Learning Probabilistic Reduced-Order Models

Towards Scientific Machine Learning and Digital Twinning

Dr. Mengwu Guo

Assistant Professor
Mathematics of Imaging & AI

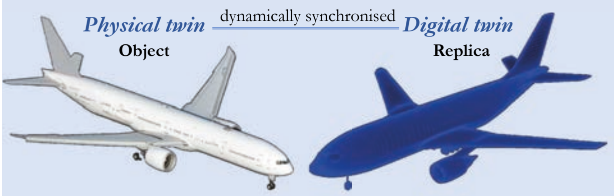
University of Twente

NHR4CES Community Workshop 2023 · February 28, 2023

Digital twin technology

Synthesis of ***physics-based modeling*** & ***machine learning*** with real-world observational ***data***

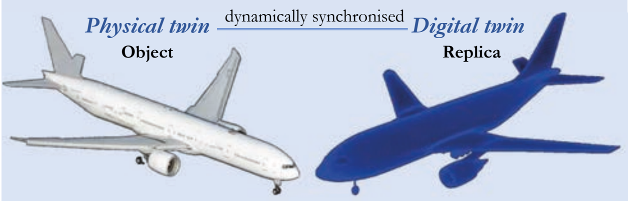
"A Digital Twin is a set of virtual information constructs that mimics the structure, context, and behavior of an individual/unique physical asset, is dynamically updated with data from its physical twin throughout its lifecycle, and informs decisions that realize value"
– *AIAA Institute Position Paper*, 2020



China: Top 10 frontier scientific questions
Singapore: Virtual Singapore
U.S.: DOE - AI and Decision Support for Complex Systems
Europe: Many initiatives in Horizon 2020 & EU

Foundation – **multi-query & real-time simulations**

Reduced-order models (ROMs) for digital twinning



$\dot{\mathbf{u}}(t;\mu) = \mathbf{h}(\mathbf{u}(t;\mu);\mu)$

$\mathbf{u}(t_0;\mu)$ → $\mathbf{u}(t;\mu) \approx \hat{\mathbf{u}}(t;\mu)$

$\mathcal{G}_e$ ***Reduce – Represent – Recover*** $\mathcal{G}_d$

$\mathbf{q}(t_0;\mu)$ → $\mathbf{q}(t;\mu)$

$\dot{\mathbf{q}}(t;\mu) = \tilde{\mathbf{h}}(\mathbf{q}(t;\mu);\mu)$

$\tilde{\mathbf{h}}(\cdot;\mu) = [\partial_{\mathbf{q}}\mathcal{G}_d(\cdot)]^\dagger\, \mathbf{h}(\mathcal{G}_d(\cdot);\mu)$

Nonlinear, time-dependent PDEs
in ***multi-query, real-time*** contexts

Full-order solvers lead to prohibitive demands!

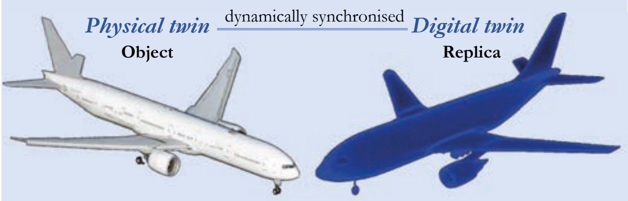
Projection-based intrusive ROMs

Access to full-order solvers
+
Problem-dependent projection formulation

$\Downarrow$

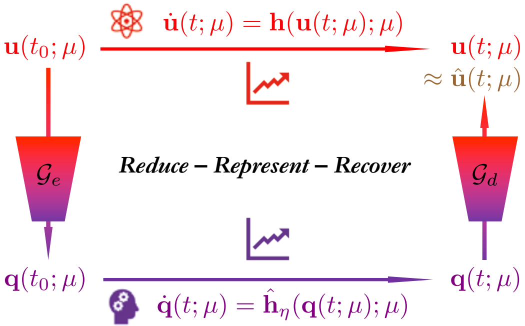
Limited industrial relevance

Reduced-order models (ROMs) for digital twinning



Nonlinear, time-dependent PDEs in ***multi-query, real-time*** contexts

Full-order solvers lead to prohibitive demands!



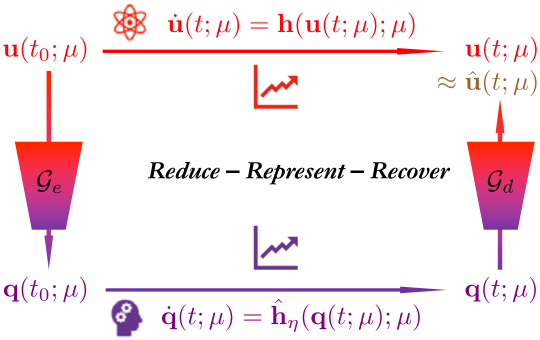
Data-driven non-intrusive ROMs

Physically-consistent,
gray-box dynamics learning

Reduced-order models (ROMs) for digital twinning

"As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality."

Albert Einstein



Data-driven non-intrusive ROMs

Physically-consistent **gray-box** dynamics learning

Big noisy data + *Intrinsic modeling uncertainties*

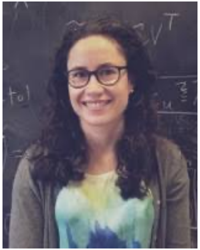
$\Downarrow$

Probabilistic Machine Learning

Gaussian process surrogate modeling

Prof. Jan Hesthaven

Dr. Zhenying Zhang



Mariella Kast

EPFL

AKSELOS

- **MG**, J. S. Hesthaven. *Comput. Methods Appl. Mech. Eng.*, 341: 807-826, 2018.
- **MG**, J. S. Hesthaven. *Comput. Methods Appl. Mech. Eng.*, 345: 75–99, 2019.
- Z. Zhang, **MG**, J. S. Hesthaven. *Comput. Methods Appl. Mech. Eng.*, 353: 491-515, 2019.
- M. Kast, **MG**, J. S. Hesthaven. *Comput. Methods Appl. Mech. Eng.*, 364: Article 112947, 2020.

Reduced-order surrogate modeling

Non-intrusively find underlying structure from snapshot data $\longrightarrow$ Machine Learning

Reduced-order solutions formulated as low-rank approximation:

$$u(x,t;\boldsymbol{\mu}) \approx \sum_{l=1}^{r} \hat{q}_l(t;\boldsymbol{\mu})\psi_l(x) \quad \text{or} \quad \mathbf{u}_h(t;\boldsymbol{\mu}) \approx \sum_{l=1}^{r} \hat{q}_l(t;\boldsymbol{\mu})\mathbf{v}_l \in \mathbb{R}^{N_h} \text{ with } r \ll N_h$$

Offline stage:

- Reduced basis vectors $\{\psi_1, \cdots, \psi_r\}$ or $\mathbf{V} = [\, \mathbf{v}_1 \,|\cdots|\, \mathbf{v}_r \,]$
  - Proper orthogonal decomposition acting on full-order snapshots [*unsupervised learning*]
- Expansion coefficients $(t,\boldsymbol{\mu}) \mapsto \hat{q}_l = \langle u, \psi_l \rangle_\Omega$ or $\mathbf{v}_l^{\mathrm{T}}\mathbf{u}_h$
  - ***Gaussian process surrogate*** models from projected snapshots [*supervised learning*]

Online stage:

- Regression output evaluation + linear combination **Simulation free**

Offline improvements: multi-fidelity data fusion

Gaussian process for regression

Prior Gaussian process (GP) with noise

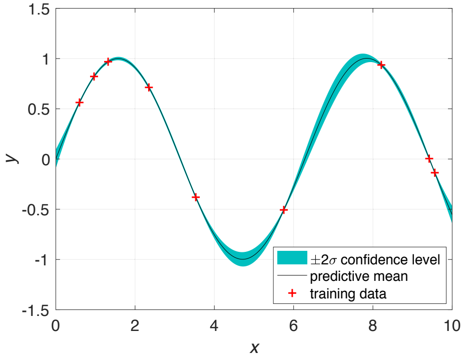
$$q(\boldsymbol{\mu}) \sim \mathcal{GP}(0, \kappa(\boldsymbol{\mu}, \boldsymbol{\mu}')), \quad y = q(\boldsymbol{\mu}) + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \sigma_y^2).$$

Posterior GP conditioned on training data

$$\begin{aligned} q^*(\boldsymbol{\mu}) | \boldsymbol{\mu}, \mathbf{X}, \mathbf{y} &\sim \mathcal{GP}(m^*(\boldsymbol{\mu}), c^*(\boldsymbol{\mu}, \boldsymbol{\mu}')), \\ m^*(\boldsymbol{\mu}) &= \kappa(\boldsymbol{\mu}, \mathbf{X}) \mathbf{K}_y^{-1} \mathbf{y}, \;\; \mathbf{K}_y = \kappa(\mathbf{X}, \mathbf{X}) + \sigma_y^2 \mathbf{I}_M, \\ c^*(\boldsymbol{\mu}, \boldsymbol{\mu}') &= \kappa(\boldsymbol{\mu}, \boldsymbol{\mu}') - \kappa(\boldsymbol{\mu}, \mathbf{X}) \mathbf{K}_y^{-1} \kappa(\mathbf{X}, \boldsymbol{\mu}'). \end{aligned}$$

Hyperparameters $\boldsymbol{\theta}$ determined by maximizing the marginal likelihood

$$\log p(\mathbf{y}|\mathbf{X}, \boldsymbol{\theta}) = -\frac{1}{2} \mathbf{y}^{\mathrm{T}} \mathbf{K}_y^{-1}(\boldsymbol{\theta}) \mathbf{y} - \frac{1}{2} \log |\mathbf{K}_y(\boldsymbol{\theta})| - \frac{M}{2} \log(2\pi).$$



The posterior mean of GPR is equivalent to a kernel ridge regression.

- **Interpolation**: $f(\boldsymbol{\mu}) = \kappa(\boldsymbol{\mu}, \mathbf{X}) \boldsymbol{\alpha}$; **RKHS norm regularization**: $\|f\|_{\mathcal{H}}^2 = \boldsymbol{\alpha}^{\mathrm{T}} \kappa(\mathbf{X}, \mathbf{X}) \boldsymbol{\alpha}$
- **Loss function**: $J(f) = \|\mathbf{y} - f(\mathbf{X})\|_2^2 + \sigma_y^2 \|f\|_{\mathcal{H}}^2$; $\arg\min_{\boldsymbol{\alpha} \in \mathbb{R}^M} J(f) = \mathbf{K}_y^{-1} \mathbf{y}$
- **The existence of noise term σ_y^2 acts as a regularization.**

Hybrid modeling for large-scale systems

Based on a predefined linear-nonlinear domain decomposition

(1) Sensitivity analysis and GPR

$$\mathbb{U}_{\Gamma}^{\mathtt{rb}}(\mu) \approx \mathbb{U}_{\Gamma}^{\mathtt{rb}}(\tilde{\mu}) \sim \mathcal{GP}$$

(2) Solve a linear reduced basis problem

$$\mathbb{U}_{\mathrm{lin}}^{\mathtt{rb}}(\mu) = [\mathbb{A}_{\mathrm{lin}}^{\mathtt{rb}}(\mu)]^{-1}[\mathbb{F}_{\mathrm{lin}}^{\mathtt{rb}}(\mu) - \mathbb{A}_{\mathrm{lin},\Gamma}^{\mathtt{rb}}(\mu)\mathbb{U}_{\Gamma}^{\mathtt{rb}}(\mu)]$$

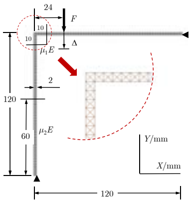
$$\begin{bmatrix} \mathbb{A}_{\mathrm{lin}} & \mathbb{A}_{\mathrm{lin},\Gamma} & \mathbf{0} \\ \mathbb{A}_{\Gamma,\mathrm{lin}} & \mathbb{A}_{\Gamma} + \mathbb{G}_{\Gamma} & \mathbb{G}_{\Gamma,\mathrm{nlin}} \\ \mathbf{0} & \mathbb{G}_{\mathrm{nlin},\Gamma} & \mathbb{G}_{\mathrm{nlin}} \end{bmatrix} \begin{bmatrix} \mathbb{U}_{\mathrm{lin}} \\ \mathbb{U}_{\Gamma} \\ \mathbb{U}_{\mathrm{nlin}} \end{bmatrix} = \begin{bmatrix} \mathbb{F}_{\mathrm{lin}} \\ \mathbb{F}_{\Gamma} \\ \mathbb{F}_{\mathrm{nlin}} \end{bmatrix}$$

(3) GPR surrogate

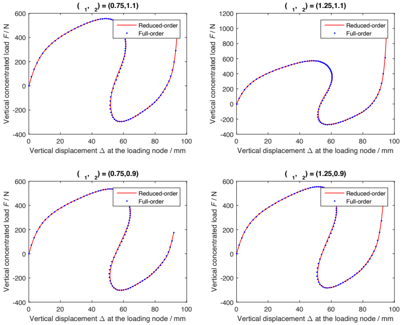
$$\mathbb{U}_{\mathrm{nlin}}^{\mathtt{rb}}(\mu) = \mathbb{U}_{\mathrm{nlin}}^{\mathtt{rb}}(\mathbb{U}_{\Gamma}^{\mathtt{rb}}, \mu_{\mathrm{nlin}}) \sim \mathcal{GP}$$

- **Sensitivity analysis** (Modified DGSM) of solution field $\mathbf{w}(\mu)$, $\hat{\nu}_i = \frac{\nu_i}{\sum_{i=1}^{d} \nu_i}$, $\nu_i = \mathbb{E}\left[\frac{\|\mathbf{w}(\mu+\delta\mu_i)\|_2}{\delta\mu_i}\right]^2$
- **Component-wise reduced order model** [Huynh et al, 2013]: static condensation + reduced basis

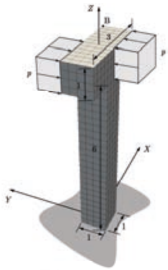
Applications in computational mechanics



	Time / s
One run F-O	2.28E+01
GPRs	4.88E+00
One run R-O	2.23E-02



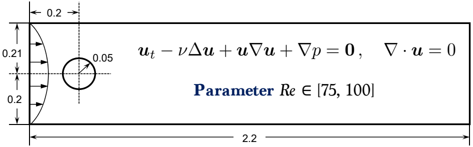
Reduce order model predictions with RB size = 10

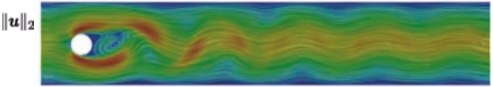


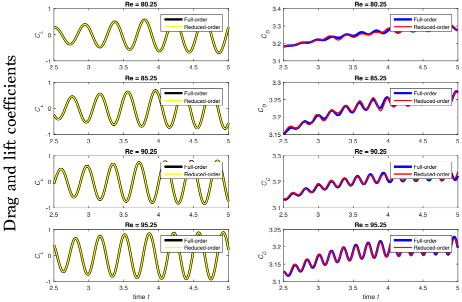
	Time / s
One run F-O	1.14E+03
GPRs	4.30E-01
One run R-O	8.83E-05

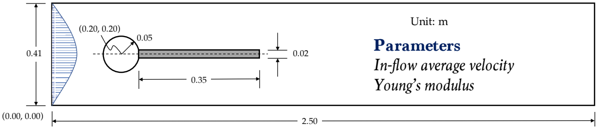
Reduced order model with RB size = 6

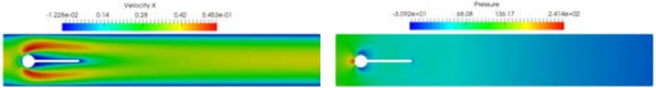
Applications in computational physics



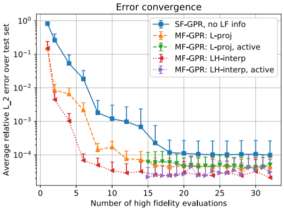




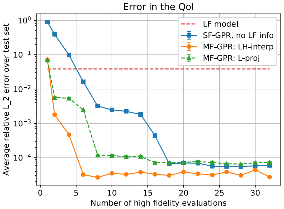




MF recovery with 80 LF and 1 HF snapshots



Global error in solution field



QoI: Lift

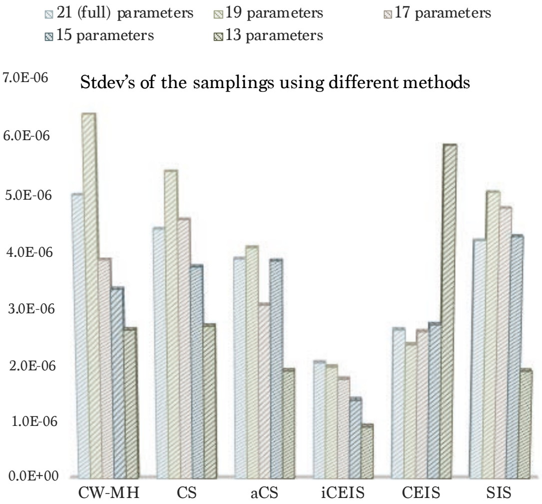
Convergence of average relative error over 225 test samples

Application to risk assessment

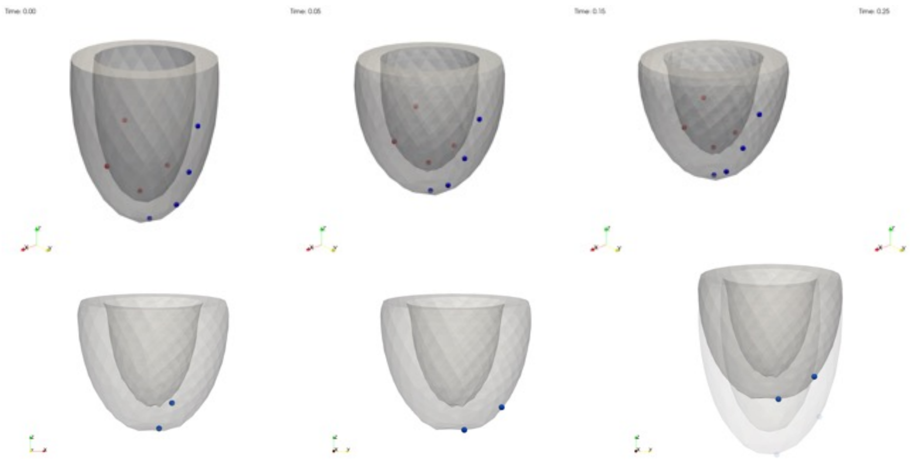
Estimation of a probability of 1E-05:
The max stress in the two nonlinear components exceeding 110.4 and 204.8 MPa

Subset simulation
- **CW-MH**
Component-**W**ise **M**etropolis-**H**astings
- **CS**
Conditional **S**ampling
- **aCS**
Adaptive **C**onditional **S**ampling

Importance sampling
- **iCEIS**
Improved **C**ross-**E**ntropy **I**mportance **S**ampling
- **CEIS**
Cross-**E**ntropy **I**mportance **S**ampling
- **SIS**
Sequential **I**mportance **S**ampling



w/ Z. Zhang and J. S. Hesthaven



Active contraction of truncated ellipsoid

— Benchmark in ***Cardiac Mechanics***

- Sensitivity analysis
- Parameter estimation

L. Cicci, S. Fresca, **MG**, A. Manzoni, P. Zunico. arXiv:2302.08216, 2023

Institution of MECHANICAL ENGINEERS
Journal of AEROSPACE ENGINEERING

Original Article

Proc IMechE Part G: J Aerospace Engineering
2019, Vol. 233(16) 5896–5912

DOI: 10.1177/0954410019890721
journals.sagepub.com/home/pig
SAGE

Non-intrusive reduced-order modeling for fluid problems: A brief review

Jian Yu[1], Chao Yan[1] and Mengwu Guo[2]

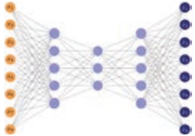
Abstract

Despite tremendous progress seen in the computational fluid dynamics community for the past few decades, numerical tools are still too slow for the simulation of practical flow problems, consuming thousands or even millions of computational core-hours. To enable feasible multi-disciplinary analysis and design, the numerical techniques need to be accelerated by orders of magnitude. Reduced-order modeling has been considered one promising approach for such purposes. Recently, non-intrusive reduced-order modeling has drawn great interest in the scientific computing community due to its flexibility and efficiency and undergoes rapid development at present with different approaches emerging from various perspectives. In this paper, a brief review of non-intrusive reduced-order modeling in the context of fluid problems is performed involving three key aspects: i.e. dimension reduction of the solution space, surrogate models, and sampling strategies. Furthermore, non-intrusive reduced-order modelings regarding to some interesting topics such as unsteady flows, shock-dominating flows are also discussed. Finally, discussions on future development of non-intrusive reduced-order modeling for fluid problems are presented.

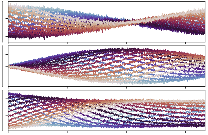
Keywords

Reduced order modeling, non-intrusive, machine learning, surrogate model

Dimensionality reduction
(Manifold learning)



Reduced-state dynamics
representation



Deterministic	*Encoding-decoding*: Linear PCA – POD Multilayer/Convolutional autoencoder Bregman autoencoder ……	Projection-inspired dense structure — OpInf Sparse structure — SINDy Time integration inside loss — neural ODE Auto-regressive — "time-stepper", R-K NN ……
Probabilistic	Variational autoencoder Gaussian process latent variable model Manifold Gaussian process $\kappa(\mathbf{u}, \mathbf{u}') = k(\mathcal{M}_{\mathrm{enc}}(\mathbf{u}), \mathcal{M}_{\mathrm{enc}}(\mathbf{u}'))$ ……	Bayesian inference Bayesian inference + sparse regression Bayesian neural networks Deep kernel learning ……

Bayesian operator inference

Prof. Karen Willcox

Shane McQuarrie

ODEN INSTITUTE FOR COMPUTATIONAL ENGINEERING & SCIENCES

- **MG**, S. A. McQuarrie, K. E. Willcox. *Comput. Methods Appl. Mech. Eng.*: Article 115336, 2022.

Data-driven reduced-order operator inference

Full-order system: a quadratic one as example

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q}(t) = \mathbf{A}\mathbf{q}(t) + \mathbf{H}[\mathbf{q}(t) \otimes \mathbf{q}(t)] + \mathbf{B}\mathbf{u}(t) + \mathbf{c}, \quad \mathbf{q}(t_0) = \mathbf{q}_0, \quad t \in [t_0, t_f]$$

Projection-based reduced-order system: $\mathbf{q} \approx \mathbf{V}\hat{\mathbf{q}}$, POD basis $\dim(\mathrm{Col}(\mathbf{V})) = r \ll N_h$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{q}}(t) = \hat{\mathbf{A}}\hat{\mathbf{q}}(t) + \hat{\mathbf{H}}[\hat{\mathbf{q}}(t) \otimes \hat{\mathbf{q}}(t)] + \hat{\mathbf{B}}\mathbf{u}(t) + \hat{\mathbf{c}}, \quad \hat{\mathbf{q}}(t_0) = \mathbf{V}^{\mathrm{T}}\mathbf{q}_0, \quad t \in [t_0, t_f]$$

Deterministic inference of reduced-order operators

Minimize the (regularized) reduced-order residual sum of squares with the projected state data

$$\min_{\hat{\mathbf{A}},\hat{\mathbf{H}},\hat{\mathbf{B}},\hat{\mathbf{c}}} \left\{ \sum_{j=0}^{k-1} \left\| \hat{\mathbf{A}}\hat{\mathbf{q}}_j + \hat{\mathbf{H}}[\hat{\mathbf{q}}_j \otimes \hat{\mathbf{q}}_j] + \hat{\mathbf{B}}\mathbf{u}_j + \hat{\mathbf{c}} - \dot{\hat{\mathbf{q}}}_j \right\|_2^2 + \mathscr{P}\left([\, \hat{\mathbf{A}} \;\; \hat{\mathbf{H}} \;\; \hat{\mathbf{B}} \;\; \hat{\mathbf{c}} \,] \right) \right\}$$

Data-driven reduced-order operator inference

Full-order system: a quadratic one as example

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{q}(t) = \mathbf{A}\mathbf{q}(t) + \mathbf{H}[\mathbf{q}(t)\otimes\mathbf{q}(t)] + \mathbf{B}\mathbf{u}(t) + \mathbf{c}, \quad \mathbf{q}(t_0) = \mathbf{q}_0, \quad t \in [t_0, t_f]$$

Projection-based reduced-order system: $\mathbf{q} \approx \mathbf{V}\hat{\mathbf{q}}$, POD basis $\dim(\mathrm{Col}(\mathbf{V})) = r \ll N_h$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{q}}(t) = \hat{\mathbf{A}}\hat{\mathbf{q}}(t) + \hat{\mathbf{H}}[\hat{\mathbf{q}}(t)\otimes\hat{\mathbf{q}}(t)] + \hat{\mathbf{B}}\mathbf{u}(t) + \hat{\mathbf{c}}, \quad \hat{\mathbf{q}}(t_0) = \mathbf{V}^{\mathrm{T}}\mathbf{q}_0, \quad t \in [t_0, t_f]$$

Deterministic inference of reduced-order operators

Minimize the (regularized) reduced-order residual sum of squares with the projected state data

$$\min_{\hat{\mathbf{A}},\hat{\mathbf{H}},\hat{\mathbf{B}},\hat{\mathbf{c}}} \left\{ \sum_{j=0}^{k-1} \left\| \hat{\mathbf{A}}\hat{\mathbf{q}}_j + \hat{\mathbf{H}}[\hat{\mathbf{q}}_j \otimes \hat{\mathbf{q}}_j] + \hat{\mathbf{B}}\mathbf{u}_j + \hat{\mathbf{c}} - \dot{\hat{\mathbf{q}}}_j \right\|_2^2 + \mathscr{P}([\,\hat{\mathbf{A}}\ \hat{\mathbf{H}}\ \hat{\mathbf{B}}\ \hat{\mathbf{c}}\,]) \right\}$$

model misspecification

$$:= \min_{\hat{\mathbf{O}}} \left\{ \|\mathbf{D}\hat{\mathbf{O}}^{\mathrm{T}} - \mathbf{R}^{\mathrm{T}}\|_F^2 + \mathscr{P}(\hat{\mathbf{O}}) \right\} = \sum_{i=1}^{r} \min_{\hat{\mathbf{o}}_i} \left\{ \|\mathbf{D}\hat{\mathbf{o}}_i - \mathbf{r}_i\|_2^2 + \mathscr{P}_i(\hat{\mathbf{o}}_i) \right\}$$

data noise   approximation error   regularization

Bayesian inference of reduced-order operators

We would like to change the following inference into a BAYESIAN fashion for **UQ**

$$\left\{\|\mathbf{D}\hat{\mathbf{o}}_i - \mathbf{r}_i\|_2^2 + \mathscr{P}_i(\hat{\mathbf{o}}_i)\right\} \to \min$$

Point estimate ⟶ Posterior distribution

Bayesian regression – likelihood and prior

- **Likelihood**

$$\mathbf{r}_i = \mathbf{D}\hat{\mathbf{o}}_i + \boldsymbol{\epsilon}_i, \qquad p(\boldsymbol{\epsilon}_i) = \mathcal{N}(\boldsymbol{\epsilon}_i|\mathbf{0}_k, \mathbf{K})$$

  - Residual $\boldsymbol{\epsilon}_i$ defined by a zero-mean Gaussian process
  - Independent noise $\mathbf{K} = \sigma_i^2\mathbf{I}_k$: $\quad p(\mathbf{r}_i|\,\mathbf{D}, \hat{\mathbf{o}}_i, \sigma_i^2) = \mathcal{N}(\mathbf{r}_i|\,\mathbf{D}\hat{\mathbf{o}}_i, \sigma_i^2\mathbf{I}_k)$

- **Operator prior**

$$p(\hat{\mathbf{o}}_i|\,\sigma_i^2, \boldsymbol{\lambda}_i) = \mathcal{N}(\hat{\mathbf{o}}_i|\,\boldsymbol{\beta}_i, \sigma_i^2\mathrm{diag}(\boldsymbol{\lambda}_i)^{-1})$$

  - w/ pre-defined mean vector, and hyper-parametrized variance

Bayesian inference of reduced-order operators

Bayesian regression – posterior

$p(\hat{\mathbf{o}}_i|\, \mathbf{D}, \mathbf{r}_i, \sigma_i^2, \boldsymbol{\lambda}_i) = \mathcal{N}(\hat{\mathbf{o}}_i|\, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \propto p(\mathbf{r}_i|\, \mathbf{D}, \hat{\mathbf{o}}_i, \sigma_i^2)\; p(\hat{\mathbf{o}}_i|\, \sigma_i^2, \boldsymbol{\lambda}_i)$, in which

$$\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_i(\mathbf{D}, \sigma_i^2, \boldsymbol{\lambda}_i) = \sigma_i^2 \left[\mathrm{diag}(\boldsymbol{\lambda}_i) + \mathbf{D}^{\mathrm{T}}\mathbf{D}\right]^{-1}$$

$$\boldsymbol{\mu}_i = \boldsymbol{\mu}_i(\mathbf{r}_i, \mathbf{D}, \boldsymbol{\lambda}_i) = \boldsymbol{\beta}_i + \left[\mathrm{diag}(\boldsymbol{\lambda}_i) + \mathbf{D}^{\mathrm{T}}\mathbf{D}\right]^{-1} \mathbf{D}^{\mathrm{T}}(\mathbf{r}_i - \mathbf{D}\boldsymbol{\beta}_i)$$

- **Posterior mean**: equivalent to Tikhonov-regularized deterministic OpInf (= *ridge regression*)
  - Resolve issues with ill-conditioned Gram matrices
  - Help stabilize the reduced-order time integration
- **Posterior variance**: quantifying the modeling uncertainty introduced by Tikhonov regularization
- **Regularization selection**: aided by maximum marginal likelihood

$$\mathcal{L}_i(\sigma_i^2, \boldsymbol{\lambda}_i) = -\frac{1}{2\sigma_i^2}\|\mathbf{r}_i - \mathbf{D}\boldsymbol{\mu}_i\|_2^2 - \frac{1}{2\sigma_i^2}\left\|\mathrm{diag}(\boldsymbol{\lambda}_i)^{1/2}\,\delta\boldsymbol{\mu}_i\right\|_2^2$$
$$-\frac{1}{2}\log\left|\mathrm{diag}(\boldsymbol{\lambda}_i) + \mathbf{D}^{\mathrm{T}}\mathbf{D}\right| + \frac{1}{2}\log(\boldsymbol{\lambda}_i)^{\mathrm{T}}\mathbf{1}_{d(r,m)} - \frac{k}{2}\log(\sigma_i^2) - \frac{k}{2}\log(2\pi)$$

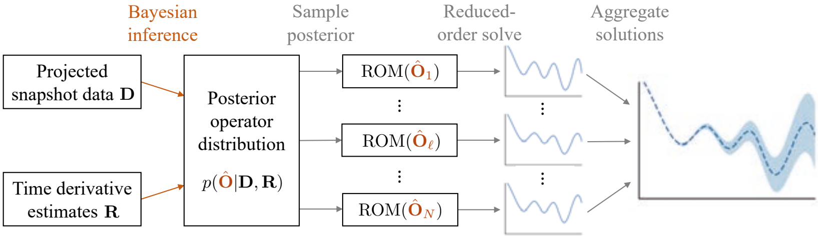
Bayesian inference of reduced-order operators

Bayesian regression – posterior

$p(\hat{\mathbf{o}}_i \mid \mathbf{D}, \mathbf{r}_i, \sigma_i^2, \boldsymbol{\lambda}_i) = \mathcal{N}(\hat{\mathbf{o}}_i \mid \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \propto p(\mathbf{r}_i \mid \mathbf{D}, \hat{\mathbf{o}}_i, \sigma_i^2)\, p(\hat{\mathbf{o}}_i \mid \sigma_i^2, \boldsymbol{\lambda}_i)$, in which

$$\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_i(\mathbf{D}, \sigma_i^2, \boldsymbol{\lambda}_i) = \sigma_i^2 \left[\mathrm{diag}(\boldsymbol{\lambda}_i) + \mathbf{D}^{\mathrm{T}}\mathbf{D}\right]^{-1}$$

$$\boldsymbol{\mu}_i = \boldsymbol{\mu}_i(\mathbf{r}_i, \mathbf{D}, \boldsymbol{\lambda}_i) = \boldsymbol{\beta}_i + \left[\mathrm{diag}(\boldsymbol{\lambda}_i) + \mathbf{D}^{\mathrm{T}}\mathbf{D}\right]^{-1} \mathbf{D}^{\mathrm{T}}(\mathbf{r}_i - \mathbf{D}\boldsymbol{\beta}_i)$$



A probabilistic reduced-order model: $p(\hat{\mathbf{q}}(t) \mid \mathbf{D}, \mathbf{R}) = \int p(\hat{\mathbf{q}}(t) \mid \hat{\mathbf{o}}_1, \ldots, \hat{\mathbf{o}}_r) \prod_{i=1}^{r} p(\hat{\mathbf{o}}_i \mid \mathbf{D}, \mathbf{r}_i)\, d\hat{\mathbf{o}}_i$

Noised Euler equations

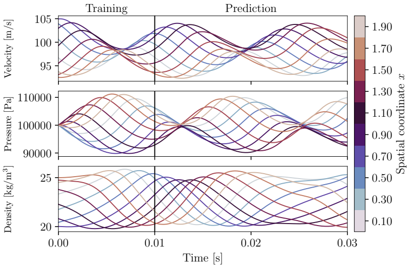
$$\begin{cases} \dfrac{\partial}{\partial t}\left[\rho\right] = -\dfrac{\partial}{\partial x}\left[\rho u\right] \\ \dfrac{\partial}{\partial t}\left[\rho u\right] = -\dfrac{\partial}{\partial x}\left[\rho u^2 + p\right] \\ \dfrac{\partial}{\partial t}\left[\rho e\right] = -\dfrac{\partial}{\partial x}\left[(\rho e + p)u\right] \end{cases}$$

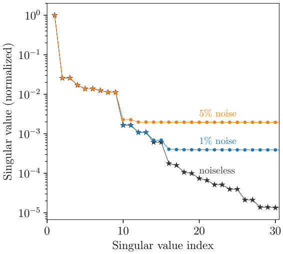
$$\Downarrow$$

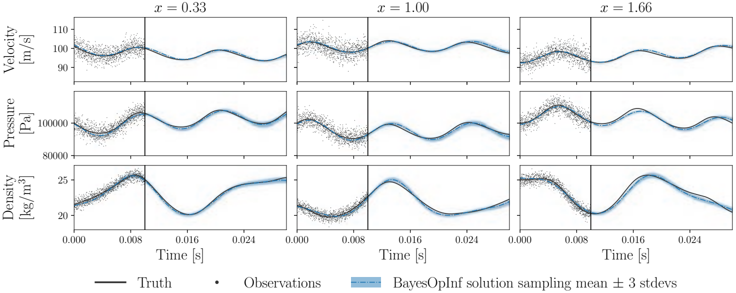
$$\begin{cases} \dfrac{\partial u}{\partial t} = -u\dfrac{\partial u}{\partial x} - \dfrac{1}{\rho}\dfrac{\partial p}{\partial x} \\ \dfrac{\partial p}{\partial t} = -\gamma p\dfrac{\partial u}{\partial x} - u\dfrac{\partial p}{\partial x} \\ \dfrac{\partial}{\partial t}\left[\dfrac{1}{\rho}\right] = -u\dfrac{\partial}{\partial x}\left[\dfrac{1}{\rho}\right] + \dfrac{1}{\rho}\dfrac{\partial u}{\partial x} \end{cases}$$

$$\Downarrow$$

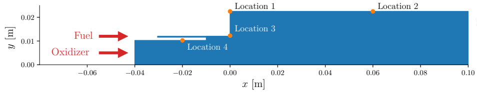
$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{q}}(t) = \hat{\mathbf{H}}[\hat{\mathbf{q}}(t) \otimes \hat{\mathbf{q}}(t)]$$



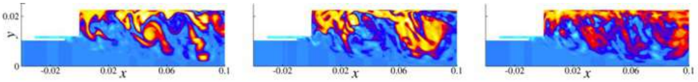




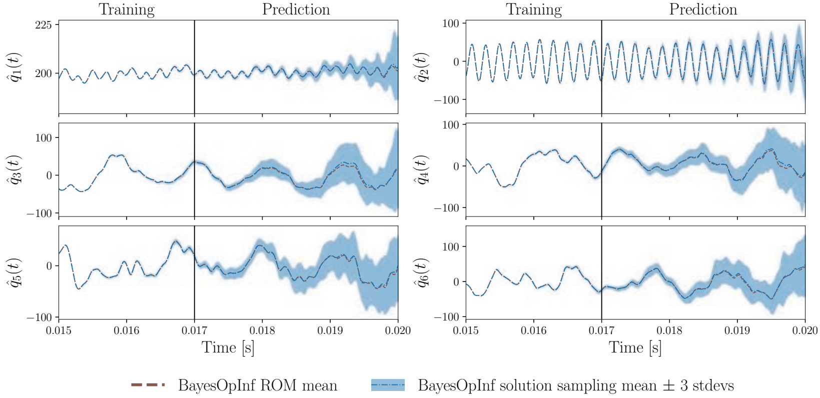
2D single-injector combustion process



McQuarrie, et al. *J. R. Soc. N. Z.* (2021)



(a) $t = t_{15000} = 0.0165$s. (b) $t = t_{20000} = 0.0170$s. (c) $t = t_{25000} = 0.0175$s.

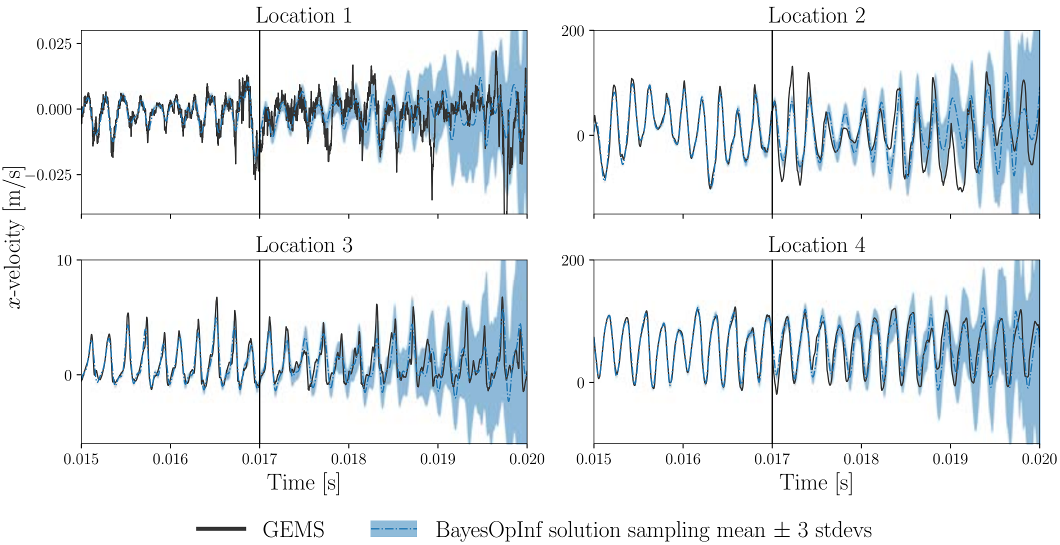


$$\frac{\partial \vec{q}_{\mathrm{c}}}{\partial t} + \nabla \cdot (\vec{K} - \vec{K}_{\mathrm{v}}) = \vec{S}$$

$$\vec{q}_{\mathrm{c}} = (\rho, \rho v_x, \rho v_y, \rho e, \rho Y_1, \rho Y_2, \rho Y_3, \rho Y_4)$$

- $\vec{K}$ inviscid flux
- $\vec{K}_{\mathrm{v}}$ viscous flux
- $\vec{S}$ source term

2D single-injector combustion process



Deep kernel learning of reduced dynamics

Prof. Christoph Brune

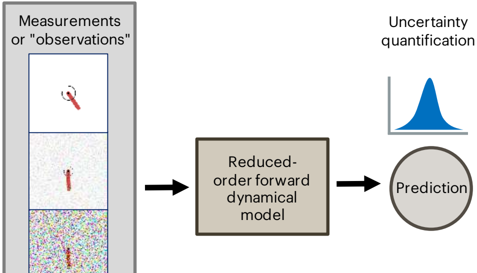
Dr. Nicolò Botteghi

UNIVERSITY OF TWENTE.

- N. Botteghi, **MG**, C. Brune. *Scientific Reports*, 12: 21530, 2022.

Deep kernel learning

Learn reduced-order latent dynamics
from high-dimensional measurements — images
corrupted by noise



Predictive data-driven **ROM** + **UQ**

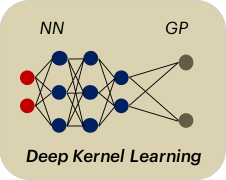
Deep Kernel Learning

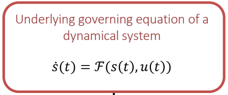
$$
\begin{aligned}
k_{\mathrm{DKL}}(\boldsymbol{x}, \boldsymbol{x}'; \mathbf{W}, \boldsymbol{\theta}) \\
&= k(\mathcal{NN}(\boldsymbol{x}; \mathbf{W}), \mathcal{NN}(\boldsymbol{x}'; \mathbf{W}); \boldsymbol{\theta})
\end{aligned}
$$

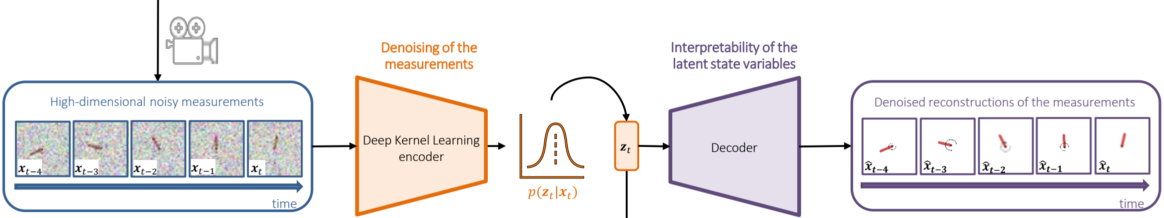
- ❑ DKL is a special manifold Gaussian process
- ❑ NN trained as GP hyperparameters
- ❑ Simpler and cheaper than Bayesian NN
- ❑ Used for **modeling, denoising, and UQ**

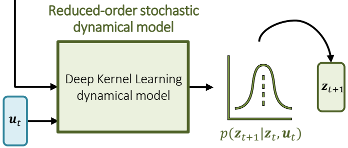
DKL: Wilson et al. PMLR, 2016. Manifold GP: Calandra et al. IJCNN, IEEE, 2016.

Deep Kernel Learning of dynamical systems from high-dimensional noisy data



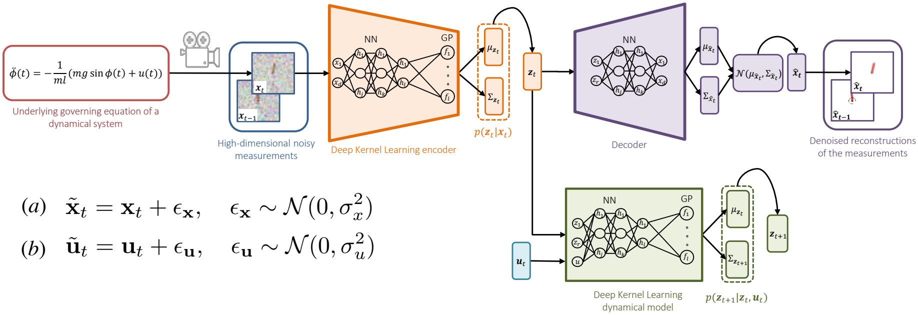






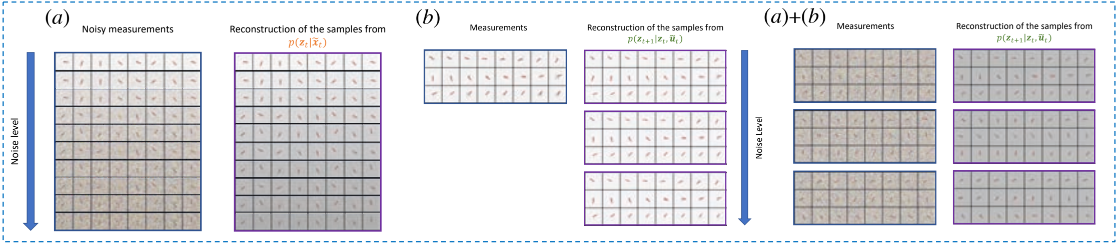
Background framework for RL & control: N. Botteghi. *Robotics Deep Reinforcement Learning with Loose Prior Knowledge*. PhD thesis, University of Twente, 2021.

Learning reduced dynamical models

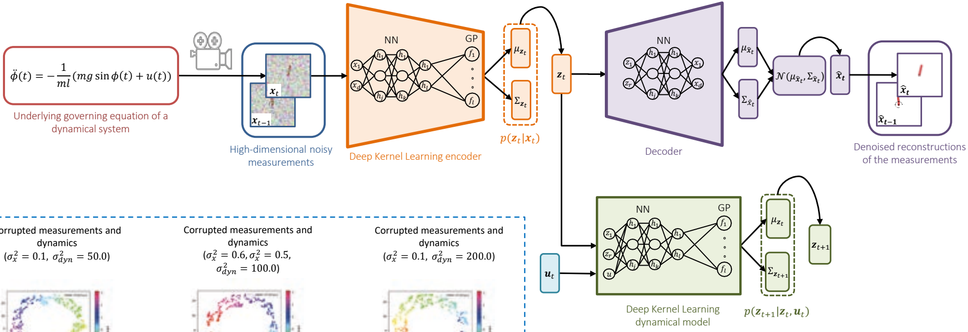


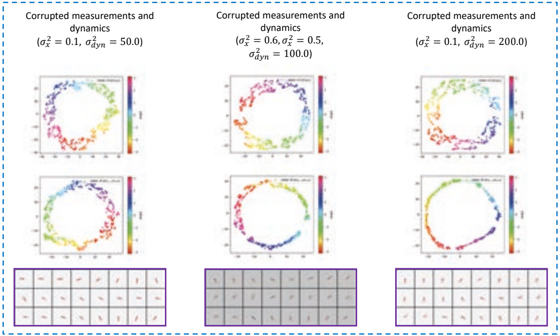
$$(a) \quad \tilde{\mathbf{x}}_t = \mathbf{x}_t + \epsilon_{\mathbf{x}}, \quad \epsilon_{\mathbf{x}} \sim \mathcal{N}(0, \sigma_x^2)$$

$$(b) \quad \tilde{\mathbf{u}}_t = \mathbf{u}_t + \epsilon_{\mathbf{u}}, \quad \epsilon_{\mathbf{u}} \sim \mathcal{N}(0, \sigma_u^2)$$



Learning reduced dynamical models





$$(a) \quad \tilde{\mathbf{x}}_t = \mathbf{x}_t + \epsilon_{\mathbf{x}}, \qquad \epsilon_{\mathbf{x}} \sim \mathcal{N}(0, \sigma_x^2)$$

$$(b) \quad \ddot{\phi}(t) = -\frac{1}{ml}(mg\sin\phi(t) + u(t) + \epsilon_{dyn})$$

$$\epsilon_{dyn} \sim \mathcal{N}(0, \sigma_{dyn}^2)$$

4TU.AMI **S**trategic **R**esearch **I**nitiative *Bridging Numerical Analysis and Machine Learning*

Two events in the Fall at CWI Amsterdam

- Workshop (Dec.) *Scientific Machine Learning: Bridging Computational Physics and Machine Learning*
- Autumn school (Oct.) *Scientific Machine Learning for Dynamical Systems*

Data-driven model reduction with *probabilistic* ML

- **Black-box**: Gaussian process surrogate modeling
- **Glass-box**: Bayesian reduced operator inference
- **Gray-box**: Deep Kernel Learning

Thank you!

Dr. Mengwu Guo

m.guo@utwente.nl

http://mengwuguo.weebly.com/