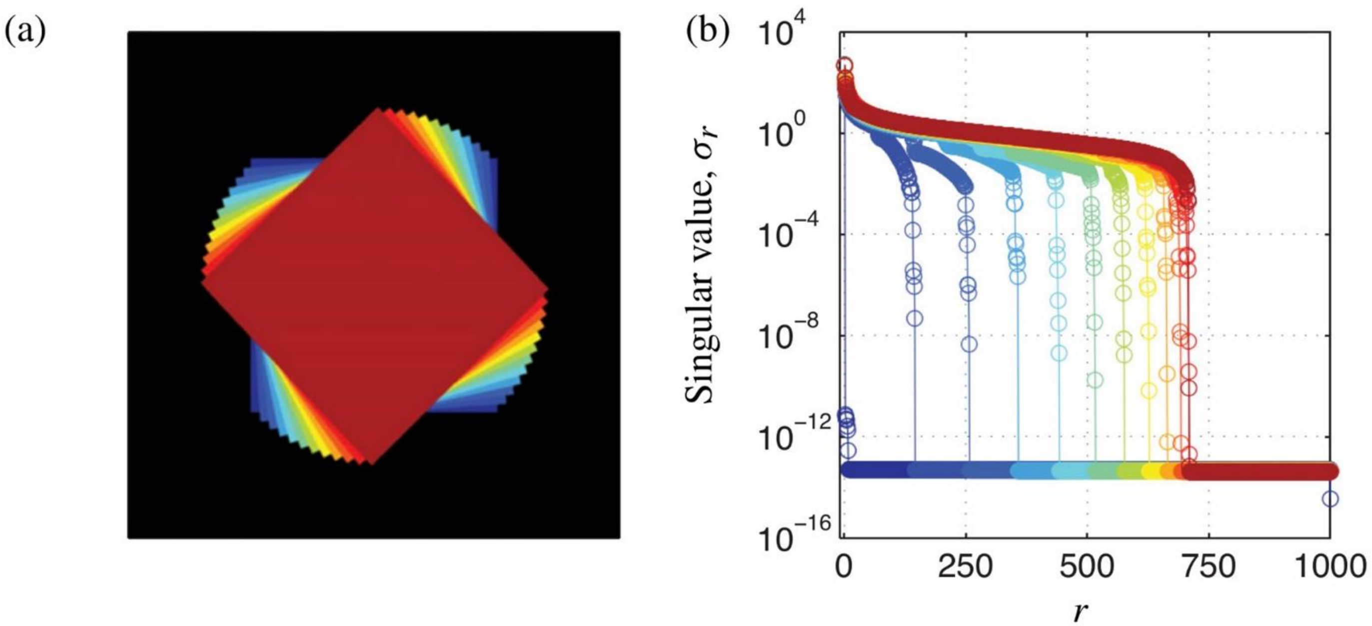
$$d_{n(M)} = \inf_{W \in \mathbb{R}^{N_y \times n}} \sup_{y \in M} \inf_{s \in \mathbb{R}^n} \|y - Ws\| \,,$$

where $M \subset \mathbb{R}^{N_y}$ is the solution manifold associated with the dynamical system. That is, the Kolmogorow n-width measures the best approximation error that can be obtained by approximating the solution manifold with a linear subspace of dimension n.

- If the Kolmogorov n -width decreases slowly with n, it indicates that very high dimensional spaces are needed, e.g., a POD with a high-dimension is needed.

Non-linear Dimension Reduction
Limitation of the SVD / POD (i)

- Consider an Image which is rotated, i.e., a black and white picture of a square,
- If it is perfectly aligned it has just one non-zero none singular value and the corresponding singular values u_1 and v_1 are the width and height of the rectangle



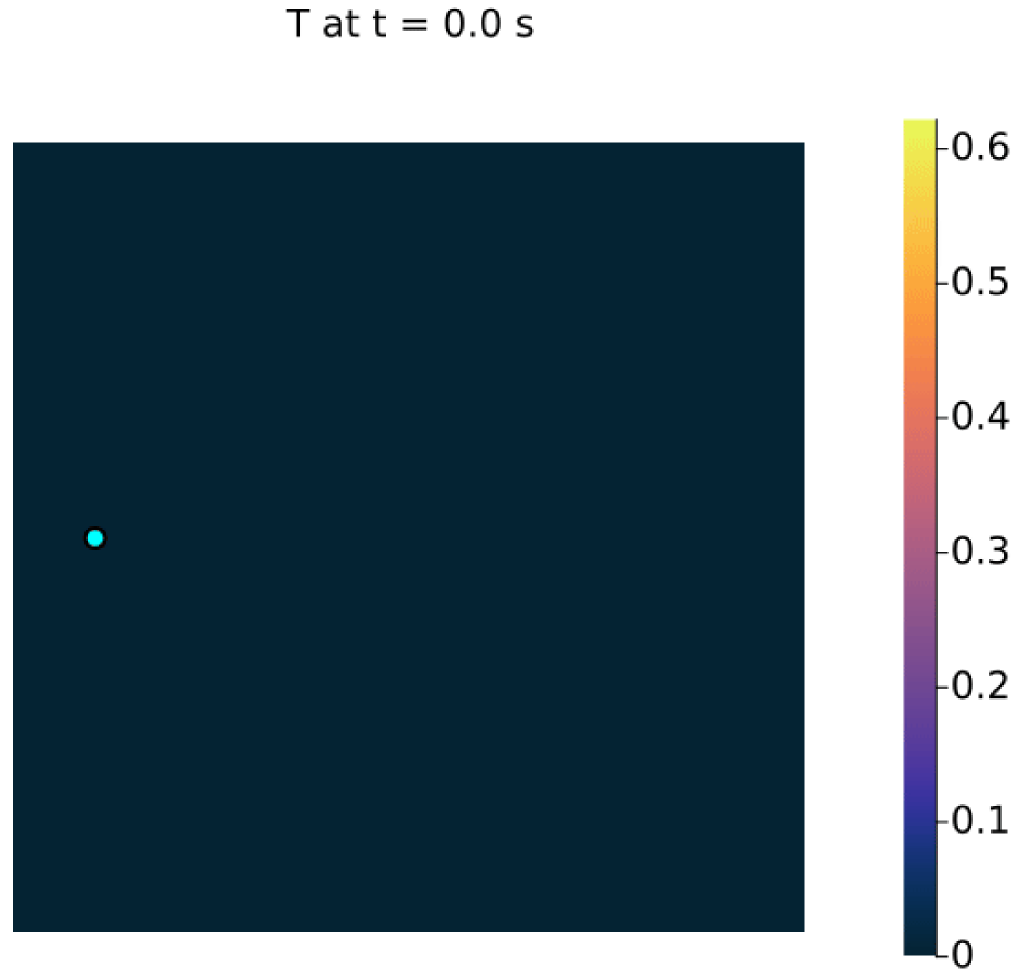
A data matrix consisting of zeros with a square sub-block of ones at various rotations (a), and the corresponding SVD spectrum, $\mathrm{diag}(S)$, (b).

- The SVD breaks down because it depends on the coordinate system the data is presented in
- The SVD rank explode when objects in the columns translate, rotate or scale, which severely limits it's use for data that has non been pre-processed appropriately

Source: SL Brunton, JN Kutz (2019): Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control, Cambridge University Press

Non-linear Dimension Reduction

Limitation of the SVD / POD (ii)



“Selective Laster Melting”

$$\partial_t T = \Delta T + q(x,t)$$

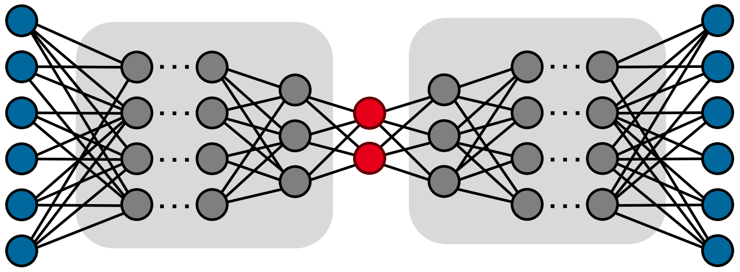
- Linear PDE
- Heating via a moving heat source
- The nightmare for ROM

https://github.com/Dirk-Hartmann/real-time-digital-twins/blob/main/scripts/pde/SLM_transient.jl

Non-linear Dimension Reduction
Autoencoder (i)

Idea [Autoencoders]: Following the idea of the success of the POD use autoencoders to perform the dimension reduction[1].

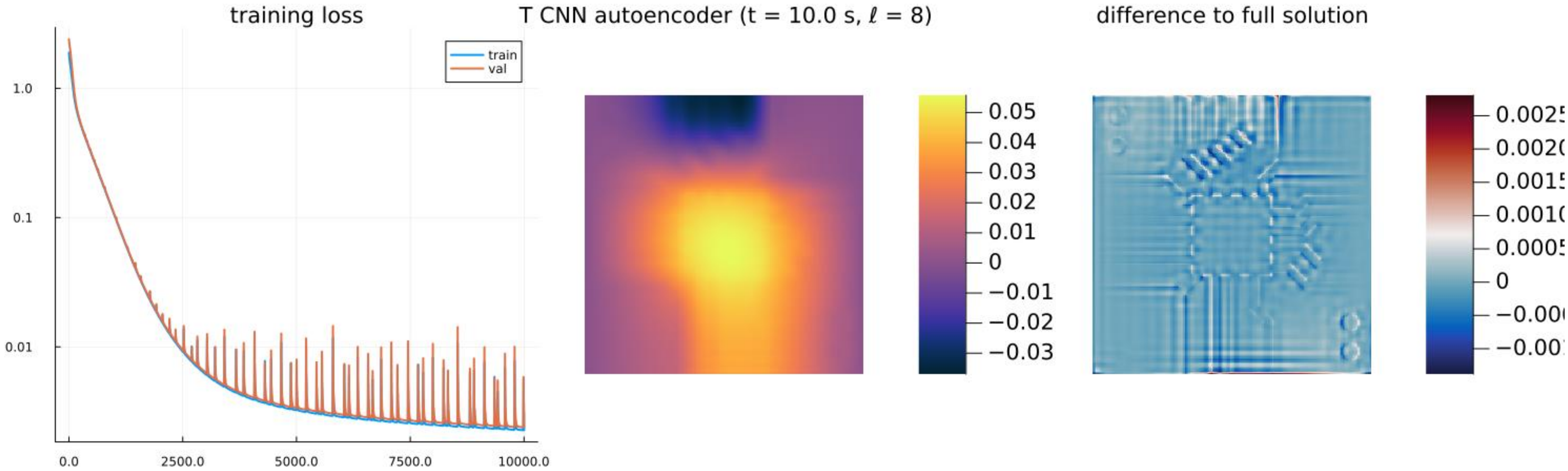
- First, an appropriate low dimensional space is identified by means of an autoencoder, i.e.,
  $$w_{enc}^*, w_{dnc}^* = \underset{w_{enc}^*, w_{dec}^*}{\operatorname{argmin}} \sum\sum \left\| \hat{y}_i^j - \mathcal{N}\mathcal{N}_{dec}\left(\mathcal{N}\mathcal{N}_{en}\left(\hat{y}_i^j; w_{enc}\right); w_{dec}\right) \right\|^2 .$$
  The compact representation s with $s = \mathcal{N}\mathcal{N}_{enc}(y)$ and $y = \mathcal{N}\mathcal{N}_{dec}(s)$ is often called the latent variable[2]



- Once the autoencoder is trained, dynamics can be identified in the tangent space of the non-linear manifold,
  $$\frac{d}{dt} s(t) = \mathcal{N}\mathcal{N}_{dyn}\left(s(t), u(t); w_{dyn}\right) t \in (0, T)$$
  $$s(0) = \mathcal{N}\mathcal{N}_{enc}(y_0)$$

1) MA Kramer (1991): "Nonlinear principal component analysis using autoassociative neural networks." AIChE journal 37.2 (1991): 233-243.
2) A variable that is not directly measured but is inferred by means of a mathematical model from one or more directly observed variables (Wikipedia)

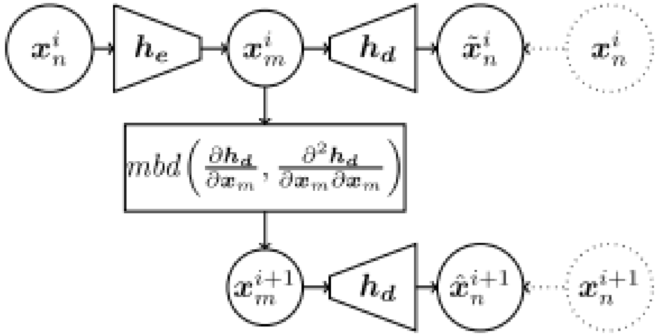
Non-linear Dimension Reduction
Autoencoder (i)



https://github.com/Dirk-Hartmann/real-time-digital-twins/blob/main/scripts/rom/PCB_AutoencoderCNN.jl

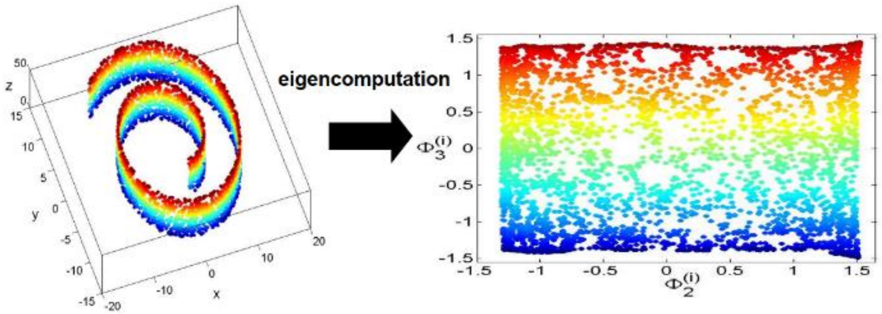
Non-linear Dimension Reduction
Further Architectures and Approaches

- Typically, the autoencoder and latent space dynamical system are trained separately but a simultaneous end to end training can be beneficial since the latent space is not fixed and takes details of the dynamics into account.



- This concept is also useful, for linear projections

- Autoencoders are not the only options, many non-linear dimension identification methods have been proposed in the last decades, e.g., Diffusion Maps or Polynomial Networks



- The same framework can also be used to learn Koopman Operators, i.e., explicitly enforcing linear models.

Source: 1) A Angeli, W Desmet F Naets (2021): Deep learning for model order reduction of multibody systems to minimal coordinates. Computer Methods in Applied Mechanics and Engineering, 373, 113517
2) RR Coifman, IG Kevrekidis, S Lafon, M Maggioni, B Nadler (2008): Diffusion maps, reduction coordinates, and low dimensional representation of stochastic systems. Multiscale Modeling & Simulation, 7(2)
3) A Bensalah, A Nouy, J Soffo (2025): Nonlinear manifold approximation using compositional polynomial networks. arXiv preprint arXiv:2502.05088.
4) JM Hokanson, PG Constantine (2018): Data-driven polynomial ridge approximation using variable projection. SIAM Journal on Scientific Computing, 40(3), A1566-A1589.

TECHNISCHE UNIVERSITÄT DARMSTADT

1 “Advanced” Machine Learning for Dynamical Systems

2 Nonlinear Dimension Reduction

3 Uncertainty Quantification

Uncertainty Quantification

Uncertainty Quantification (UQ) studies all sources of error and uncertainty; including the following systematic and stochastic measurement error; ignorance; limitations of theoretical models; limitations of numerical representations of those models: limitations of the accuracy and reliability of computations, approximations, and algorithms; human error. A more precise definition of UQ is the end-to-end study of the reliability of scientific inference.

- Understanding of uncertainty is crucial for certification, prediction, software verification and validation, parameter estimation, data assimilation, and inverse problems:
- UQ deals with two major Problem “classes”:

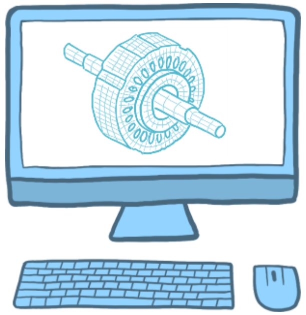
Forward Problem: Next to predicting the behavior of the system given all uncertain parameters and knowledge, the Forward problem focuses on the propagation of uncertainty

Inverse Problem: Given limited knowledge about the system, Inverse Problems determine some model properties based on the observed behavior of the real. Corresponding problems are very often hugely underdetermined.

Source: Sullivan, T. J. (2015). *Introduction to uncertainty quantification* (Vol. 63). Springer.

Uncertainty Quantification
Overview

Forward Techniques



Ensemble Forecasting

Polynomial Chaos

Stochastic Differential Equations

...

Inverse Techniques

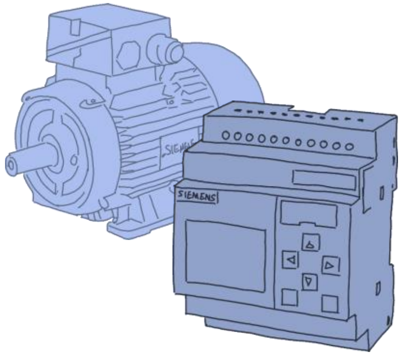
...

Adjoint Parameter Estimation

Kalman Filters

Bayesian Estimation

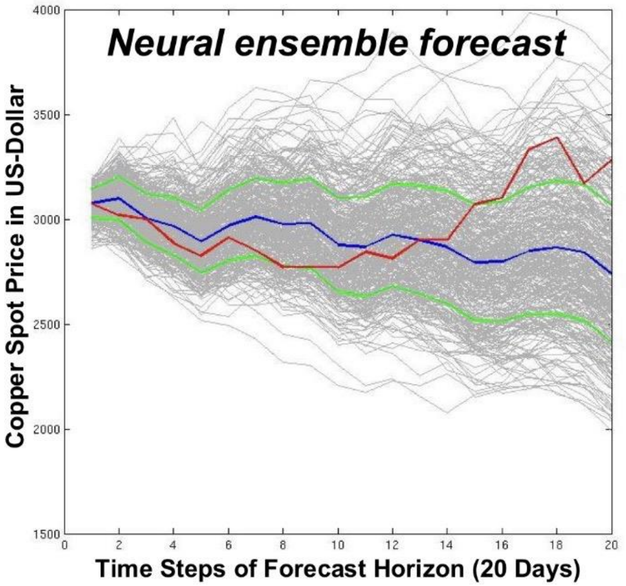
Gaussian Processes



Sources: Williams, C. K., & Rasmussen, C. E. (2006). *Gaussian processes for machine learning*. Cambridge, MA: MIT press.
Särkkä, S., & Solin, A. (2019). *Applied stochastic differential equations* .Cambridge University Press.
Pei, Y., Biswas, S., Fussell, D. S., & Pingali, K. (2019). An elementary introduction to Kalman filtering. *Communications of the ACM*, *62*(11), 122-133
Sullivan, T. J. (2015). *Introduction to uncertainty quantification* (Vol. 63). Springer..

Uncertainty Quantification
Ensemble Forecasting in Recurrent Neural Networks

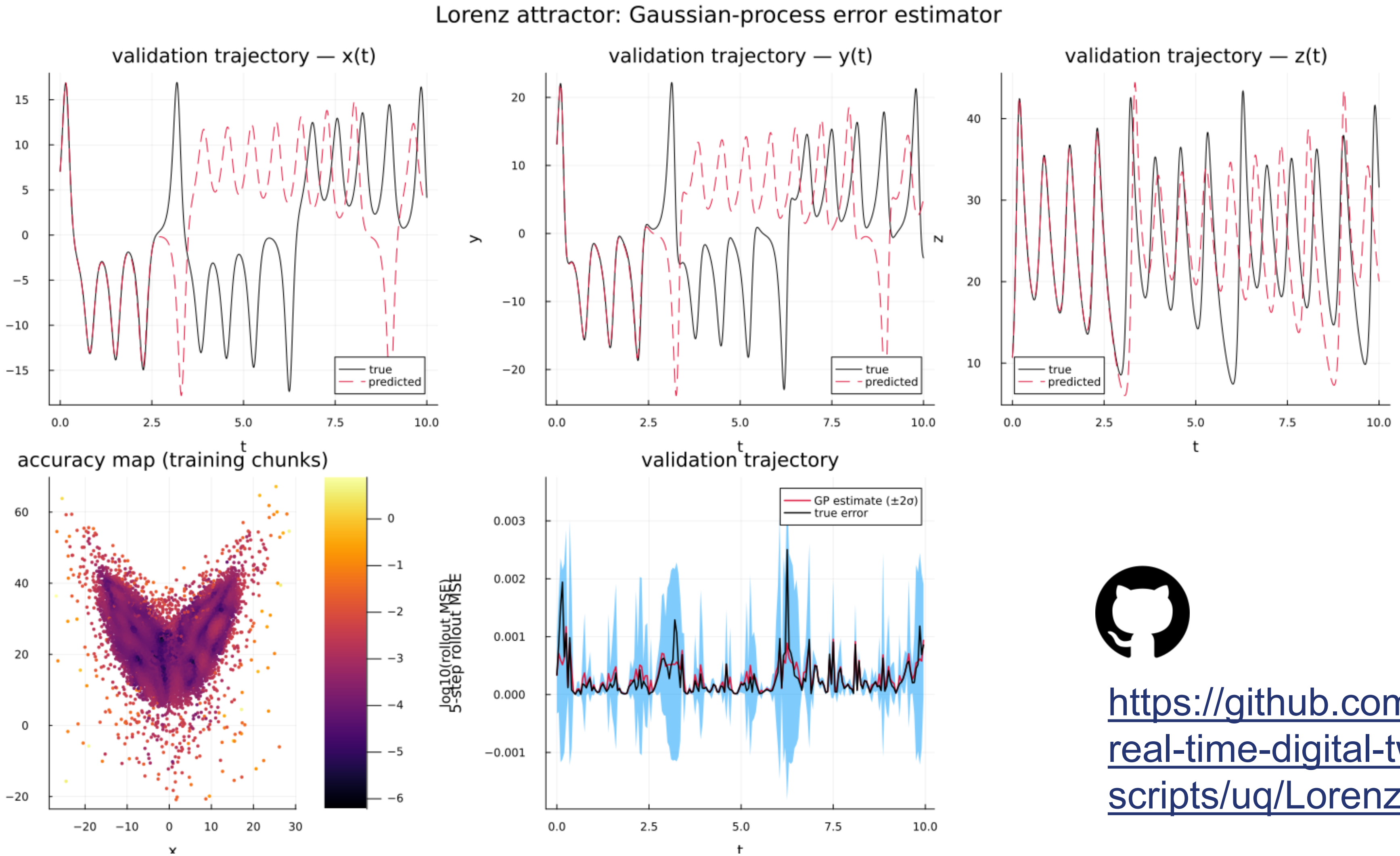
Copper Price Prediction vs. Actual Spot Price
(2008-12-15 – 2009-01-09)



- ► Ensemble forecasts can be an efficient way to characterize uncertainty.
- ► Given a finite set of data, there exists many perfect models of the past data, showing future scenarios caused by different estimations of the states.
- ► Since every ensemble member is a reasonable forecast and many models might fit perfectly we do not know which is the true one

Conjecture: forecast uncertainty = model uncertainty

Uncertainty Quantification
Gaussian Process Error Prediction



https://github.com/Dirk-Hartmann/real-time-digital-twins/blob/main/scripts/uq/Lorenz_NN_GP-ErrorEstimator.jl

Dirk Hartmann
Contact

Dr. Dirk Hartmann

S2|17 125
Schlossgartenstraße 8
64289 Darmstadt

E-mail dirk.hartmann@tu-darmstadt.de

Homepage: https://dirk-hartmann.github.io
TU Darmstadt: https://www.temf.tu-darmstadt.de/digitaltwin
LinkedIn: https://www.linkedin.com/in/dirkhartmann

The Art of the Possible – Blog: https://blogs.sw.siemens.com/art-of-the-possible