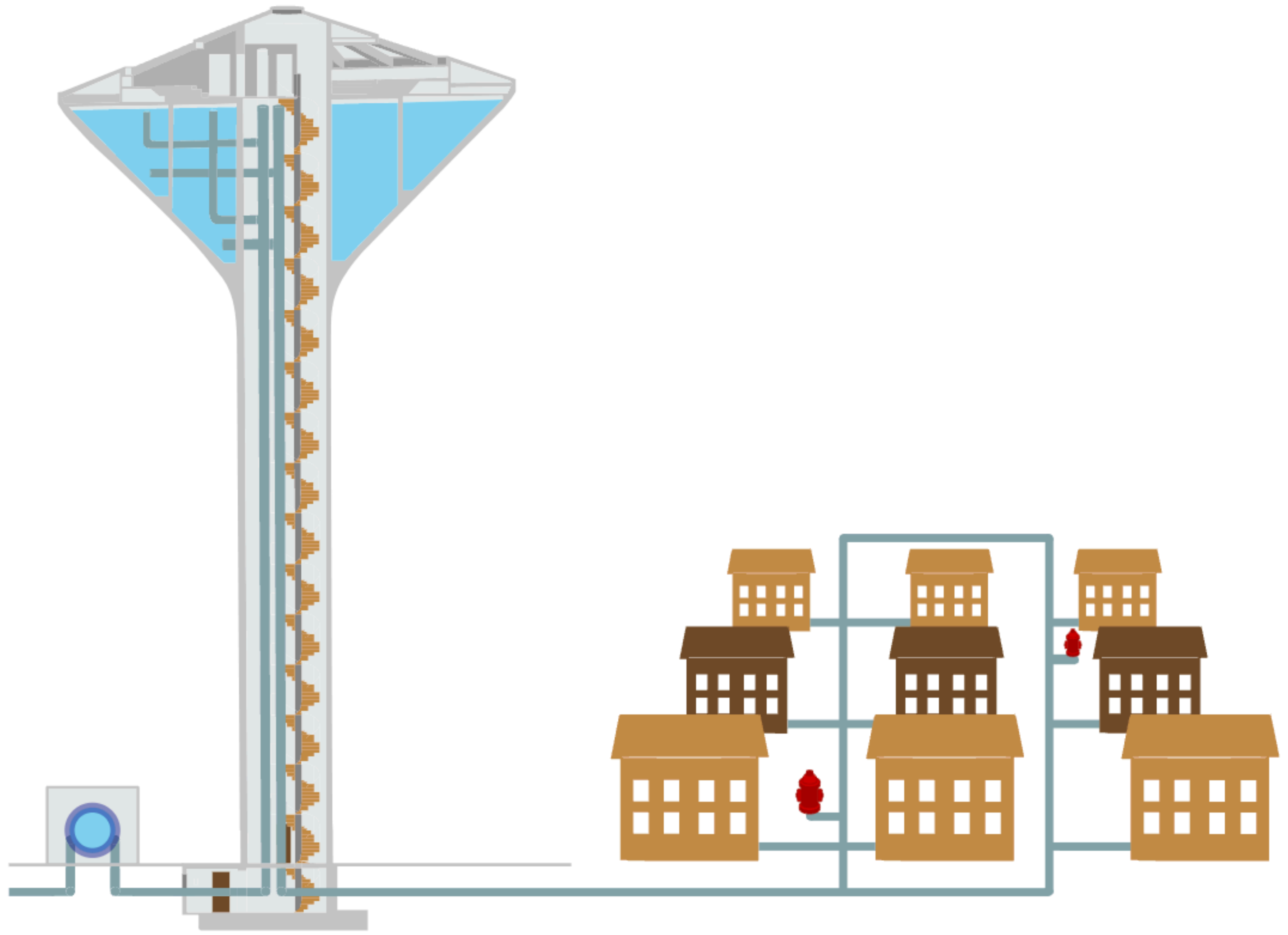
VANDERBILT UNIVERSITY

Composable Physics-Informed Learning with Uncertainty Quantification

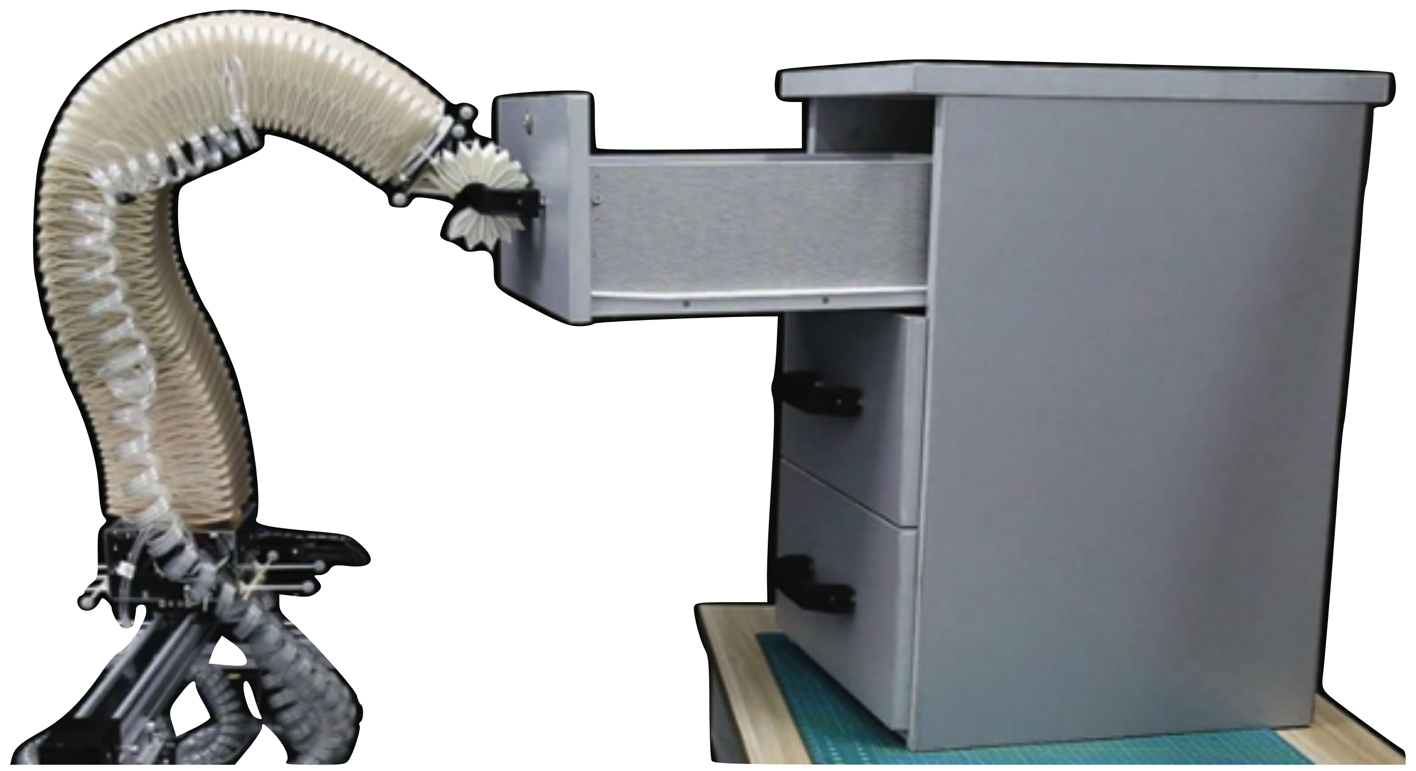
Thomas Beckers
Vanderbilt University

ACC Workshop on Physics-informed Machine Learning for Modeling, Control, and Optimization
July 9th, 2024

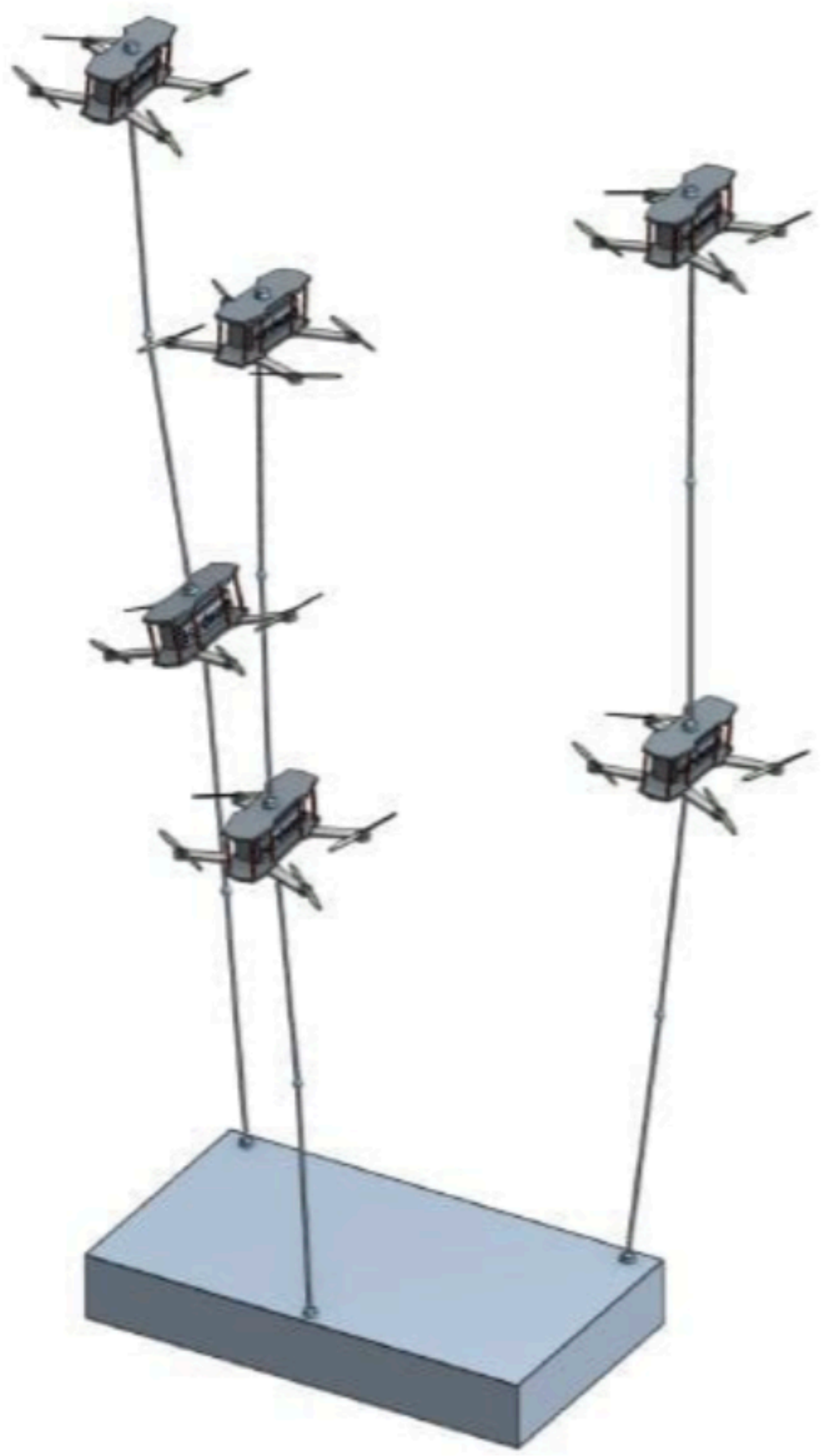
Motivation



Water-distribution network [Wikipedia]



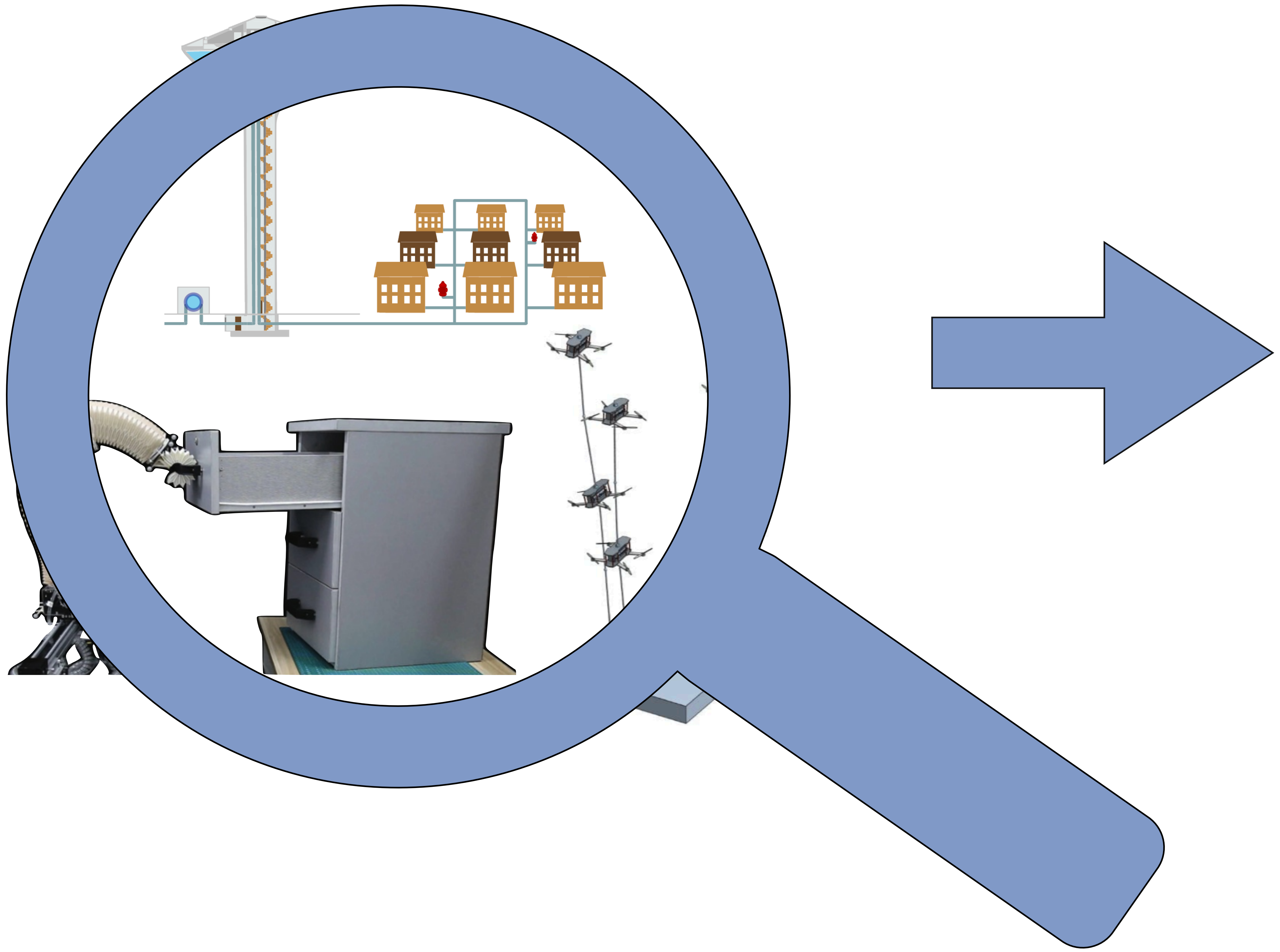
Soft robot manipulation [Jiang Hao]

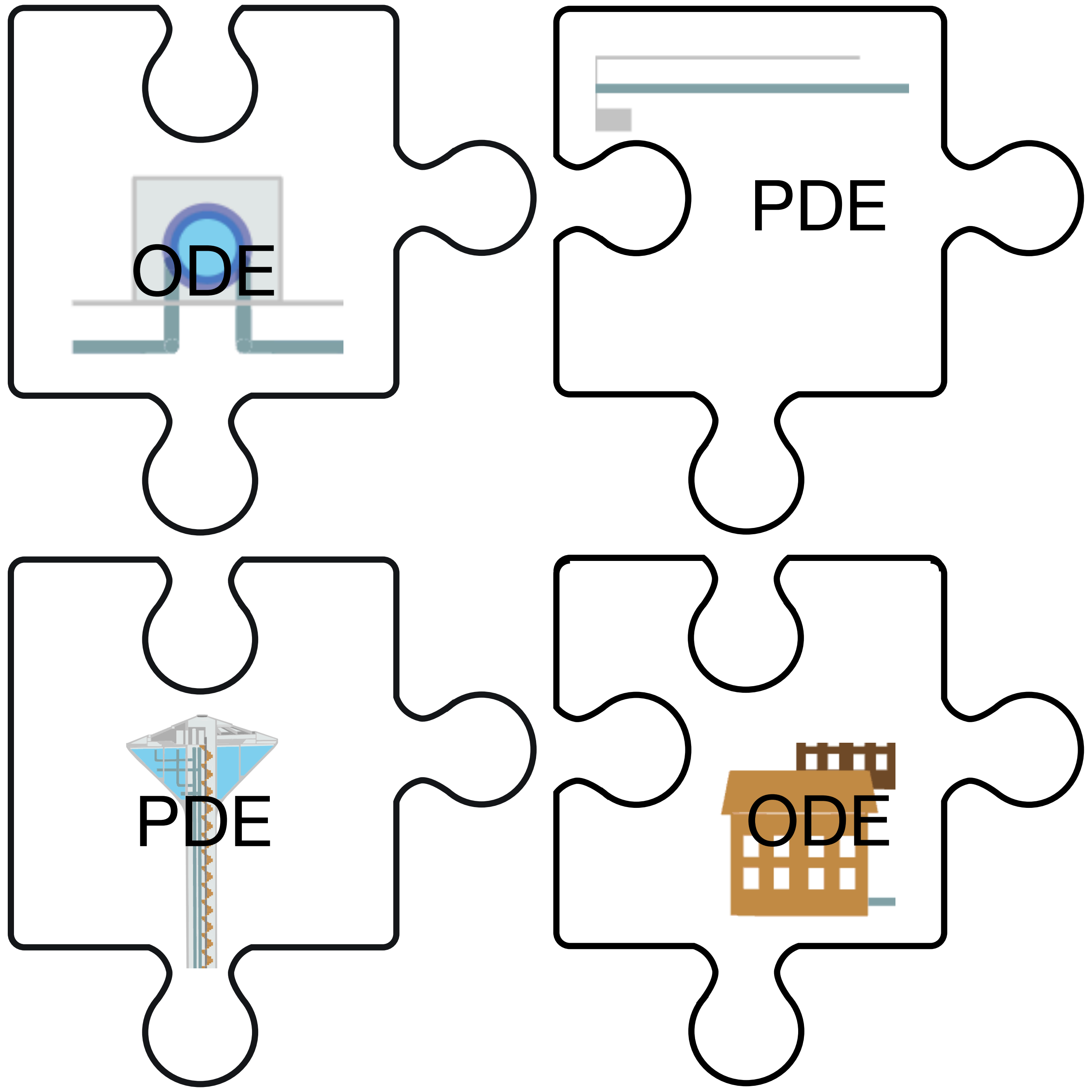


Cooperative drones [Muhammad Usama]

Need of **accurate and reliable models** **of complex systems**
for control, test & verification, optimization, fault-detection, …

Combosable structure



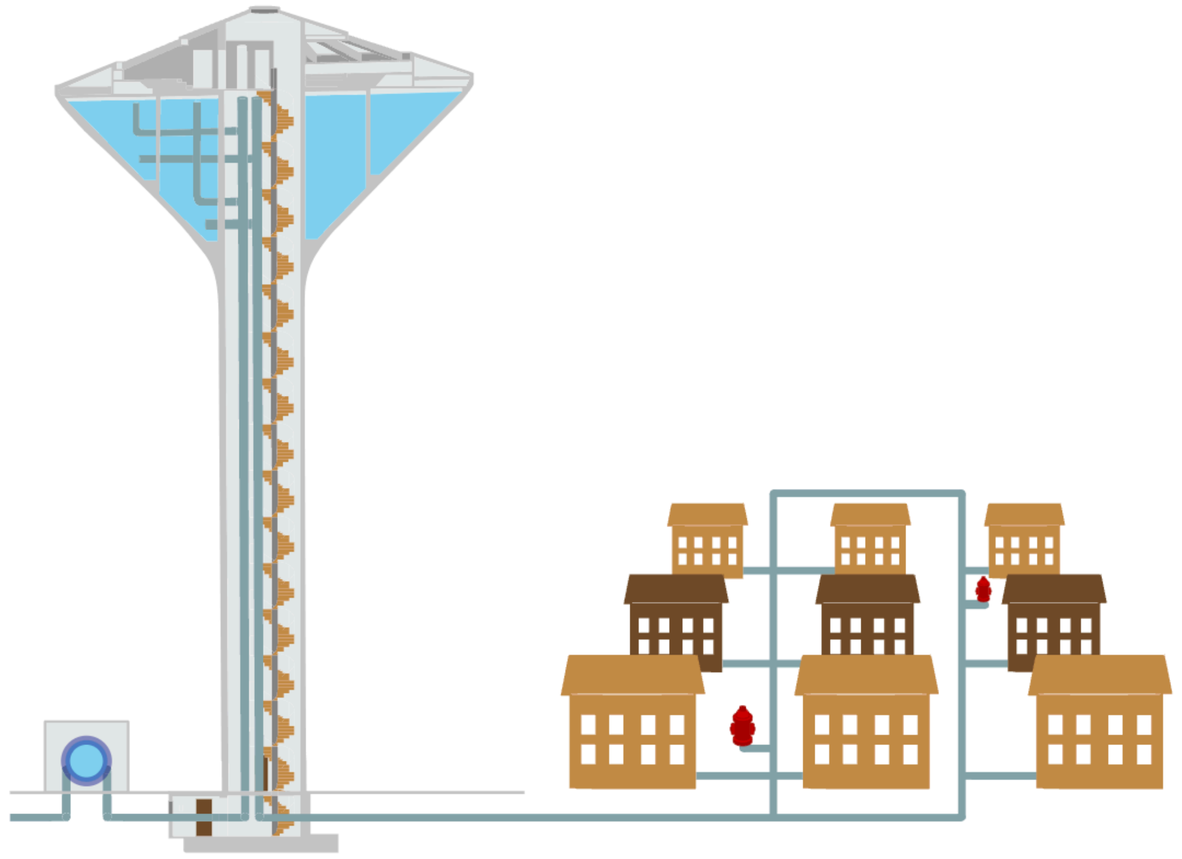


Many complex systems are composed of coupled subsystems

=> Exploiting the composable structure for modeling

- Separation (domain knowledge, verification, optimization, …)
- Parallelization (time-efficient modeling)
- Interpretability (understanding the impact)

Modeling of complex systems



Physics-based

Pump

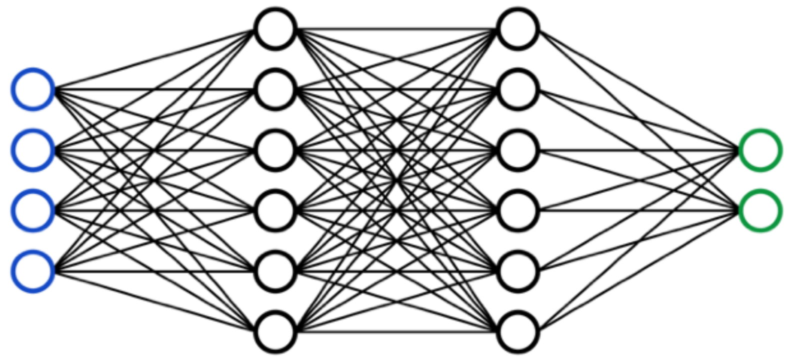
$$\dot{x} = f(x, u) = \dots$$

Pipe

$$0 = f\left(\frac{\partial x}{\partial t}, \frac{\partial x}{\partial z}, \frac{\partial^2 x}{\partial z^2}\right) = \dots$$

- ✓ Generalization
- ✓ Composable models
- ✗ Model selection (nonlinear)
- ✗ Time-consuming

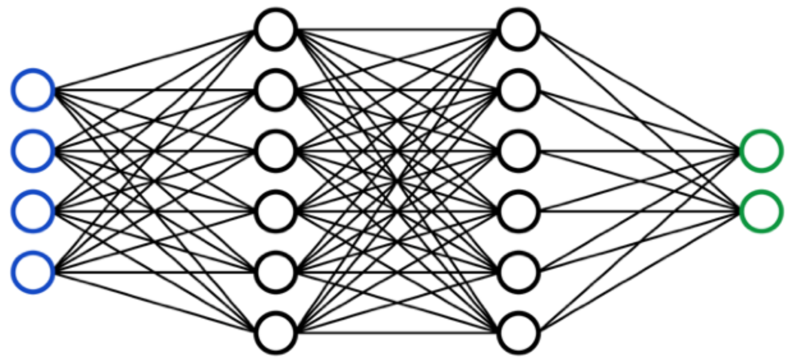
Learning-based



- ✓ Expressive models
- ✓ Minimal expert knowledge
- ✗ Physical correctness
- ✗ Trustworthiness
- ✗ Composition

Modeling of complex systems

Solutions for

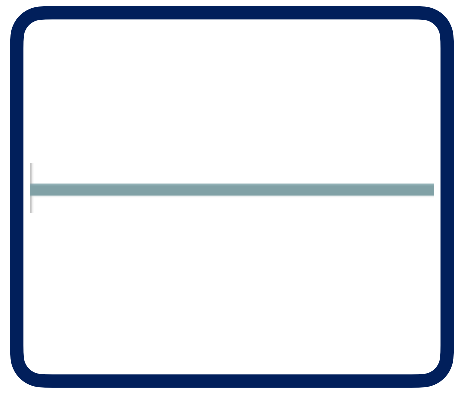
Learning-based	✓ Physical correctness	✓ Trustworthiness
	- SINDy - PINNs - Neural ODE - Neural PDE - …	- Bayesian methods - Conformal prediction - Ensemble Methods - Monte Carlo - …

Pump model

Physically correct
and trustworthy

+

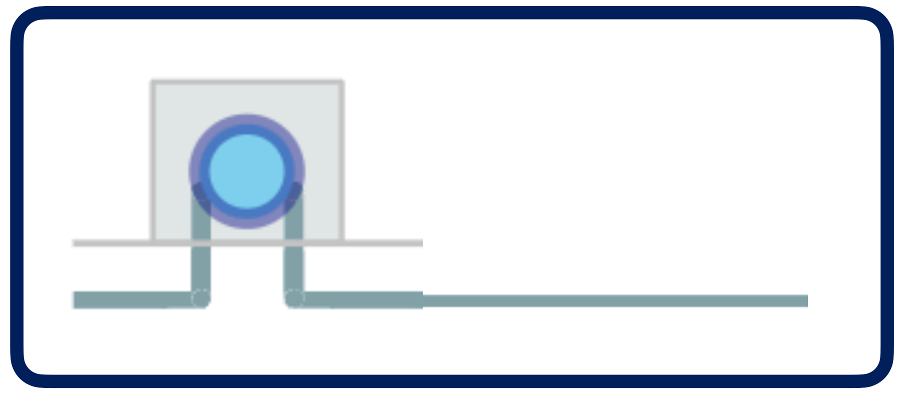
Pipe model



Physically correct
and trustworthy

=

Composed model



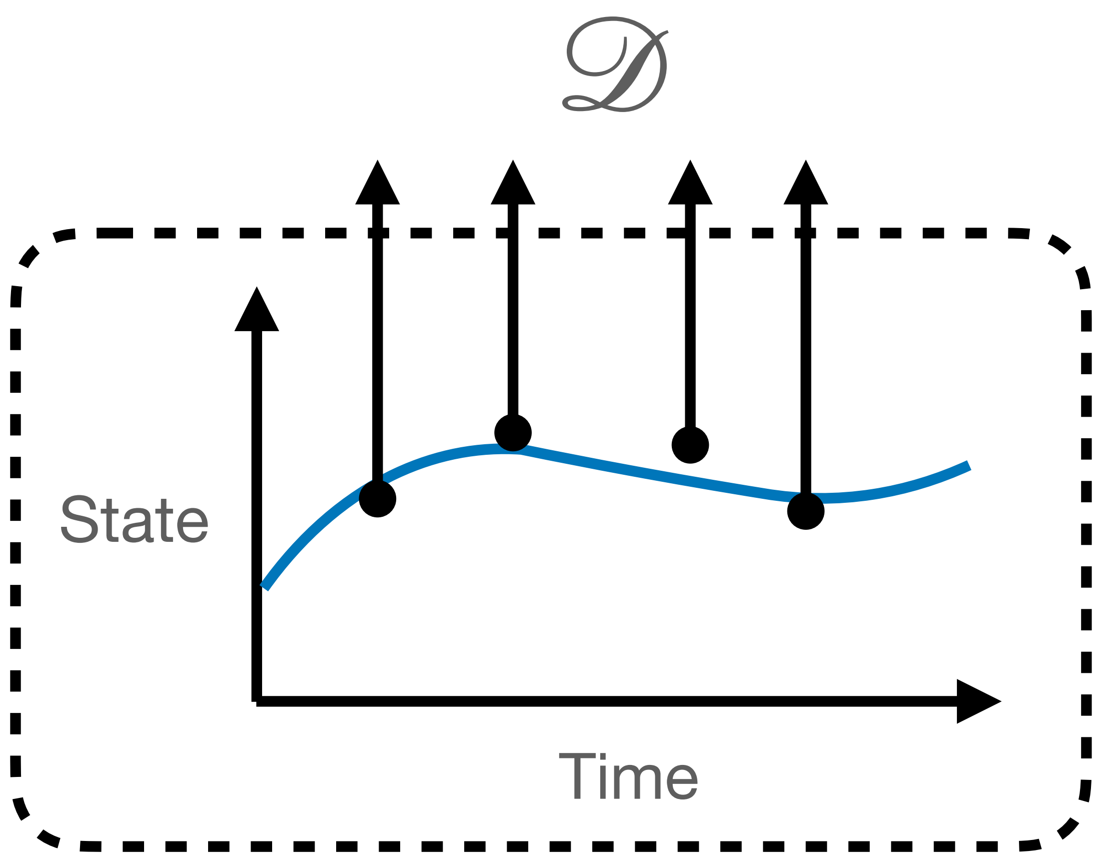
Physically correct ?
Trustworthy ?

Problem statement

Physical (ODE) system $\dot{x}(t) = \overbrace{f(x) + G(x)}^{\text{unknown}} u(t)$

Noisy measurements $\tilde{x}_i = x(t_i) + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma_n^2 I)$

Dataset $\mathcal{D} = \{(t_i, \tilde{x}_i, u_i)\}_{i=0}^{N-1}$



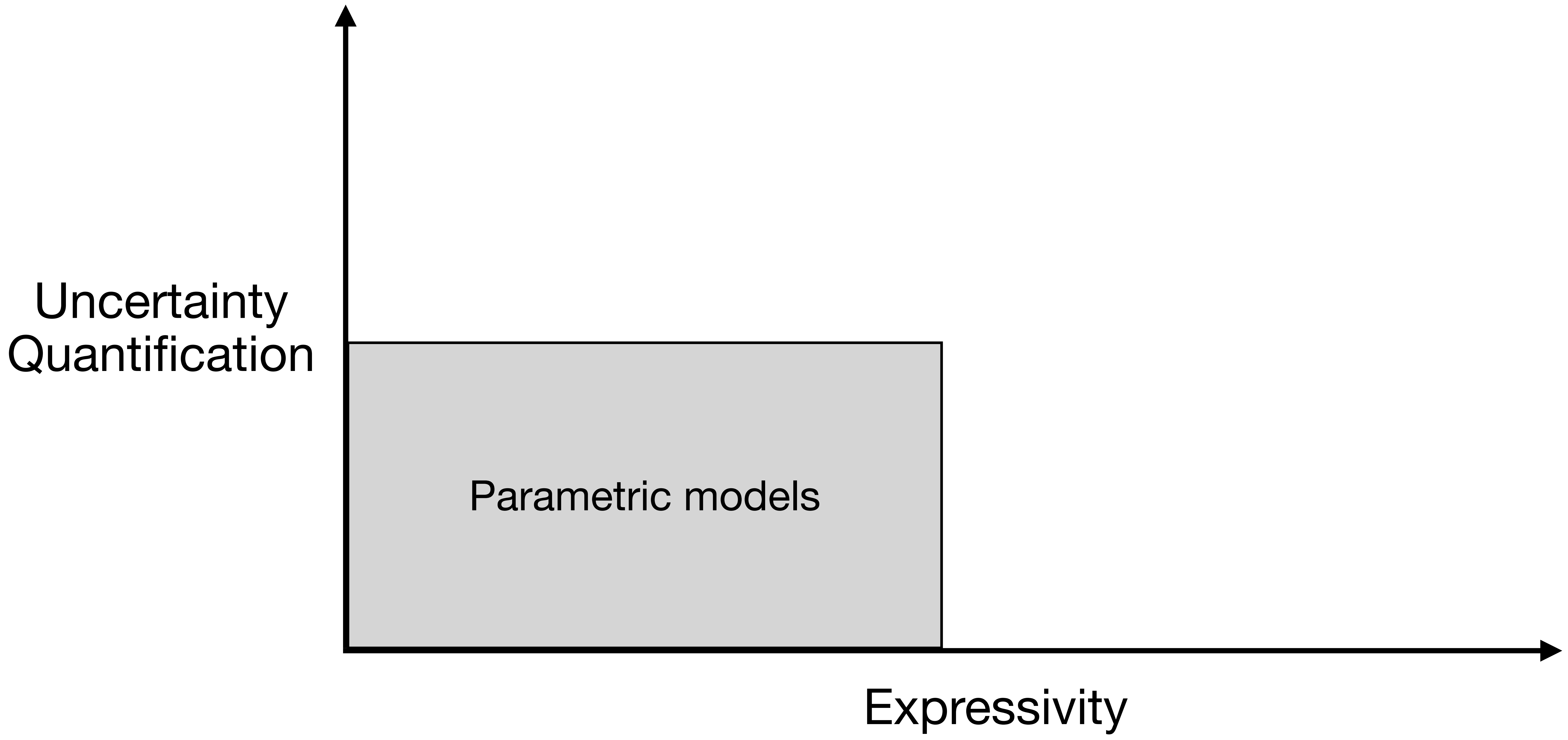
System

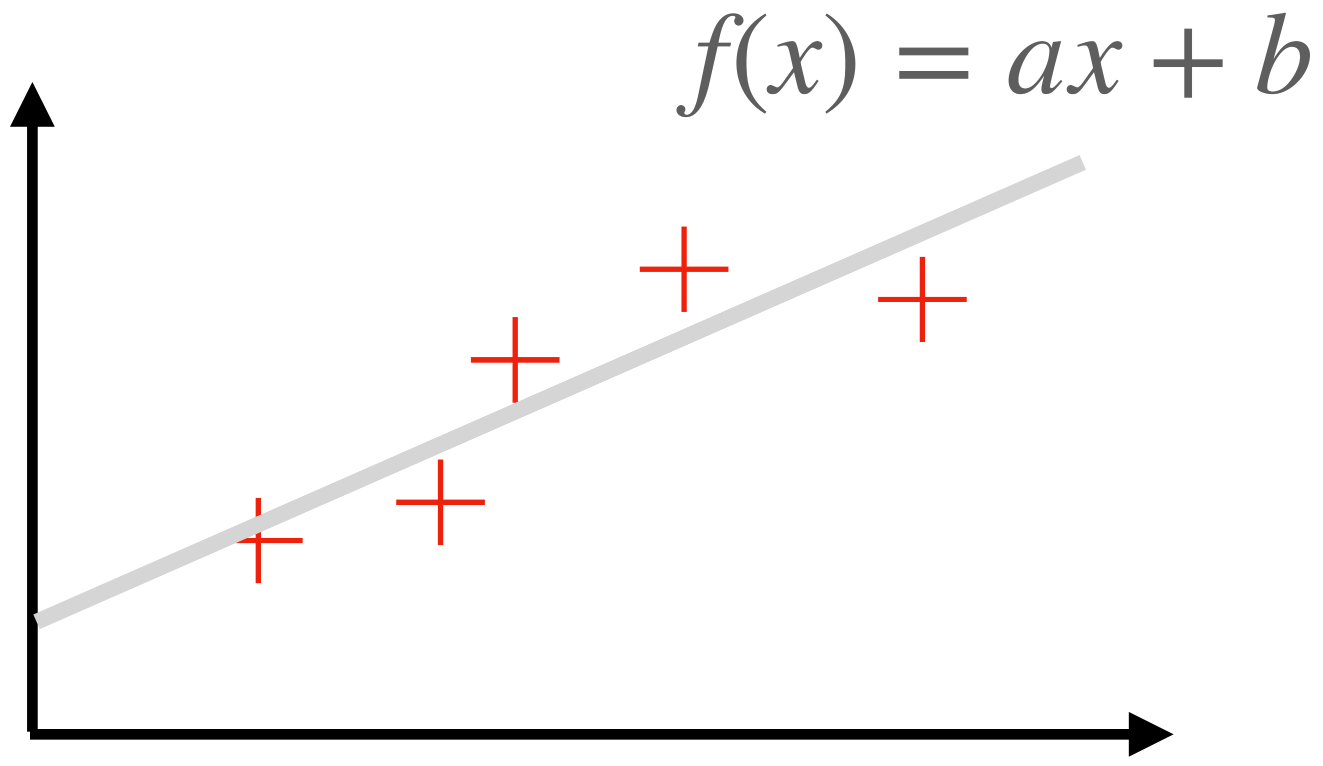
Goal

Given the dataset $\mathcal{D}$, we aim to find a dynamical model that

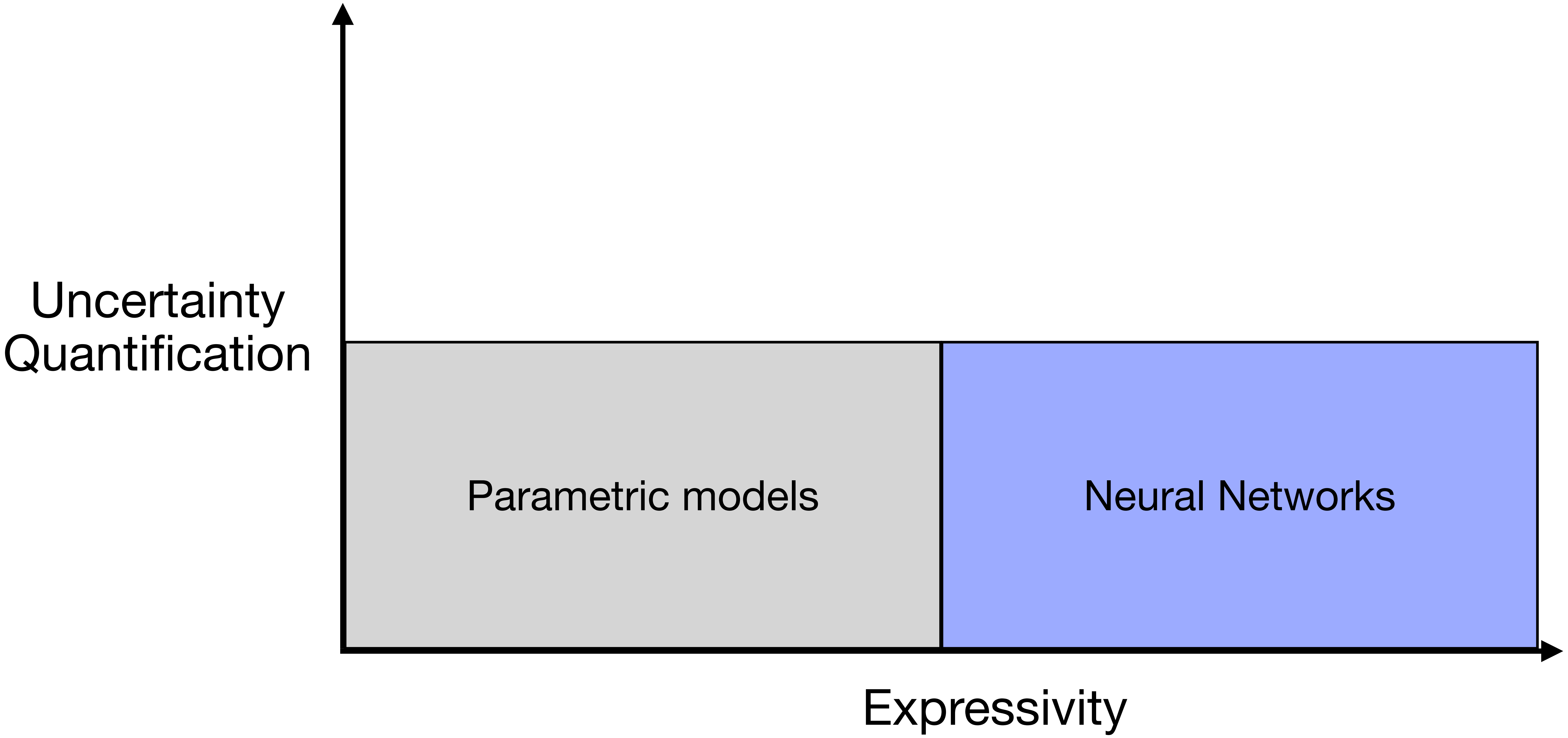
- represents a large class of systems -> Expressive
- allows uncertainty quantification -> Reliable
- respects physics -> Generalizes well
- Composable -> beneficial for modeling and control

Learning methods

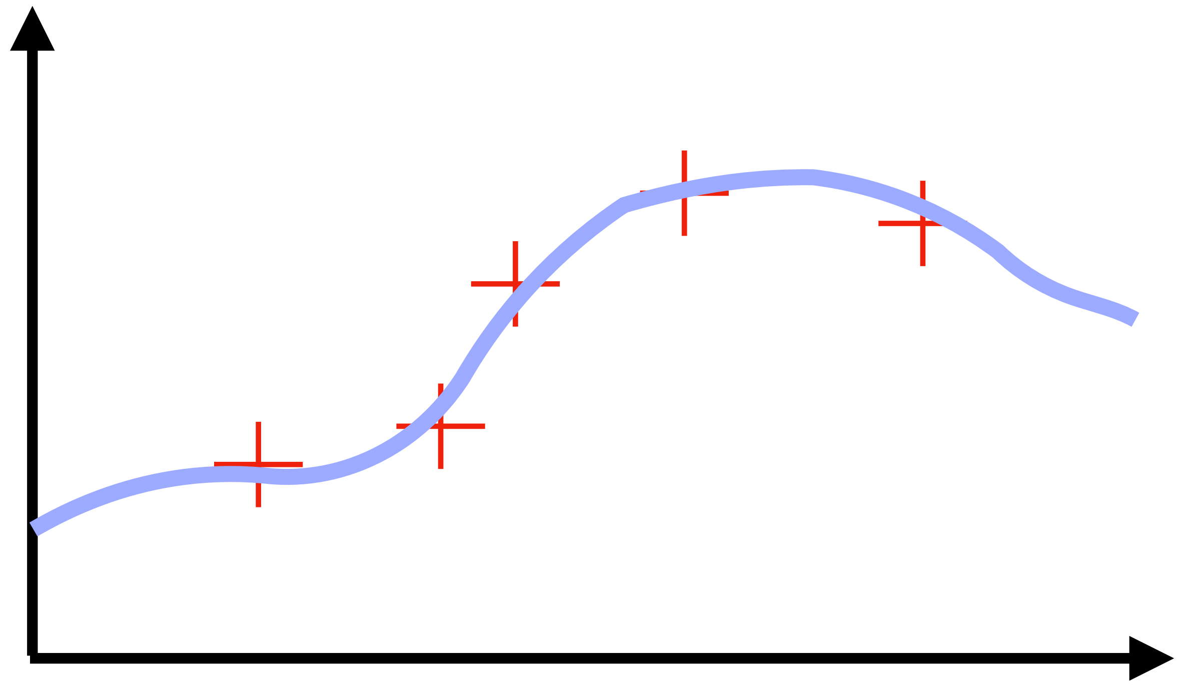




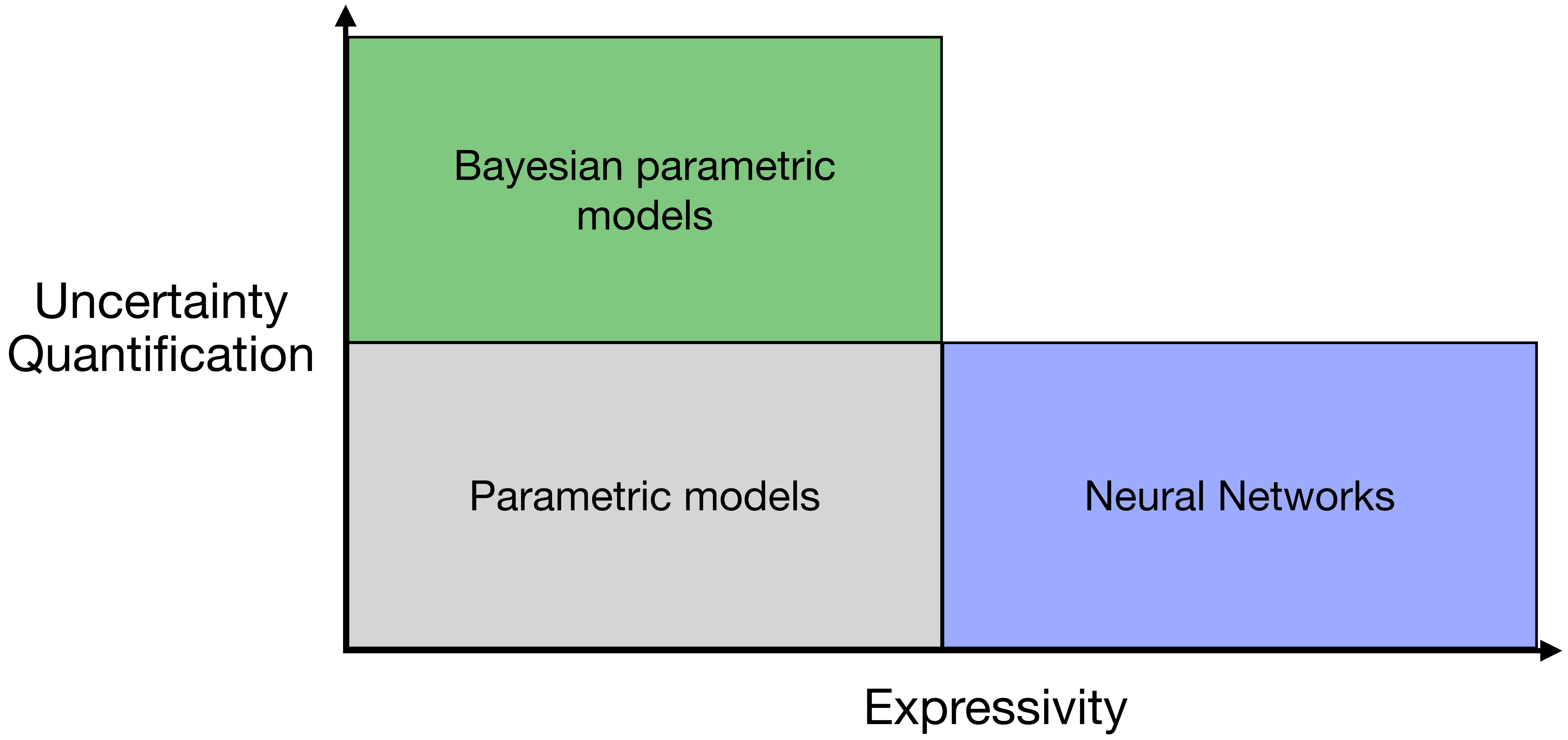
Learning methods

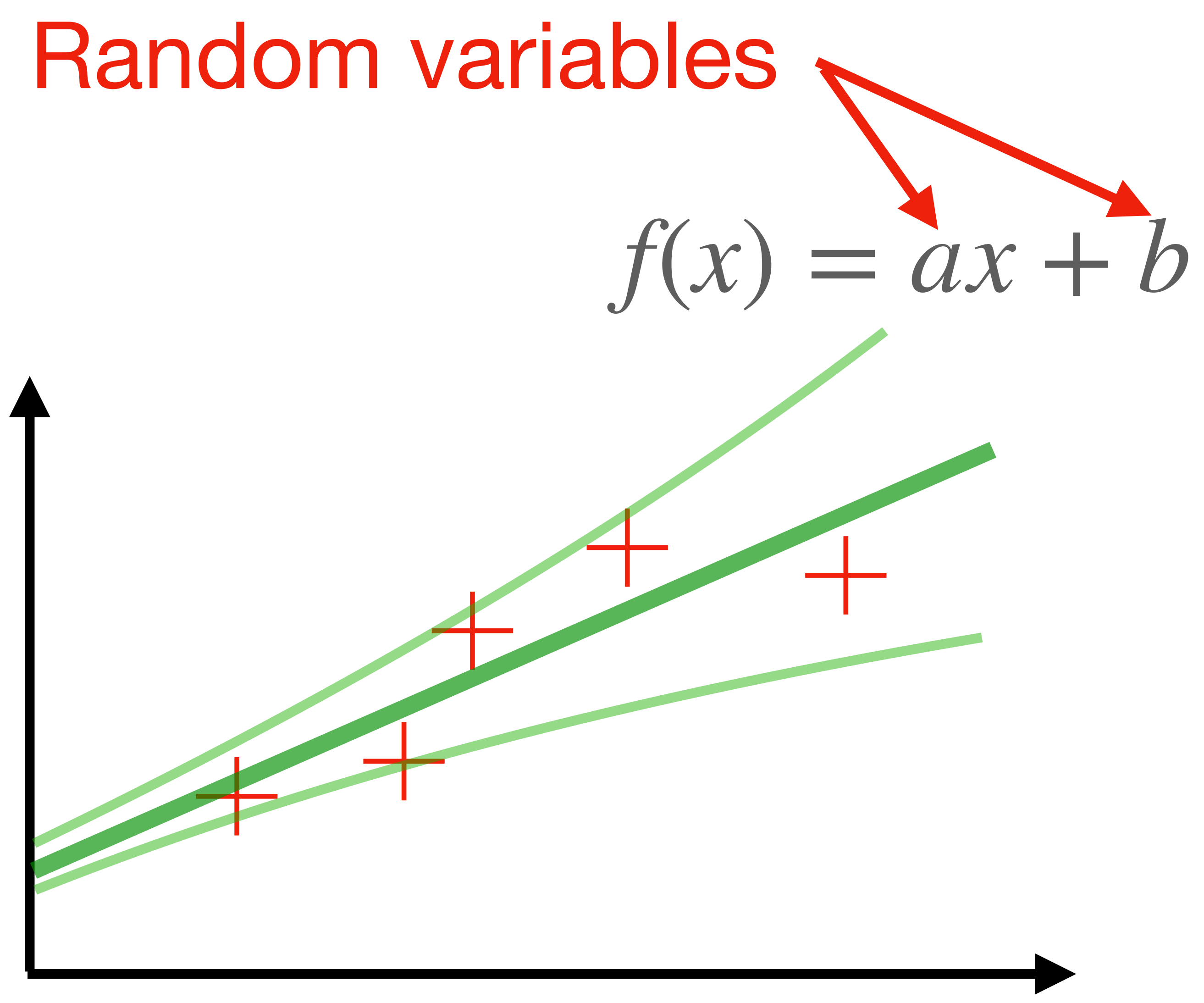


$$f(x) = \sum w_i \sigma_i(x)$$

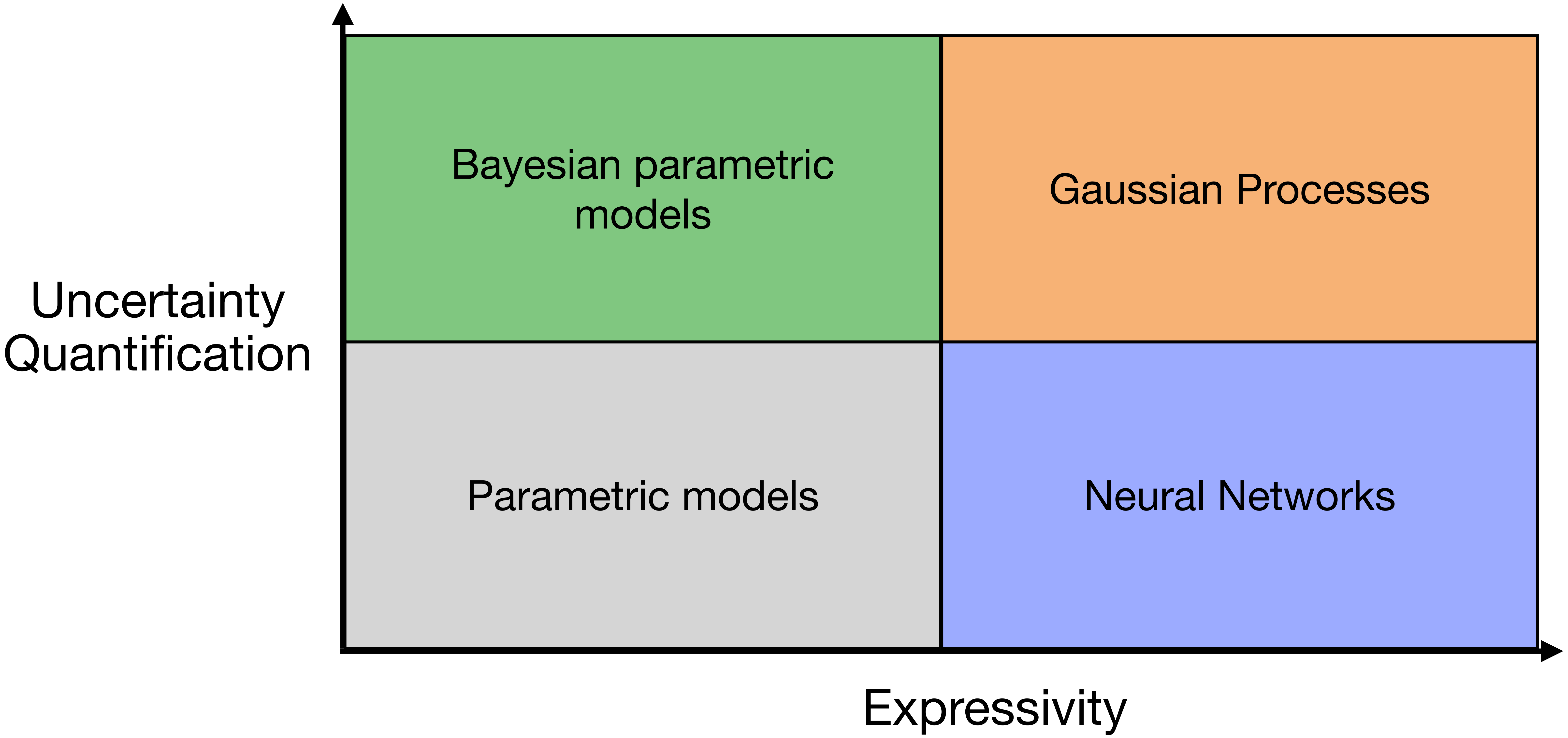


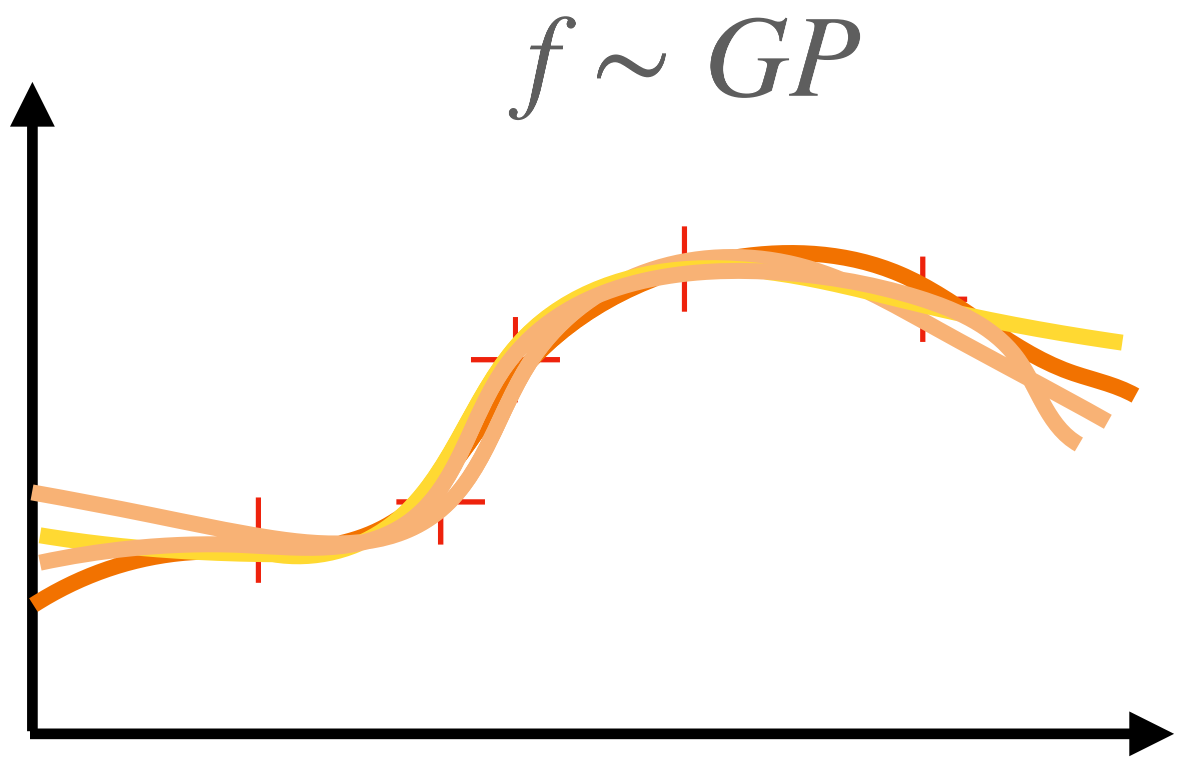
Learning methods





Learning methods





Gaussian Process: Prior knowledge + nonparametric + uncertainty quantification

Gaussian process

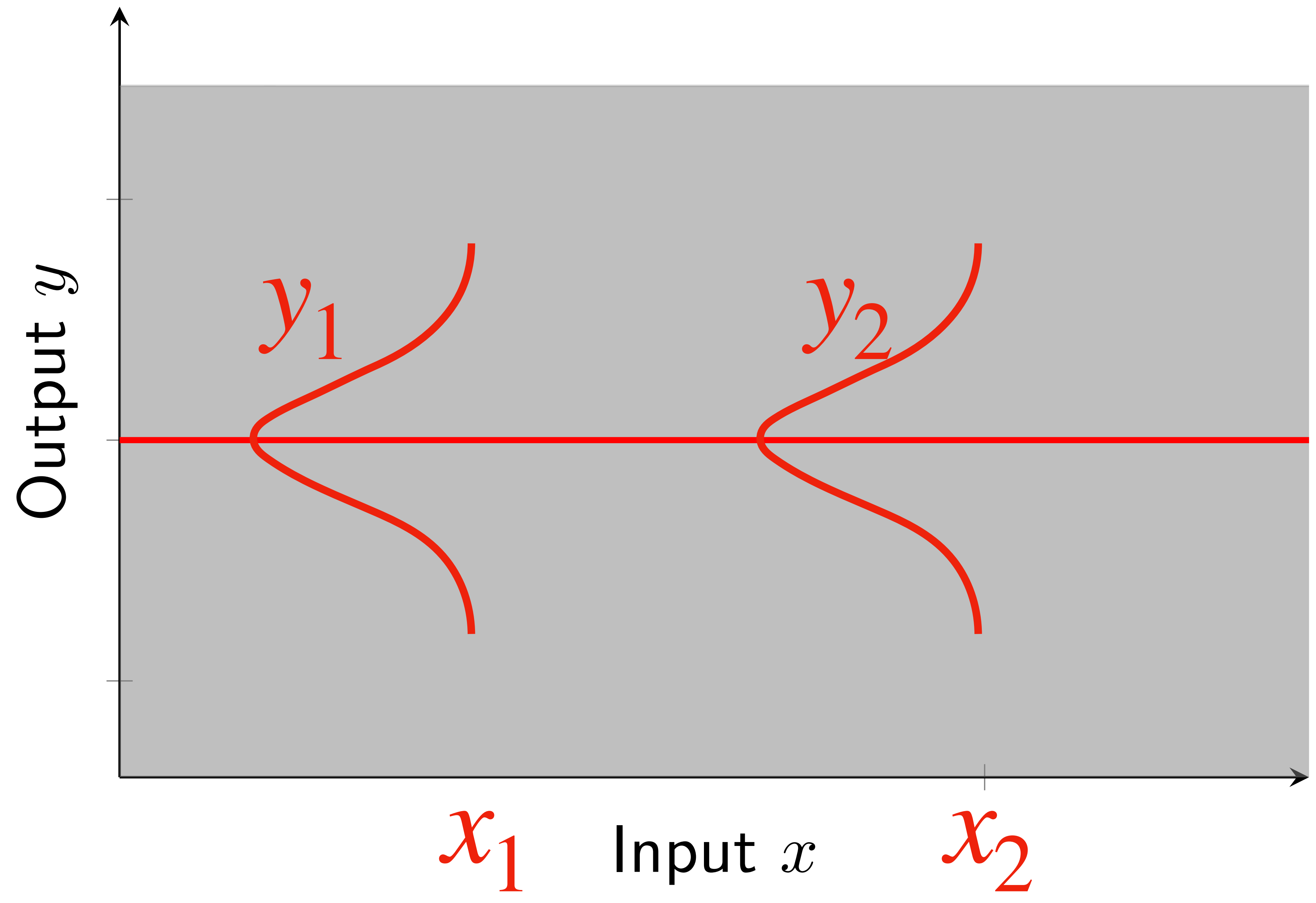
Prior: Gaussian distribution over function space

$$f(x) \sim GP(m(x), k(x, x'))$$

Mean function

Covariance

A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution.



$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m(x_1) \\ m(x_2) \end{bmatrix}, \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) \\ k(x_2, x_1) & k(x_2, x_2) \end{bmatrix} \right)$$

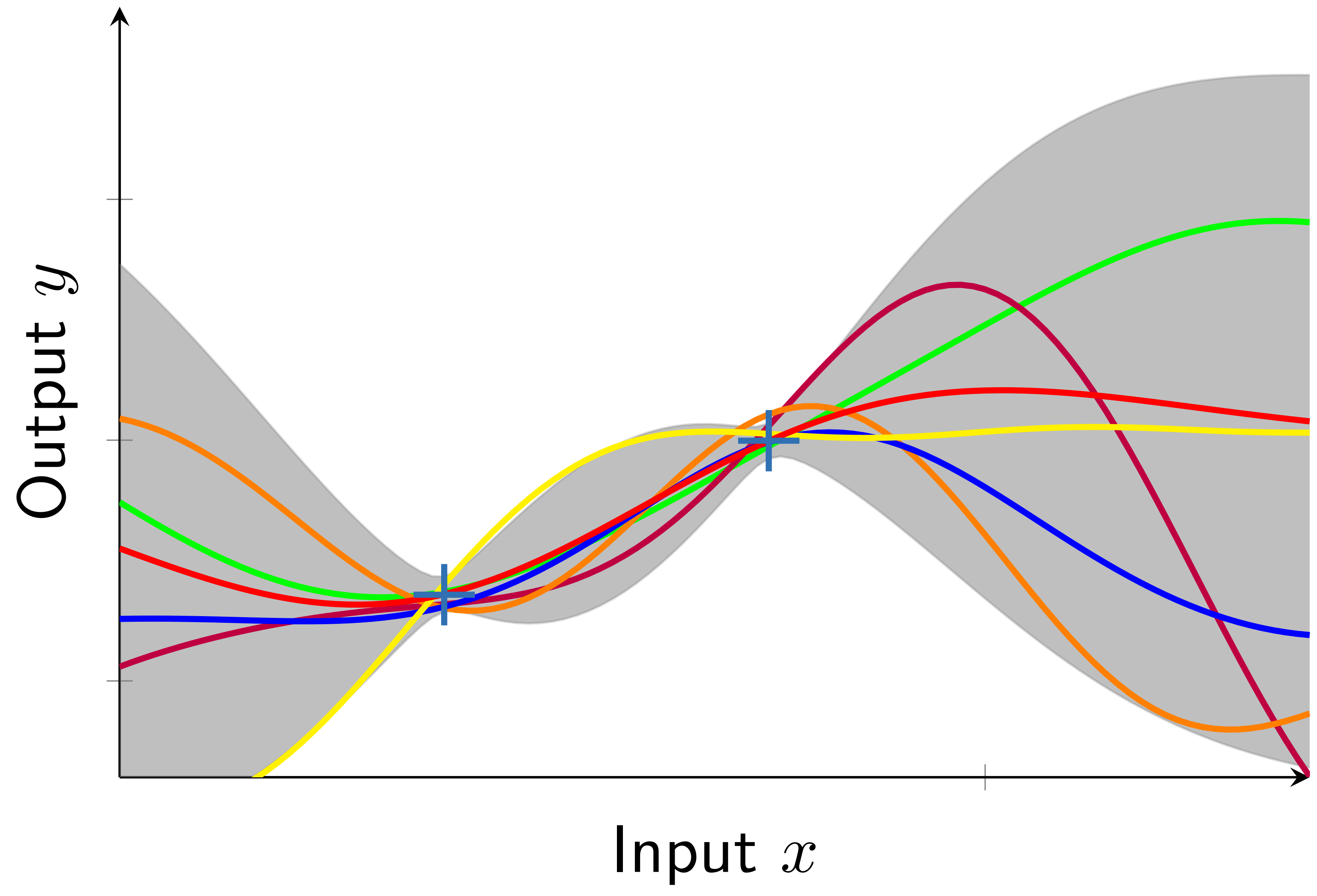
Gaussian process

Prior: Gaussian distribution over function space

$$f(x) \sim GP(m(x), k(x, x'))$$

Mean function

Covariance

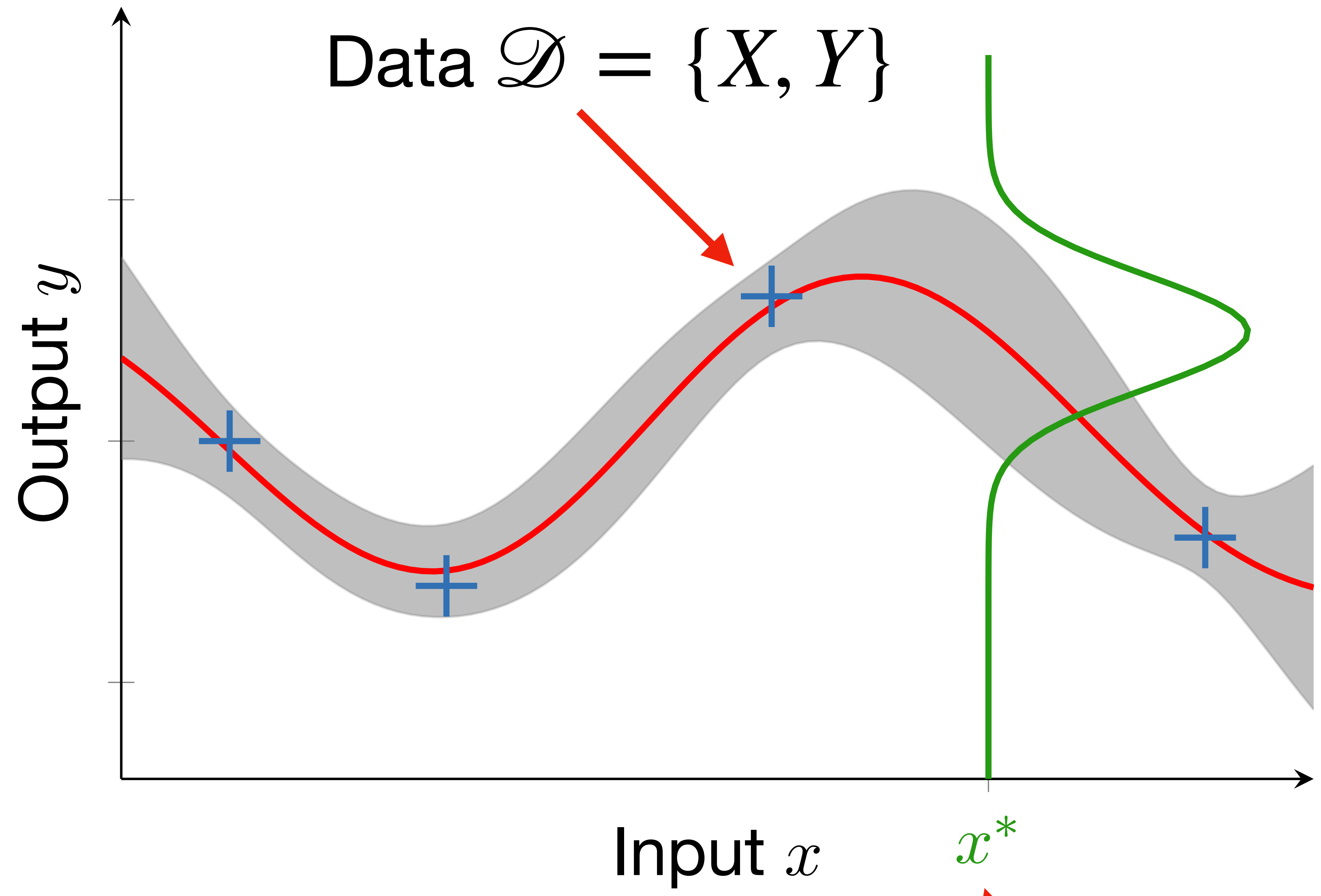


Gaussian process

Prior: Gaussian distribution over function space

$$f(x) \sim GP(m(x), k(x, x'))$$

Mean function Covariance



Posterior

$$\mu(y^* \mid x^*, \mathscr{D}) = m(x^*) + k(x^*, X)[K(X, X) + \sigma_n^2 I]^{-1}(Y - m(X))$$

$$\Sigma(y^* \mid x^*, \mathscr{D}) = k(x^*, x^*) - k(x^*, X)[K(X, X) + \sigma_n^2 I]^{-1} k(x^*, X)^\top$$

C.E. Rasmussen. Gaussian Processes for Machine Learning. 2006

Prior knowledge

Posterior mean

Data points   Covariance

$$\mu(y\,|\,x, \mathscr{D}) = m(x) + \sum_{i}^{N} w_i k(x, X_i)$$

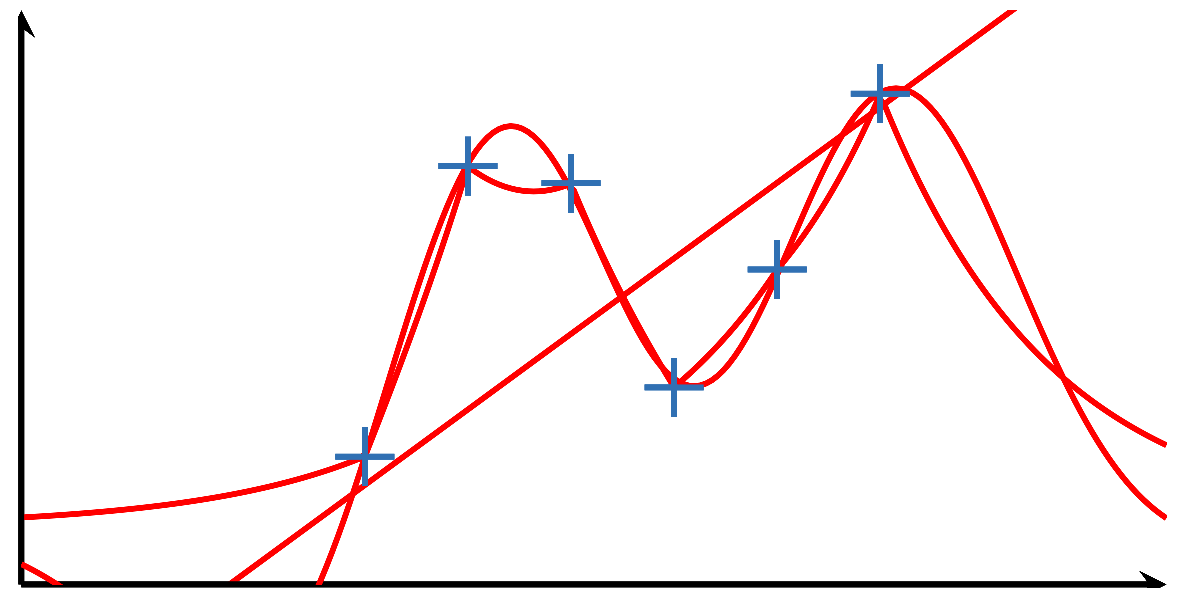
Prior mean function   Weighting factor

Via mean function

$$m(x) = \ldots$$

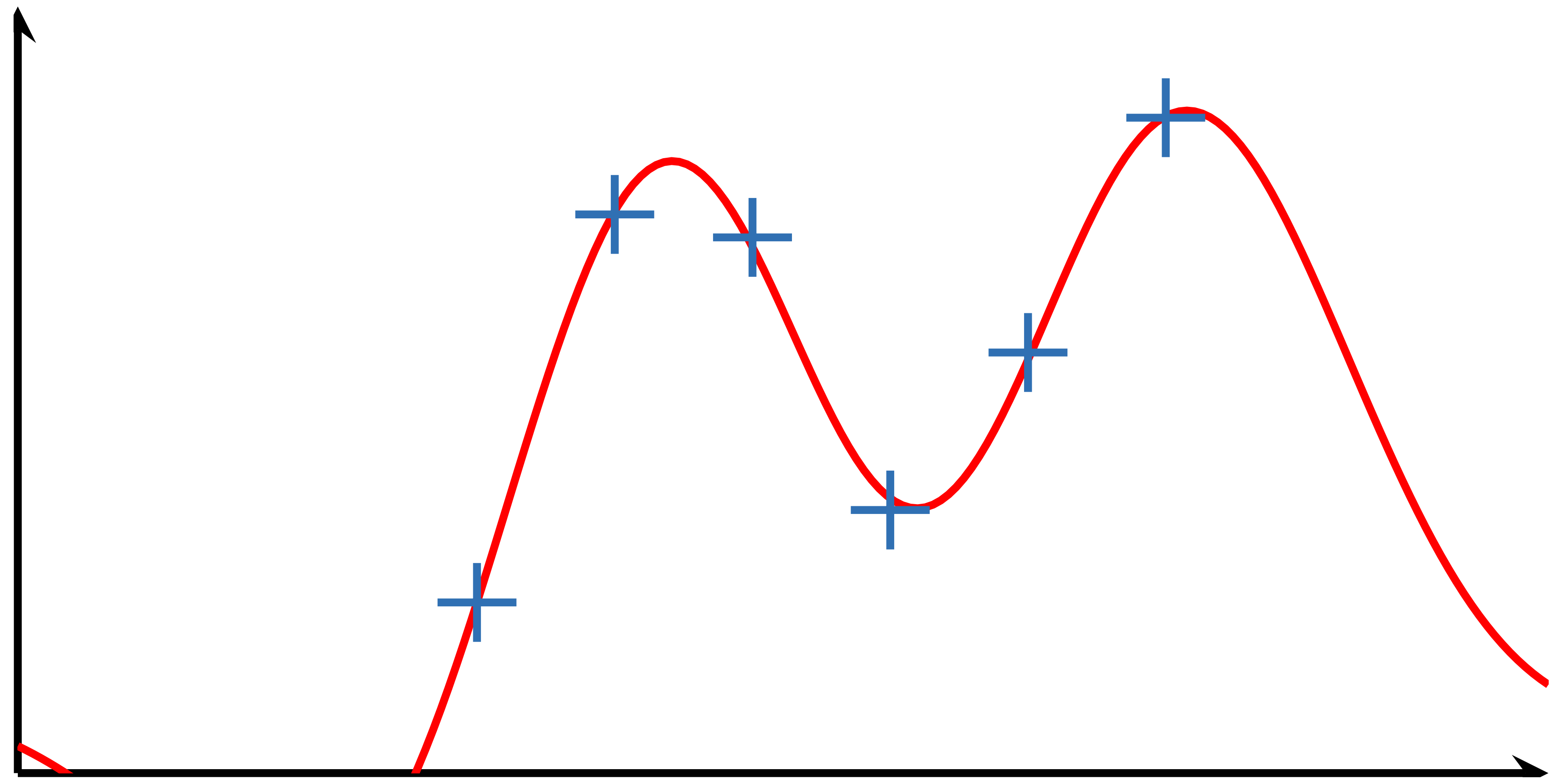
Discrepancy bias

Via covariance function



Inductive bias

Covariance function



Linear: $k(x, x') = x^\top x$

Matern: $k(x, x') = \sigma_f^2 \exp\left(\frac{\|x - x'\|}{2l}\right)$

Squared exponential: $k(x, x') = \sigma_f^2 \exp\left(\frac{\|x - x'\|^2}{2l^2}\right)$

?

→ **Encode physics**

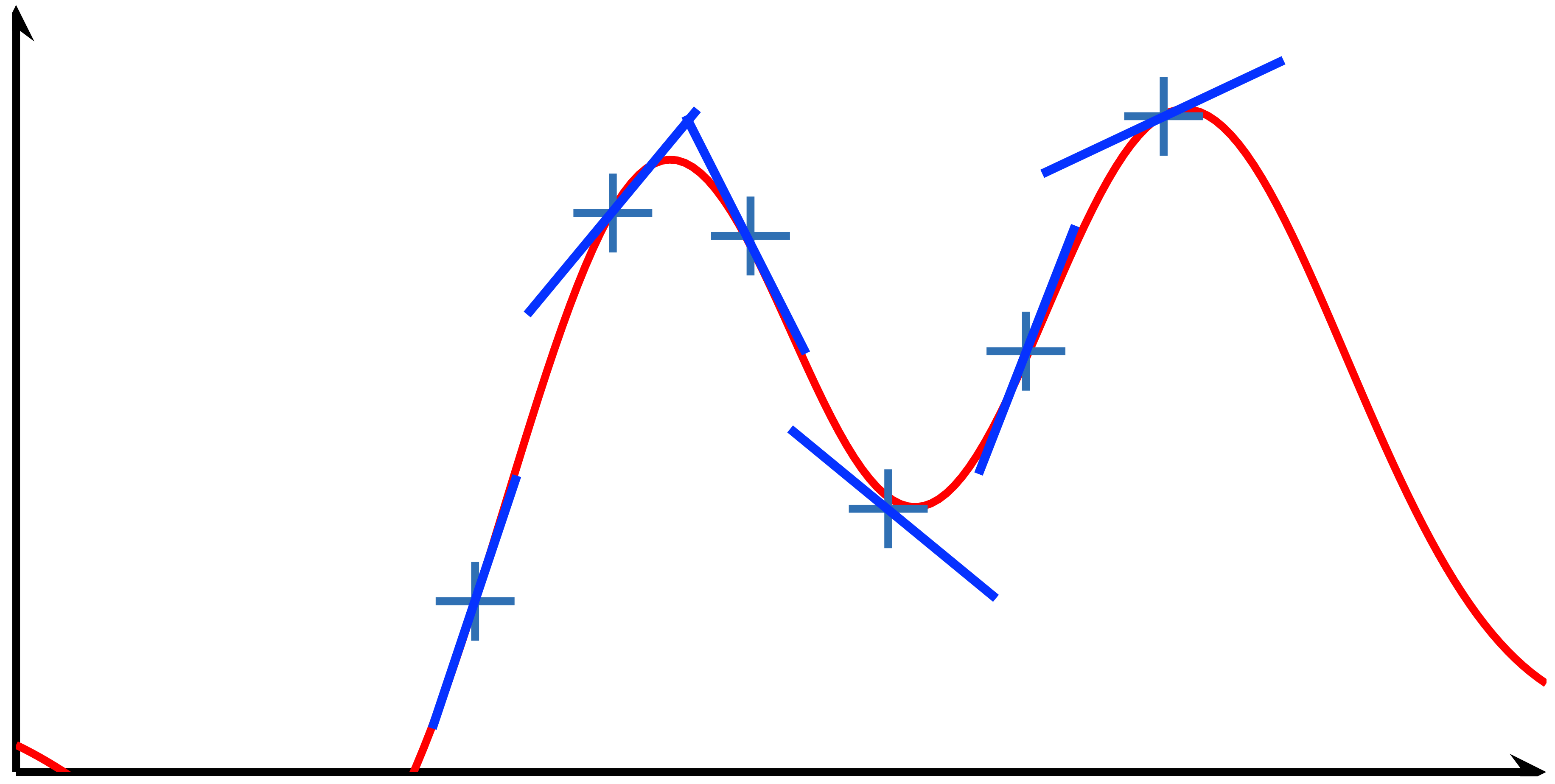
Linear operators

GPs are closed under linear operators

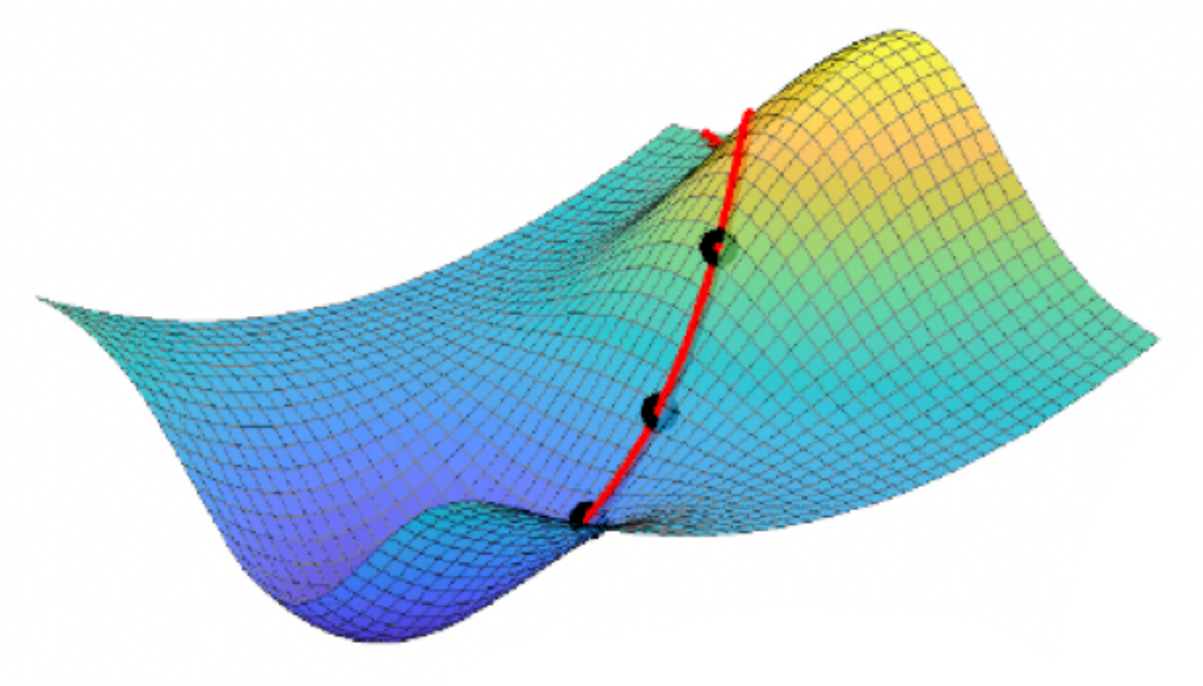
$$\boxed{f(x) = \mathscr{L}_x g(x) \sim GP(\mathscr{L}_x \mu, \mathscr{L}_x k \mathscr{L}_x^\top)}$$

Example: Differential operator $\mathscr{L}_x = \dfrac{\partial}{\partial x}$

$$\begin{bmatrix} y \\ Y' \end{bmatrix} = \mathcal{N}\left(0, \begin{bmatrix} k(x,x) & \frac{\partial}{\partial x}k(x,X) \\ \frac{\partial}{\partial x}k(x,X)^\top & \frac{\partial}{\partial x}^\top k(X,X)\frac{\partial}{\partial x} \end{bmatrix}\right)$$



Physics model



Port-Hamiltonian systems

Interconnection Dissipation Hamiltonian

$$\dot{x} = [J(x) - R(x)]\frac{\partial H}{\partial x} + G(x)u$$

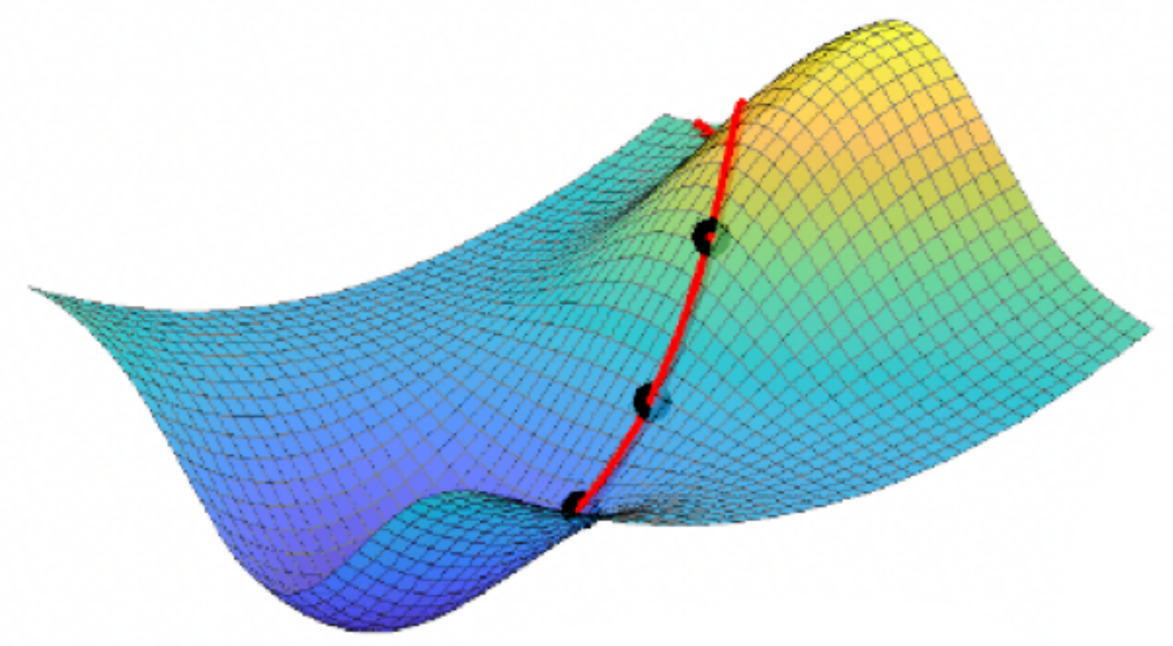
Input

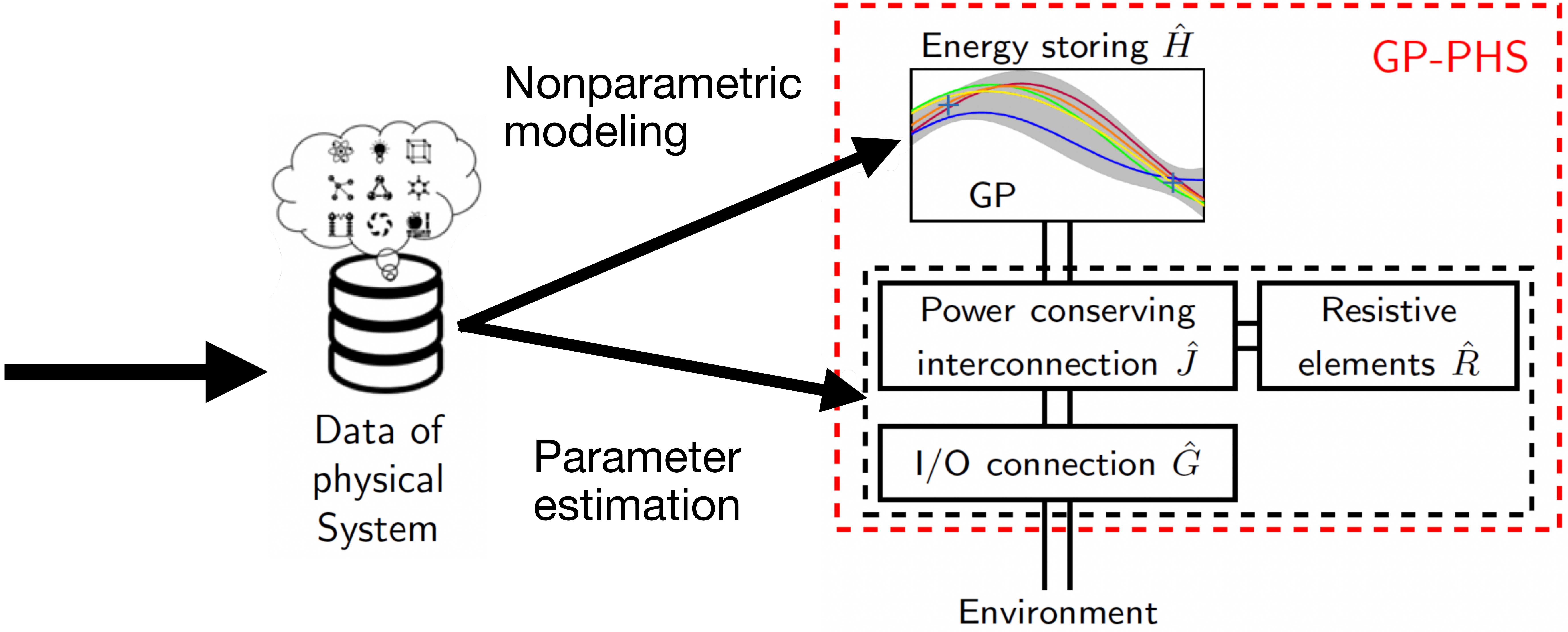
$$y = G(x)^\top \frac{\partial H}{\partial x}$$

Output

- Skew-symmetric J ensures energy conservation
- Input and output defines a passive systems
- Connection with another PHS preserves the structure
- Suitable for multi-domain applications via energy flows

Physics-constrained learning

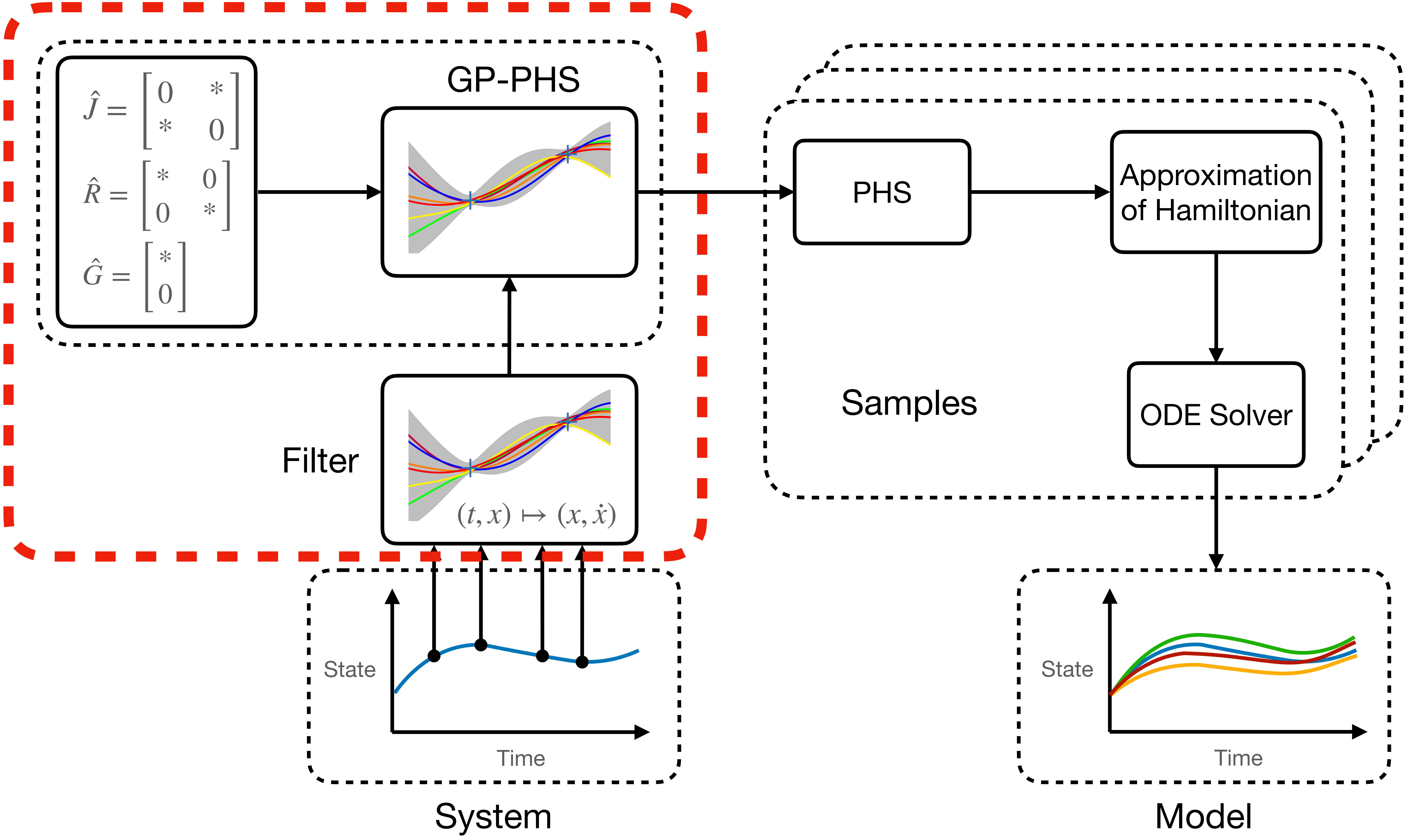




- Physics-constrained Gaussian Process model
- GP-PHS includes all possible PHS under the GP prior on the Hamiltonian

T. Beckers, et al. Gaussian Process Port-Hamiltonian Systems: Bayesian Learning with Physics Prior. CDC, 2022

GP-PHS



Thomas Beckers

GP-PHS training

GP prior on Hamiltonian

$$\dot{x} = \boxed{[J(x) - R(x)]\frac{\partial H}{\partial x}} + G(x)u \quad \longrightarrow \quad \dot{x} \sim GP(\hat{G}(x)u, k_{phs}(x, x'))$$

Linear operator

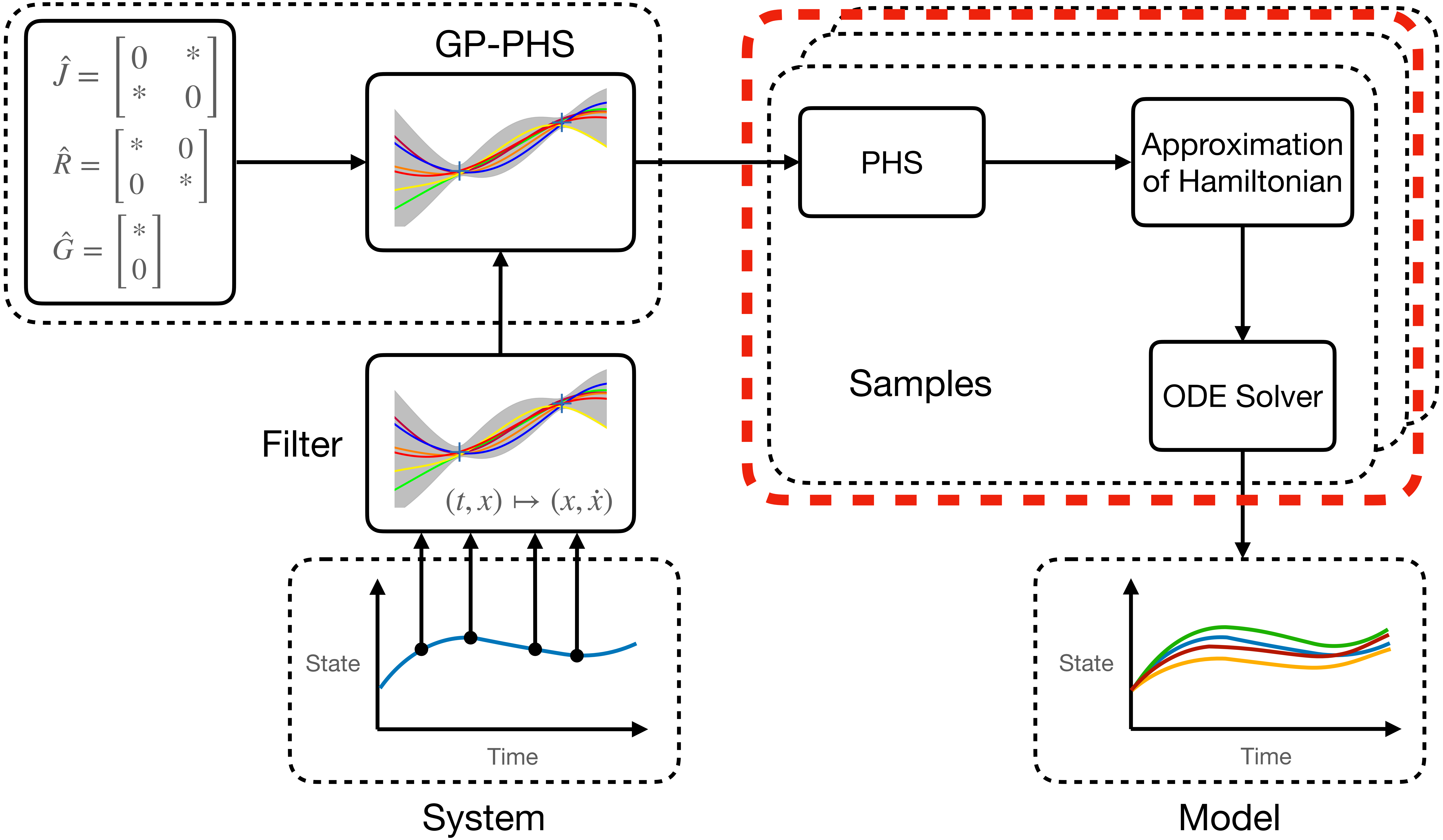
Port-Hamiltonian Kernel

$$k_{phs}(x, x') = \sigma_f^2 [\hat{J}(x) - \hat{R}(x)] \left[\frac{\partial}{\partial x \partial x'} \underbrace{\exp(-\|x - x'\|_\Lambda^2)}_{\text{Squared exp. kernel}} \right] \overbrace{[\hat{J}(x') - \hat{R}(x')]^\top}^{\text{Parametrized estimates}}$$

Hyperparameters $\quad \varphi = [\varphi_J, \varphi_R, \varphi_G, \Lambda, \sigma_f] \in \mathscr{R}^{n_J + n_R + n_G + n + 1}$

Minimizing NLML $\quad -\log p(\dot{X} \mid \varphi, X) \sim \dot{X}^\top [\underbrace{K_{phs} + \Delta}_{\text{Uncertainty of filter}}]^{-1} \dot{X} + \log |K_{phs} + \Delta|$

GP-PHS



Thomas Beckers

GP-PHS sampling

PHS

$$\begin{bmatrix} \dot{X} \\ \hat{f}(x^*) \end{bmatrix} = \mathcal{N}\left(0, \begin{bmatrix} K_{phs} & k_{phs}(X, x^*) \\ k_{phs}(X, x^*)^\top & k_{phs}(x^*, x^*) \end{bmatrix}\right)$$

$$\dot{x} = \hat{f}(x, \omega)$$

Not callable
(computational expensive)

Approximation of Hamiltonian

$$\begin{bmatrix} \dot{X} \\ \hat{H}(x^*) \end{bmatrix} = \mathcal{N}\left(0, \begin{bmatrix} K_{phs} & k_{\dot{x}H}(X, x^*) \\ k_{\dot{x}H}(X, x^*)^\top & k_{HH}(x^*, x^*) \end{bmatrix}\right)$$

$$\hat{H}(x^*, \omega) \xrightarrow{\text{Approx.}} \hat{H}^*(x^*)$$

ODE Solver

$$\dot{x} = [\hat{J}(x) - \hat{R}(x)]\frac{\partial \hat{H}^*}{\partial x} + \hat{G}(x)u \xrightarrow{\text{Solver}} x(t)$$

Approximation of H: Callable and structure preserving

Theoretical results

PHS

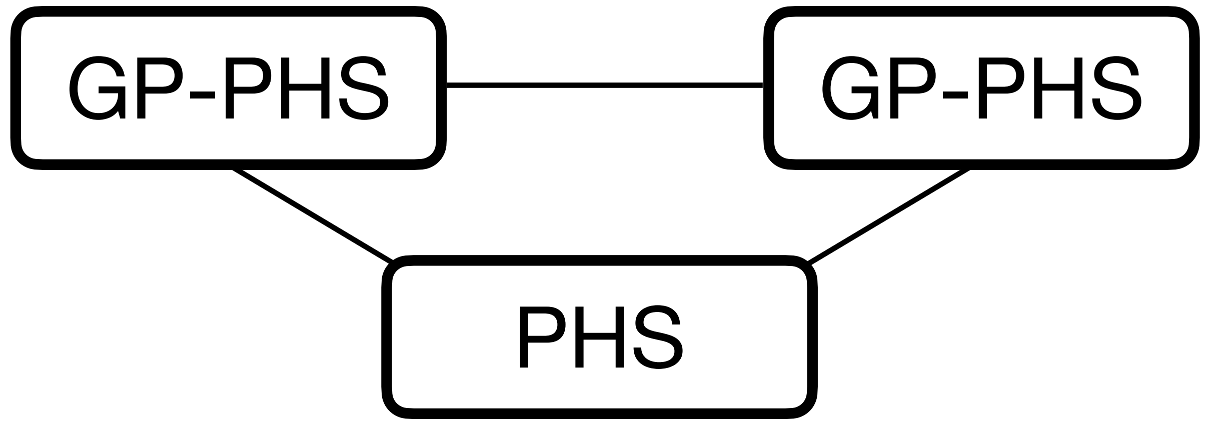
Consider a GP-PHS trained on the dataset $\mathscr{D}_x$. Let $\hat{H}^*: \mathscr{R}^n \to \mathscr{R}$ be a smooth and bounded function approximator of $\hat{H}(\,\cdot\,, \omega)$. Then,

$$\dot{x} = [\hat{J}(x) - \hat{R}(x)]\, \nabla_x \hat{H}^*(x) + \hat{G}(x)u$$
$$y = \hat{G}(x)^\top \nabla_x \hat{H}^*(x)$$

describes a Port-Hamiltonian system that is passive with respect to the supply rate $u^\top y$.

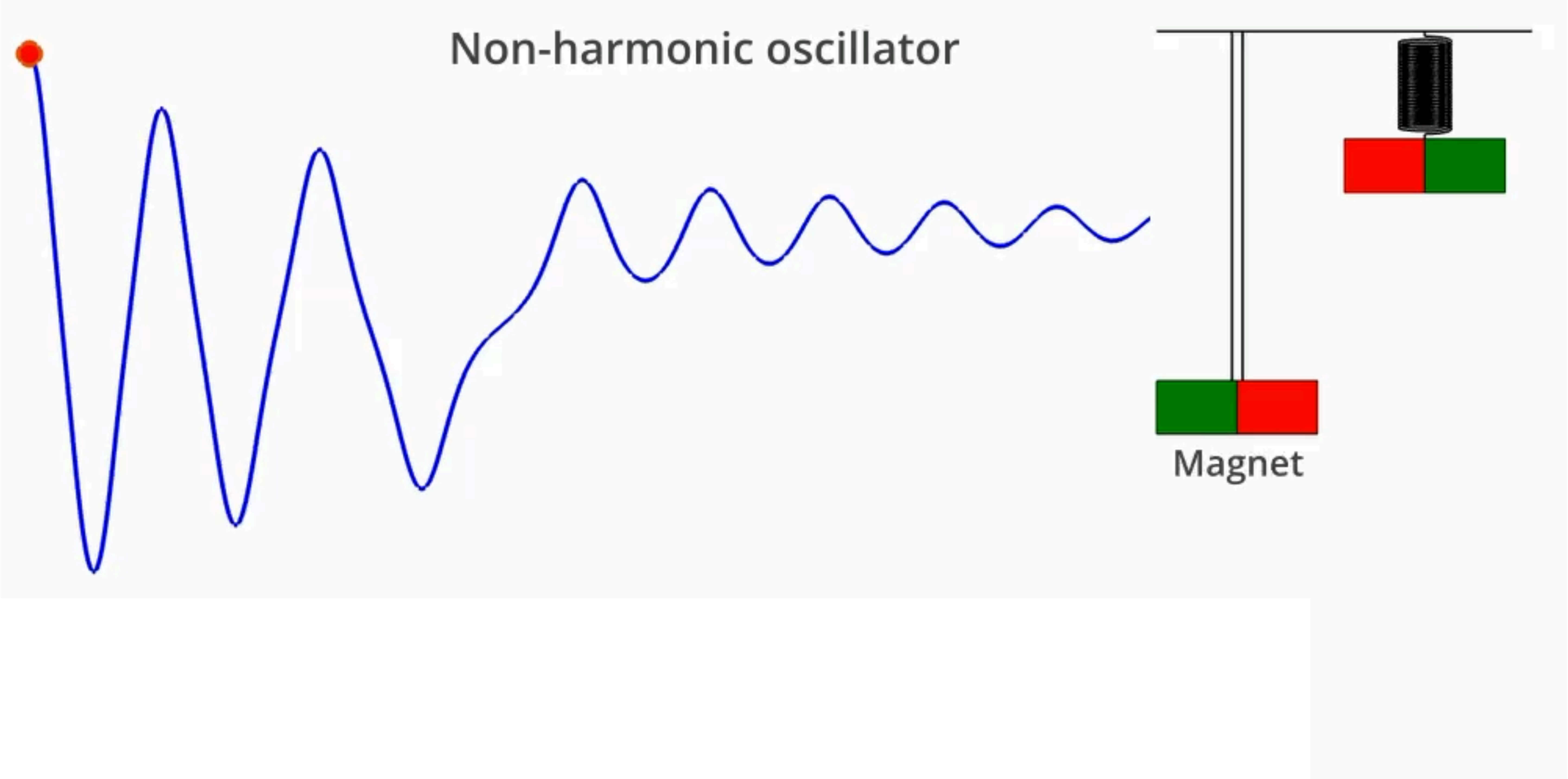
Interconnection property

Consider two GP-PHS. Then, the feedback interconnection $u_1 = -y_2$ and $u_2 = y_1$ yields a GP-PHS.

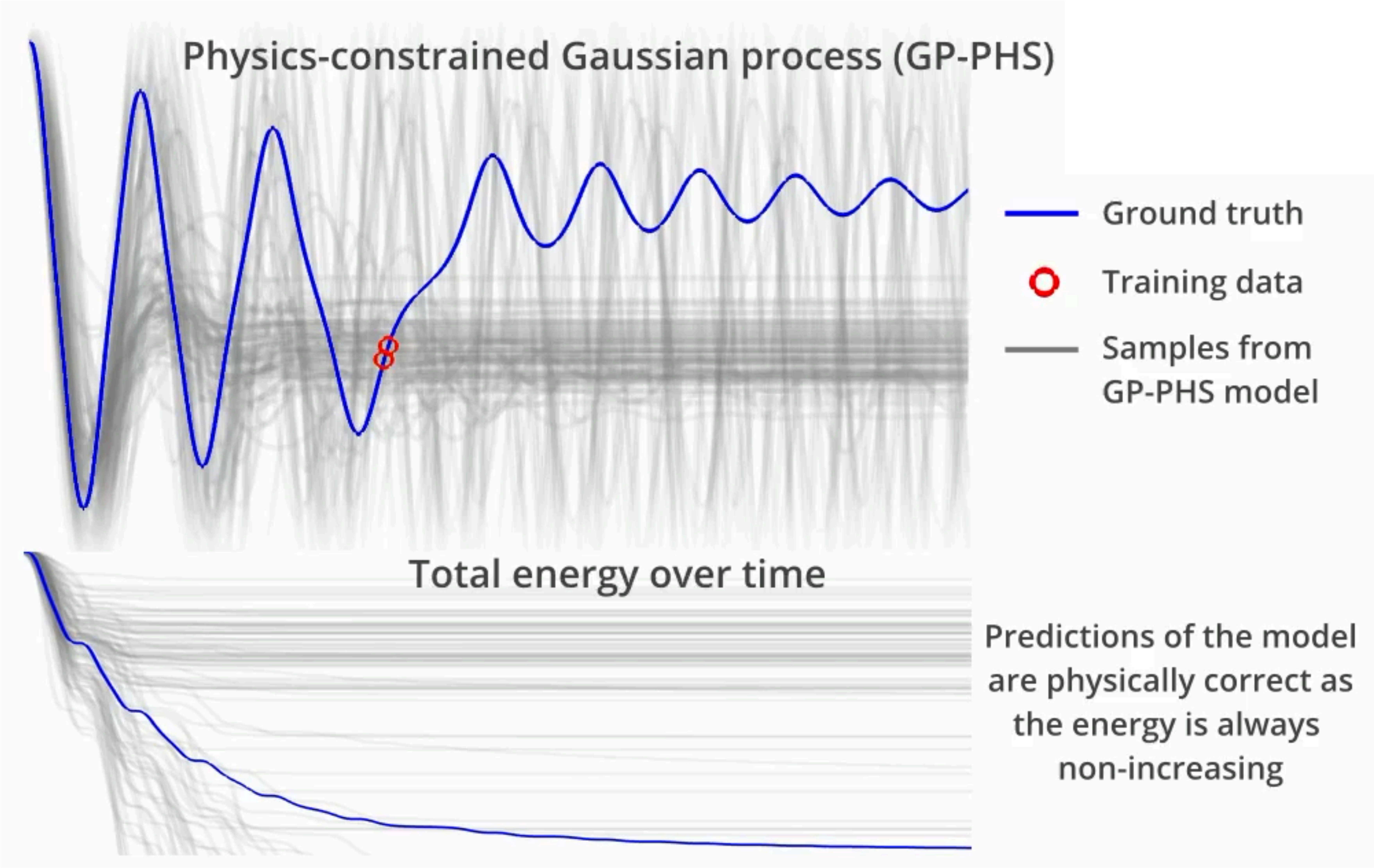


$\Rightarrow$ allows composable learning

Example



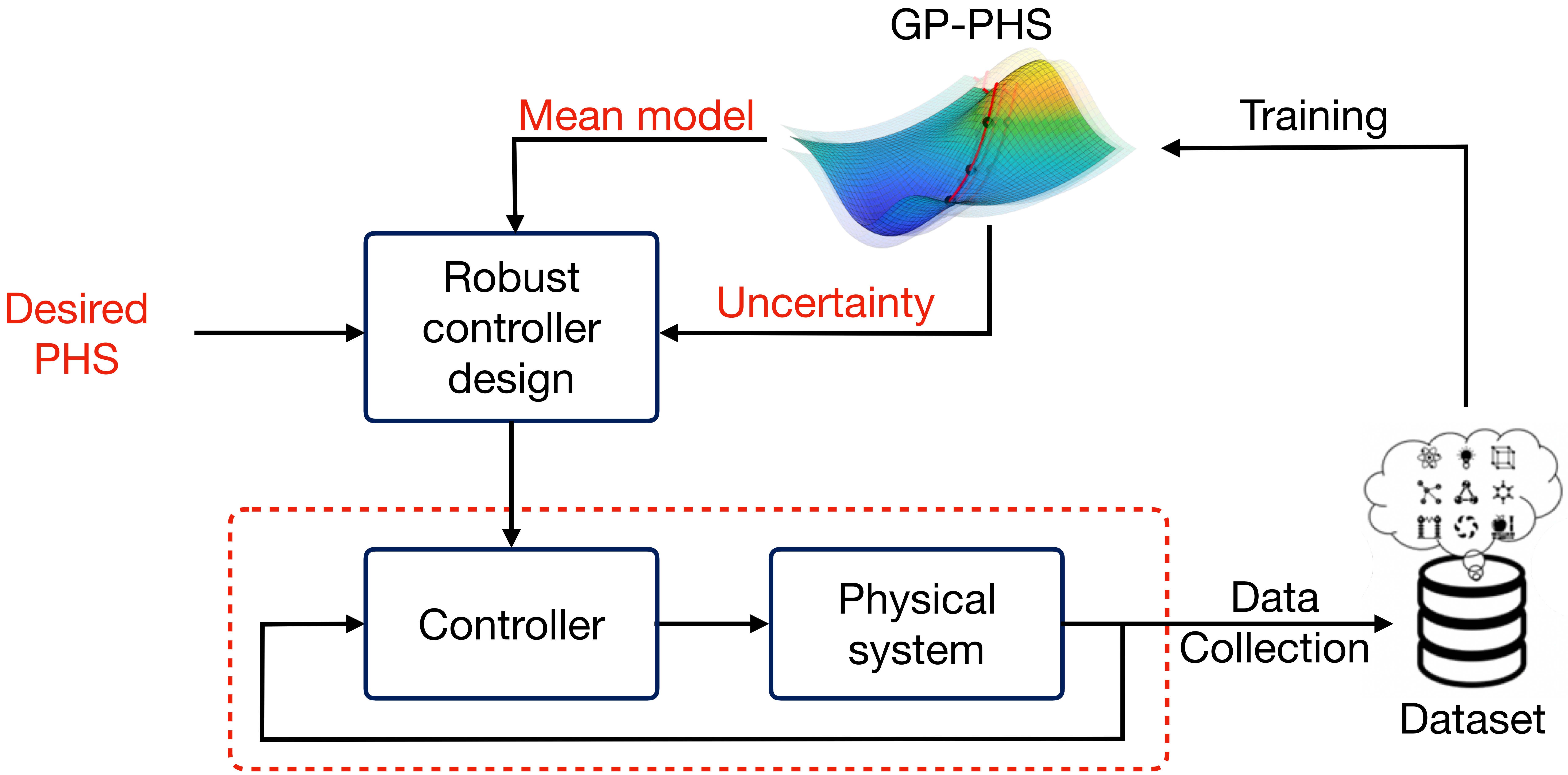
Example



Physics-constrained Gaussian process (GP-PHS)
Ground truth
Training data
Samples from GP-PHS model
Total energy over time
Predictions of the model are physically correct as the energy is always non-increasing

Control

Design of a stabilizing controller for the physical system with unknown dynamics



T. Beckers, Data-driven Bayesian Control of Port-Hamiltonian Systems. CDC, 2023

Controller design

Closed-loop dynamics = Desired PHS

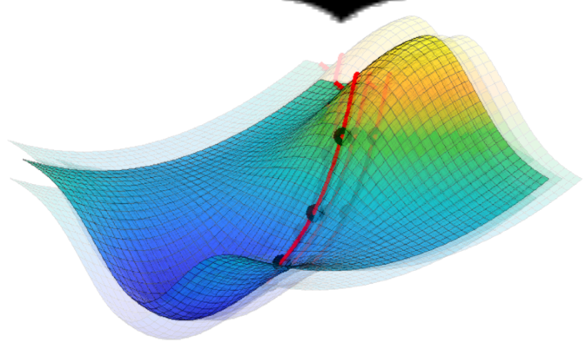
$$\dot{x} = [J_d(x) - R_d(x)]\nabla H_d$$

If satisfied

IDA-PBC PDE

Physical system | Desired PHS

$$G^{\perp}(x)\underbrace{[J(x) - R(x)]\nabla H} = G^{\perp}(x)[J_d(x) - R_d(x)]\nabla H_d$$



GP-PHS model

Mean $\mu(x)$

Uncertainty bound $\eta(x)$

Stabilization criterion

$$G^{\perp}(x)\mu(x) = G^{\perp}(x)[J_d(x) - R_d(x)]\nabla H_d$$

Robustness criterion

$$[\nabla H_d]^{\top}\eta(x) \leq [\nabla H_d]^{\top}R_d(x)\nabla H_d$$

Closed-loop system with desired equilibrium $\xrightarrow[\text{less uncertainty}]{\text{more data}}$ Closed-loop system with desired PHS dynamics

Control approach

Find $J_d(x), R_d(x), H_d(x)$ for all $x \in \Omega$ such that

- Desired properties — Performance, equilibrium x_d
- Stabilization criterion — $G^{\perp}(x)\mu(x) = G^{\perp}(x)[J_d(x) - R_d(x)]\nabla H_d$
- Robustness criterion — $[\nabla H_d]^{\top}\eta(x) \leq [\nabla H_d]^{\top}R_d(x)\nabla H_d$

$$\dot{H}_d = [\nabla H_d]^{\top}(J_d(x) - R_d(x))\nabla H_d + [\nabla H_d]^{\top}\eta(x)$$
$$= \boxed{-[\nabla H_d]^{\top}R_d(x)\nabla H_d + [\nabla H_d]^{\top}\eta(x)}$$

How to solve?

- "Classic" approaches: Fixing J_d, R_d or fixing H_d
- Physics-informed NN

Theoretical results

Stability guarantee

Let the Stabilization criterion and the Robustness criterion be satisfied. Then, the control law

$$u(x) = [\hat{G}^\top(x)\hat{G}(x)]^{-1}\hat{G}^\top(x)(\underbrace{[J_d(x) - R_d(x)]\nabla H_d}_{\text{Desired PHS}} - \underbrace{\mu\left(x, \mathscr{D}\right)}_{\text{Nominal model}}$$

leads to a closed-loop system with **a stable equilibrium** x_d **with probability** $(1 - p)$.

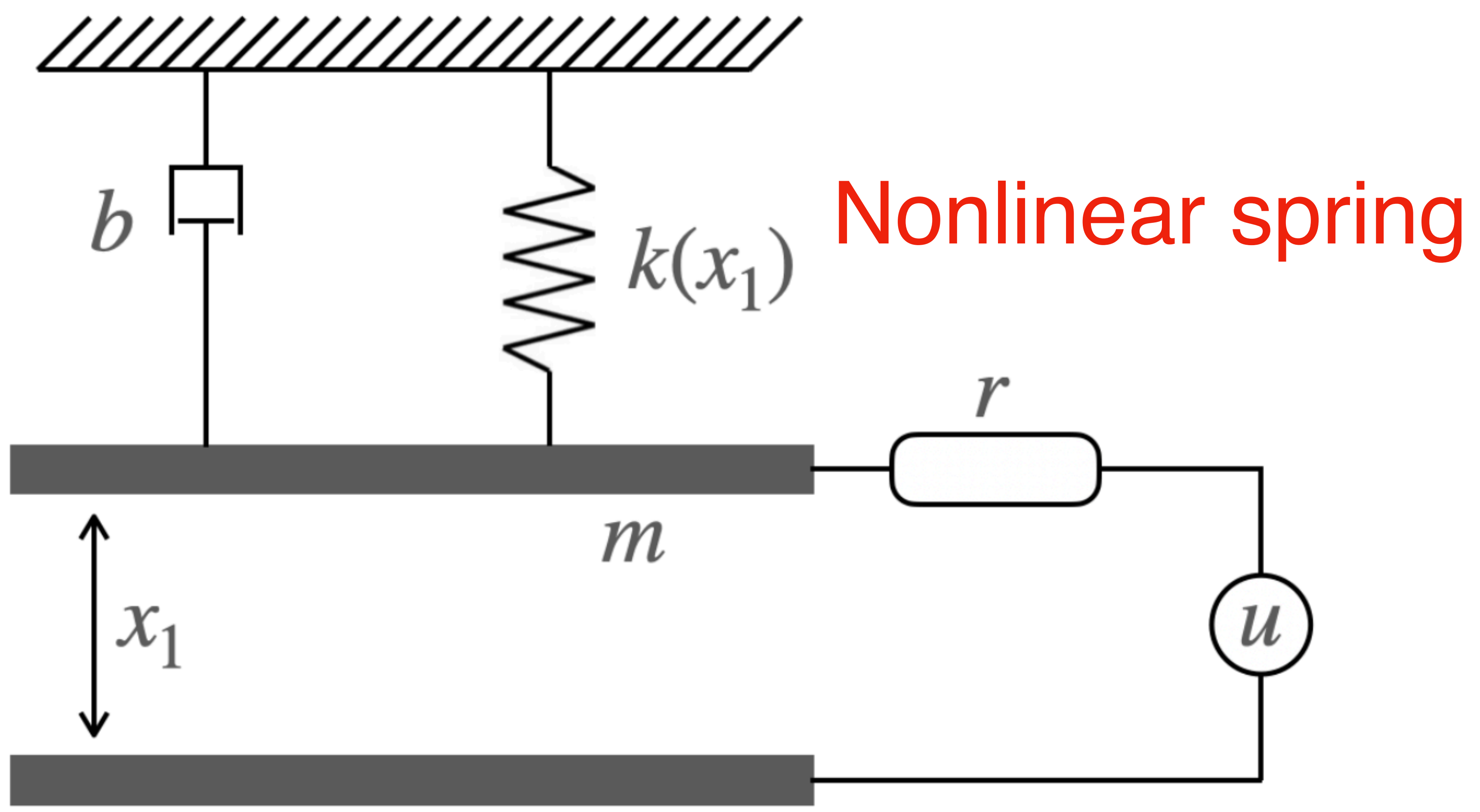
Asymptotic stability

The equilibrium x_d will be **asymptotically stable** on Ω with probability $(1 - p)$ if

- x_d is an isolated minimum of H_d and
- the largest invariant set in $\{x \in \Omega \,|\, [\nabla H_d]^\top R_d(x) \nabla H_d = 0\}$ equals x_d

Simulation

Electrostatic micro-actuator

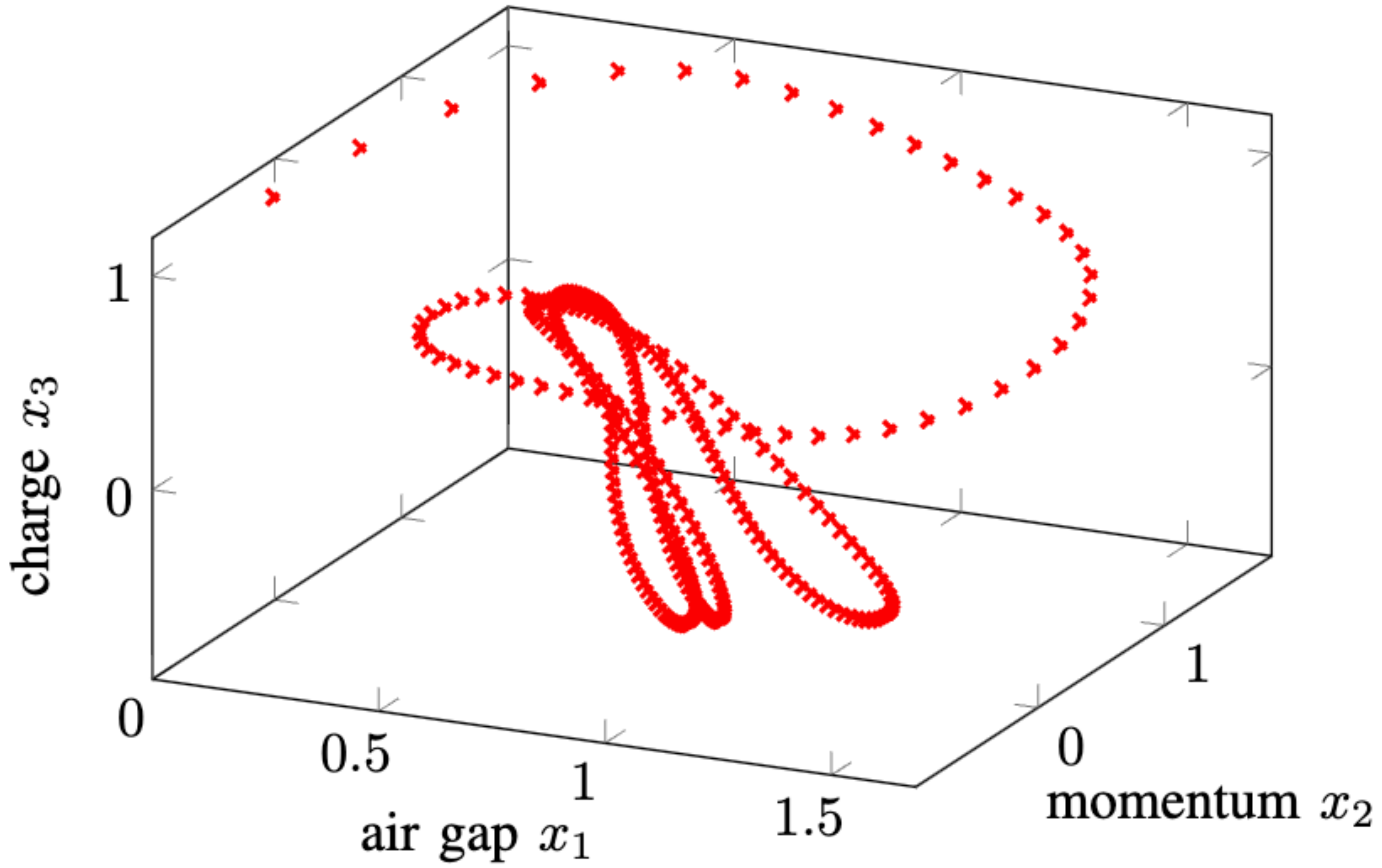


$$H(x) = \frac{1}{4}10(x_1 - x^*)^4 + \frac{1}{2m}x_2^2 + \frac{x_1}{2A\epsilon}x_3^2$$

Ground truth Hamiltonian

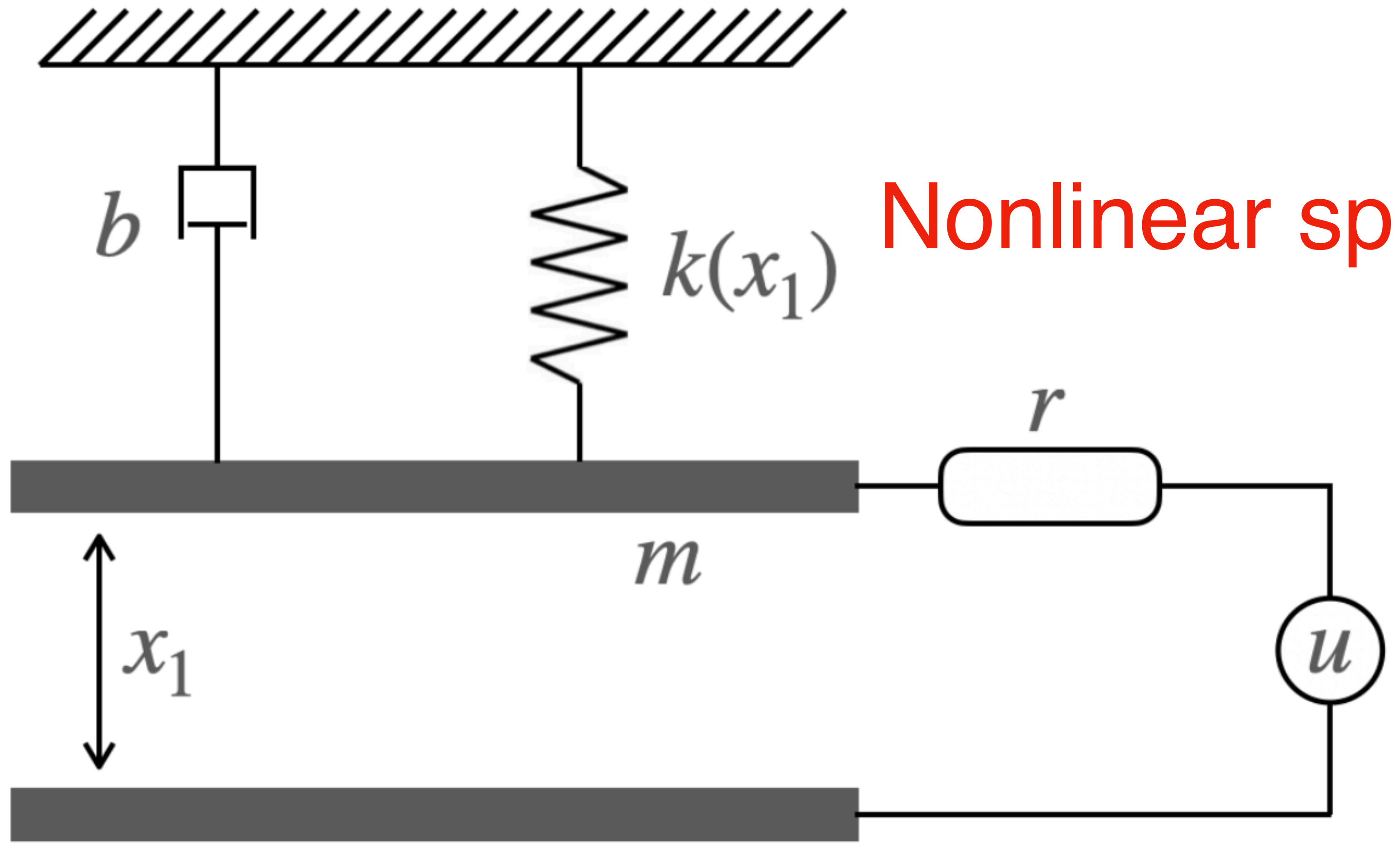
Collecting training data

- sinusoidal input
- 300 state observations



Simulation

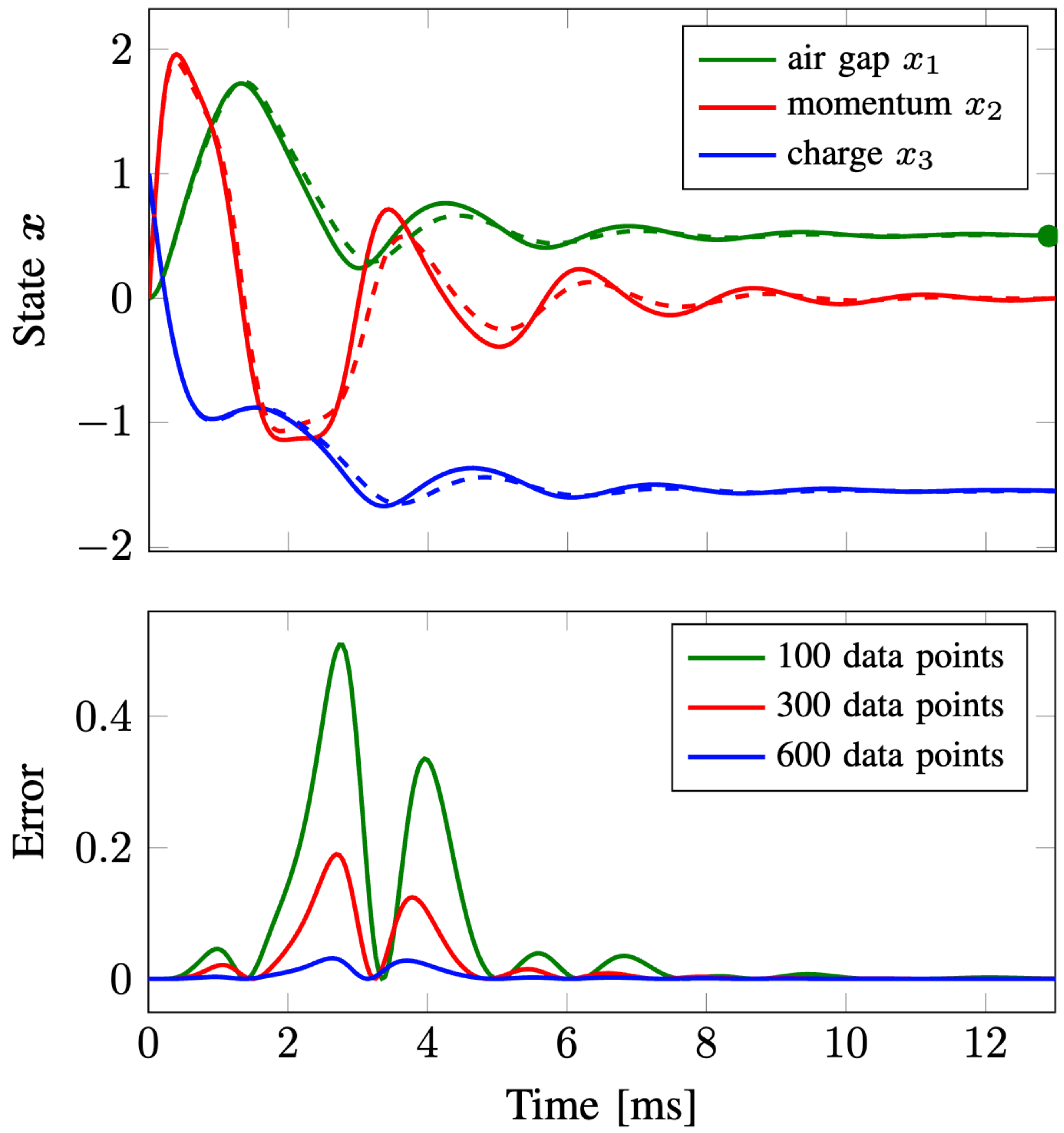
Electrostatic micro-actuator



Nonlinear spring

$$H(x) = \frac{1}{4}10(x_1 - x^*)^4 + \frac{1}{2m}x_2^2 + \frac{x_1}{2A\epsilon}x_3^2$$

Ground truth Hamiltonian



GP-PHS based controller: State converges to desired equilibrium

Extension to PDEs

	ODE model	PDE model
Dynamics	$\dot{x}(t) = [J(x) - R(x)]\frac{\partial H}{\partial x} + G(x)u$	$\frac{\partial x}{\partial t}(t,z) = [J - R]\delta_x \mathscr{H} + Gu$ $w = B_{\mathscr{Z}}(\delta_x \mathscr{H}, u)$
Energy	$H(x)$	$\mathscr{H}(x) = \int_{\mathscr{Z}} H(z,x) dV$
Interconnection	Skew-symmetric $J(x)$	Matrix differential operator J
Training data	$\mathscr{D} = \{t_i, x(t_i), u(t_i)\}_{i=0}^{i=N_t-1}$	$\mathscr{D} = \{t_i, z_j, x(t_i, z_j), u(t_i)\}_{i=j=0}^{i=N_t-1, j=N_z-1}$
Output	Samples of state over time	Samples of state over time and spatial

K. Tan, P. Li, and T. Beckers, Physics-Constrained Learning for PDE Systems with Uncertainty Quantified Port-Hamiltonian Models. L4DC, 2024

GP-dPHS example

Vibration of a string with nonlinear stress-strain curve

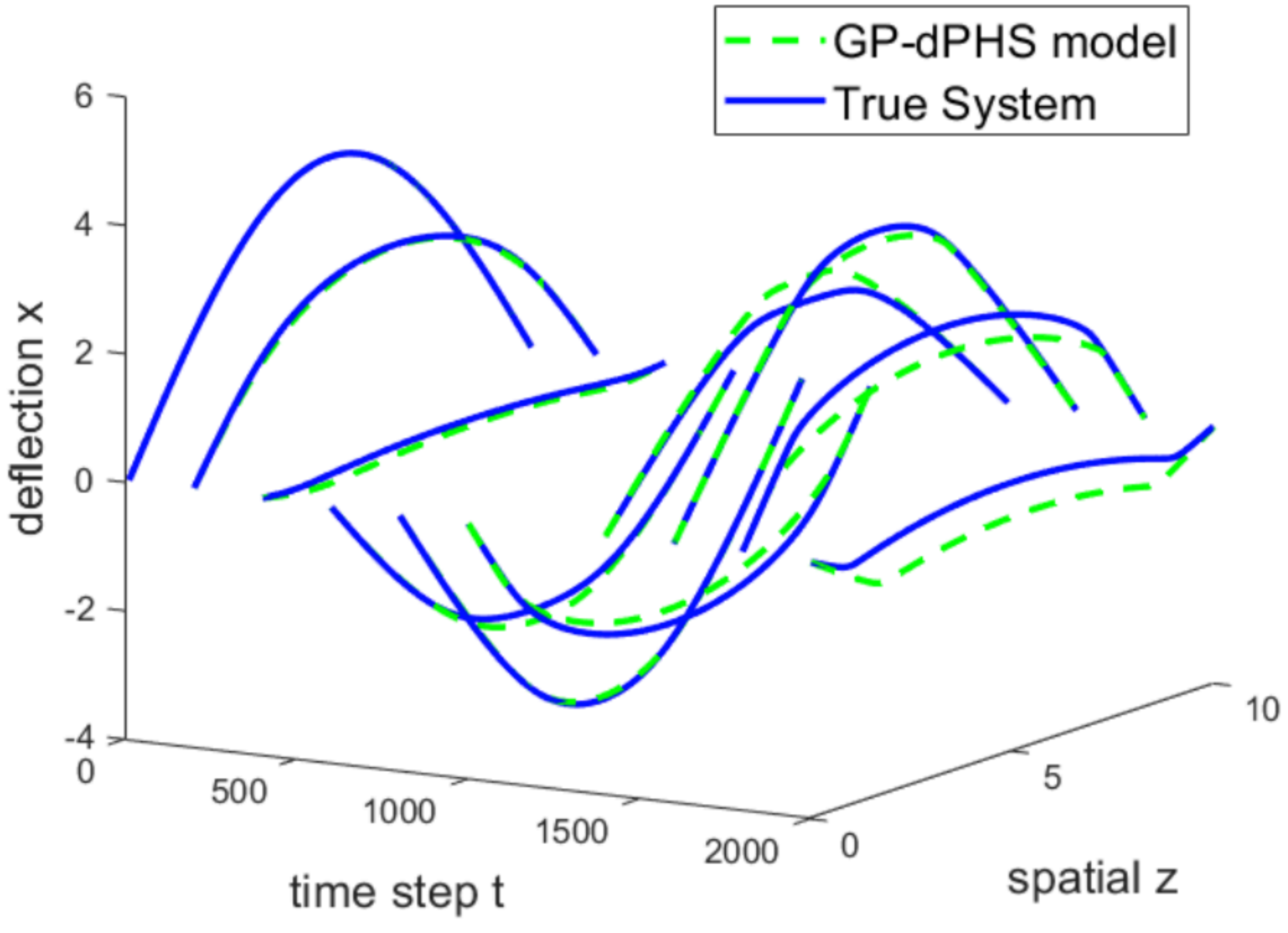
$$\frac{\partial^2 x}{\partial t^2} - f\left(\frac{\partial x}{\partial z}\right)\frac{\partial^2 x}{\partial z^2} = 0,$$

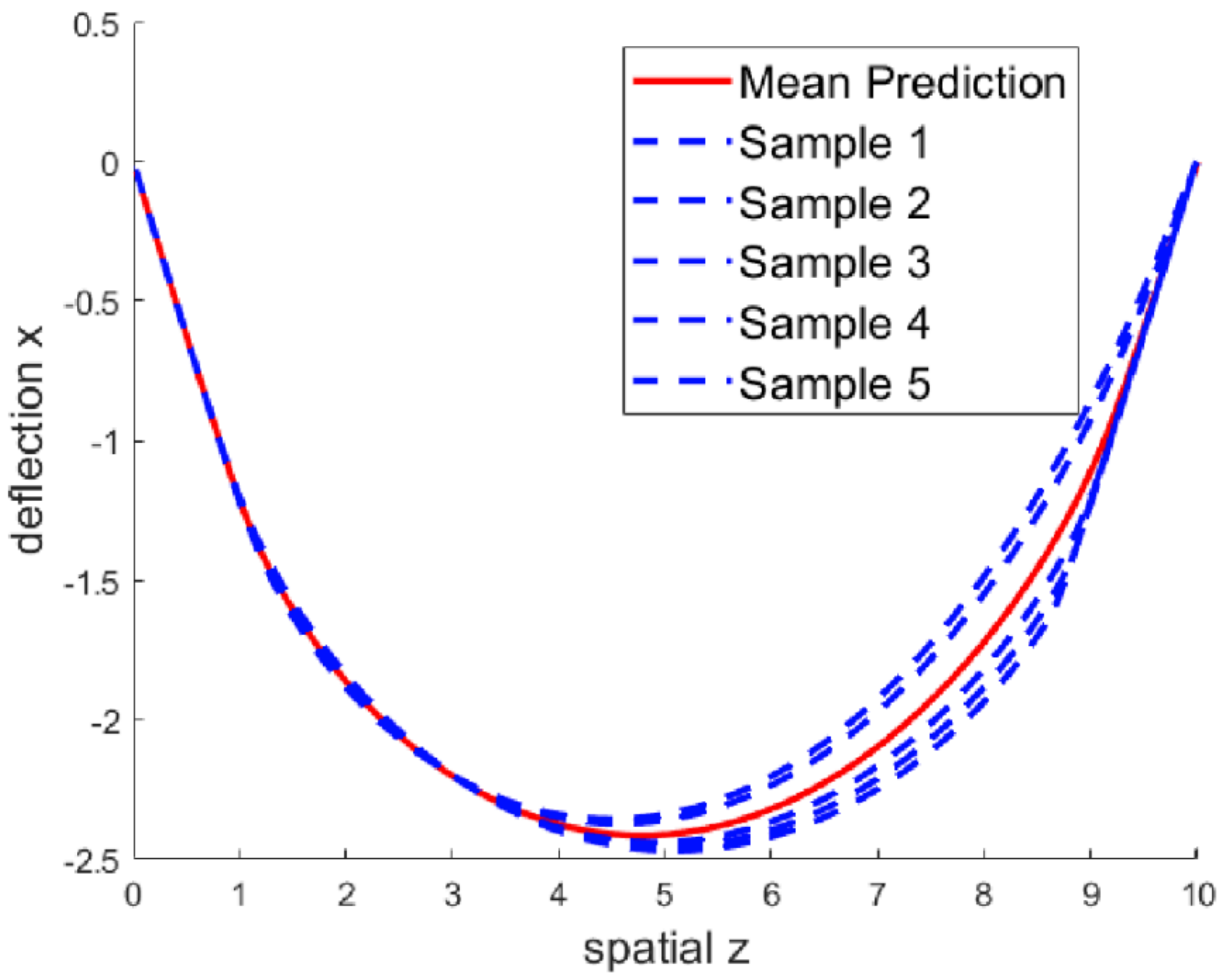
In PHS form

$$\frac{\partial x}{\partial t}$$

$$\frac{\partial}{\partial t}\begin{bmatrix} p(t,z) \\ q(t,z) \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial z} & 0 \end{bmatrix}}_{J-R} \delta_{pq}\mathcal{H}(p,q)$$

$$\frac{\partial x}{\partial z}$$





K. Tan, P. Li, and T. Beckers, Physics-Constrained Learning for PDE Systems with Uncertainty Quantified Port-Hamiltonian Models. L4DC, 2024

Composition

GP-dPHS

GP-dPHS

$$\frac{\partial}{\partial t}\begin{bmatrix} p(t,z) \\ q(t,z) \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & \frac{\partial}{\partial z} \\ \frac{\partial}{\partial z} & 0 \end{bmatrix}}_{J-R} \delta_{pq}\mathscr{H}(p,q)$$

Vibration of a string with
nonlinear stress-strain curve

Energy flow

$\longrightarrow$

$\longleftarrow$

PHS

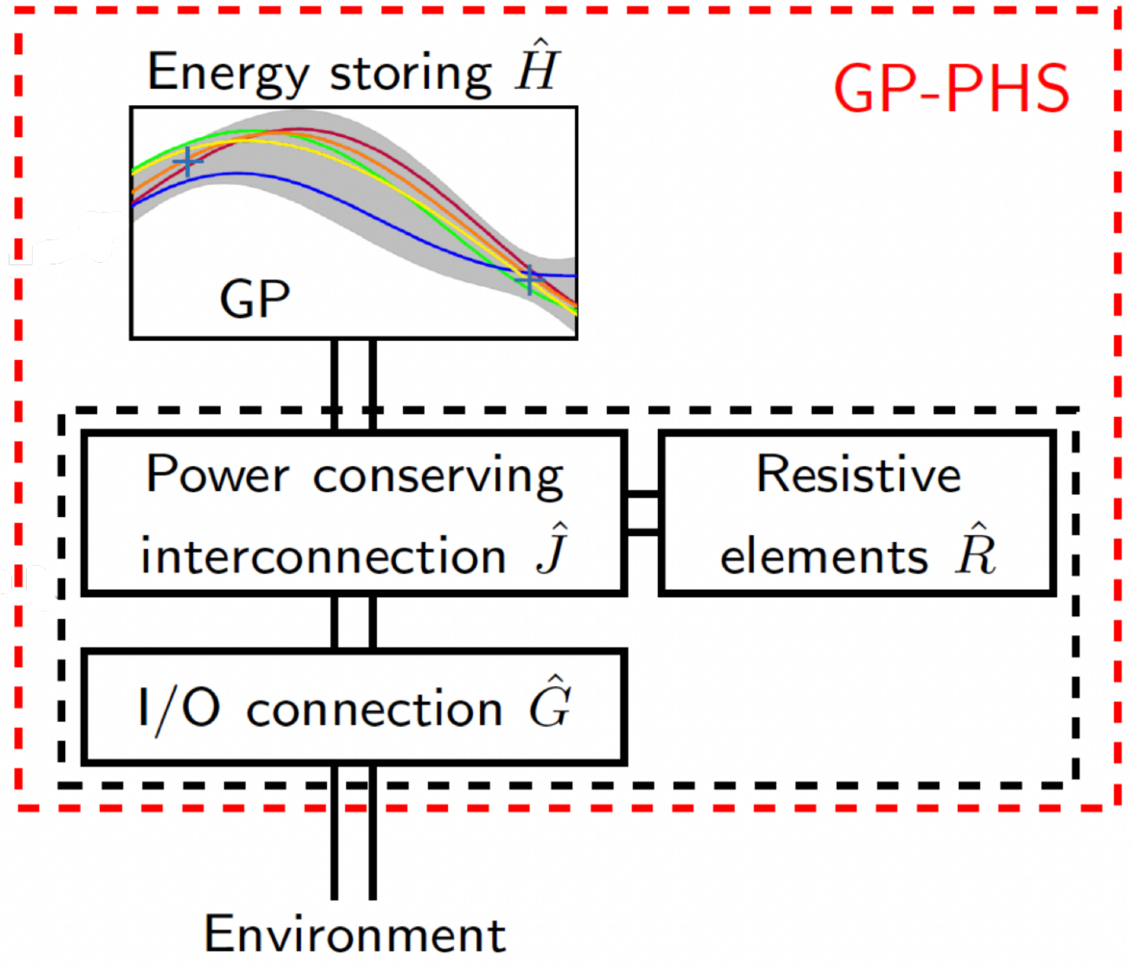
$$\begin{bmatrix} \dot{p}(t) \\ \dot{q}(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ -1 & -r \end{bmatrix}}_{J-R} \nabla H(p,q)$$

Mass-spring-damper system

Conclusion

Gaussian Process Port-Hamiltonian Systems

- Model is physically consistent and allows uncertainty quantification
- Preserve the interconnection property and passivity characteristic
- Robust control of physical system based on GP-PHS
- Extension to switching systems



Challenges

- Real-time uncertainty quantification for complex physics-informed models
- Uncertainty of physics?

www.tbeckers.com **thomas.beckers@vanderbilt.edu**