

GPS: Gaussian Process Subspace Regression for Model Reduction

Ruda Zhang¹ Simon Mak² David Dunson^{1,2}

¹Department of Mathematics, Duke University

²Department of Statistical Science, Duke University

INFORMS QSR best paper competition, Oct 25, 2021

ruda.zhang@duke.edu

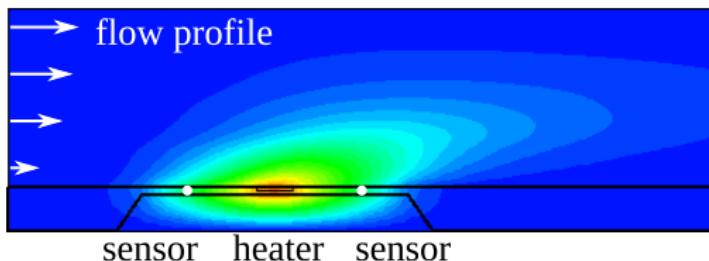
 @ZHANG_Ruda



Motivation: Parametric Studies of Computational Models

Anemometer:

- ▶ flow-speed MEMS device
- ▶ calibration
- ▶ convection-diffusion PDE
- ▶ linear ODE ($n = 29,008$)

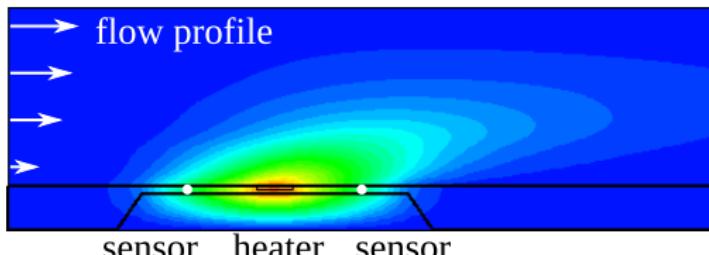


The MORwiki Community. Anemometer. Model Order Reduction Wiki, 2018.

Motivation: Parametric Studies of Computational Models

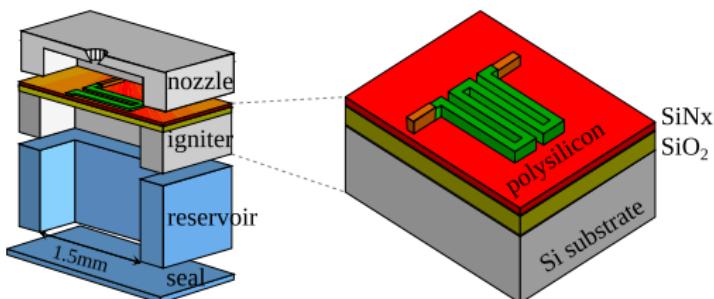
Anemometer:

- ▶ flow-speed MEMS device
- ▶ calibration
- ▶ convection-diffusion PDE
- ▶ linear ODE ($n = 29,008$)



Microthruster:

- ▶ solid propellant micro-rocket
- ▶ design
- ▶ heat transfer PDE
- ▶ linear ODE ($n = 4,257$)



The MORwiki Community. Anemometer. Model Order Reduction Wiki, 2018.



Oberwolfach Benchmark Collection. Thermal model. Model Order Reduction Wiki, 2018.

Reduced Order Modeling via Projection to Subspaces

Physics

$$x(0), u(t)$$



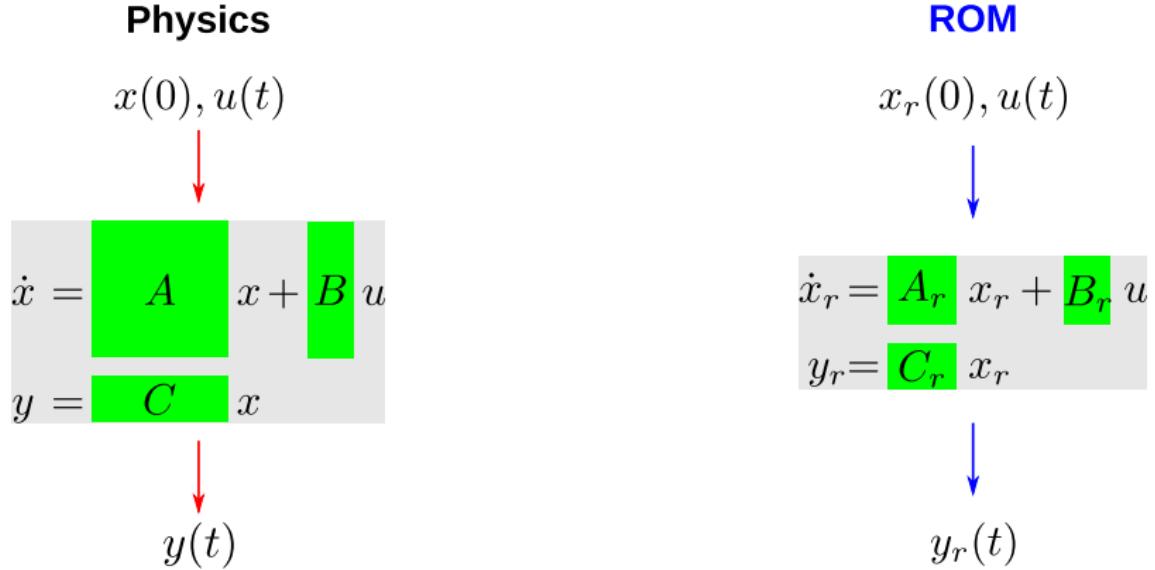
$$\begin{aligned}\dot{x} &= \begin{array}{|c|} \hline A \\ \hline \end{array} x + \begin{array}{|c|} \hline B \\ \hline \end{array} u \\ y &= \begin{array}{|c|} \hline C \\ \hline \end{array} x\end{aligned}$$



$$y(t)$$

$$\begin{aligned}O(n) &\sim O(n^2) \\ n > O(10^5)\end{aligned}$$

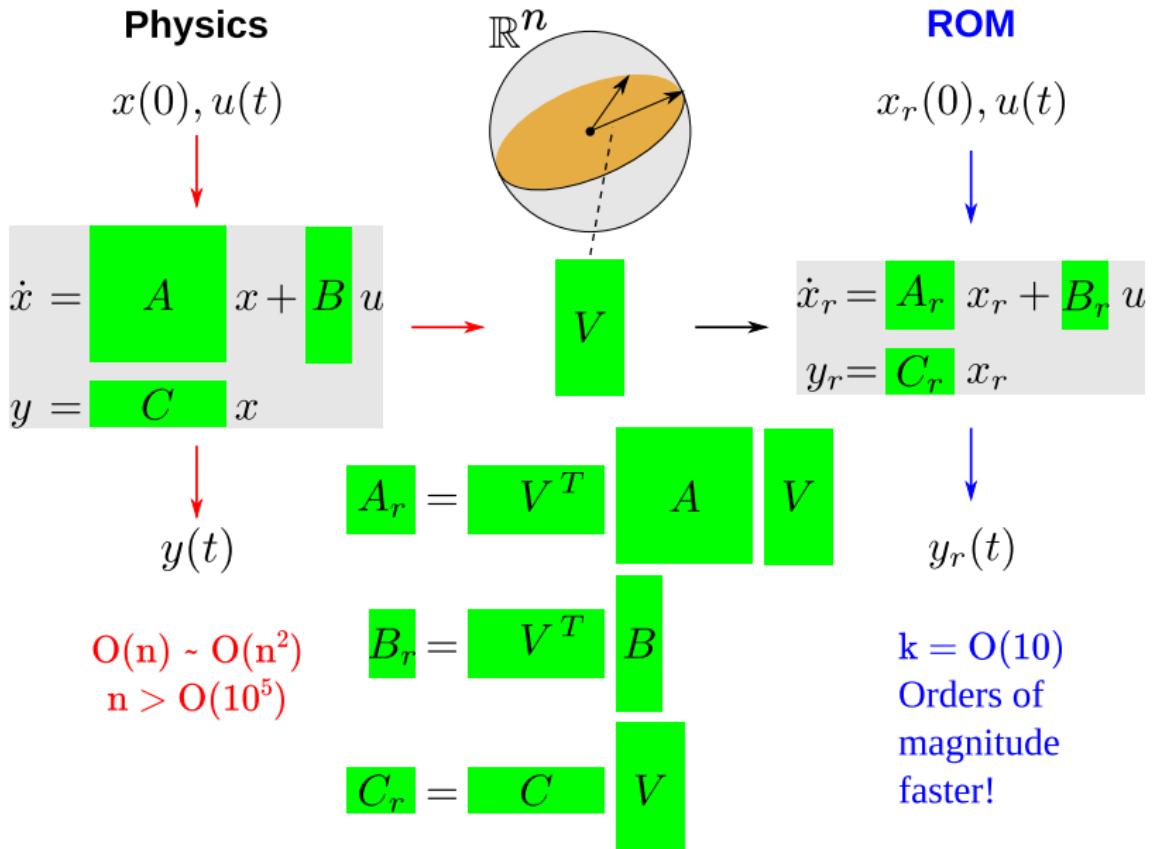
Reduced Order Modeling via Projection to Subspaces



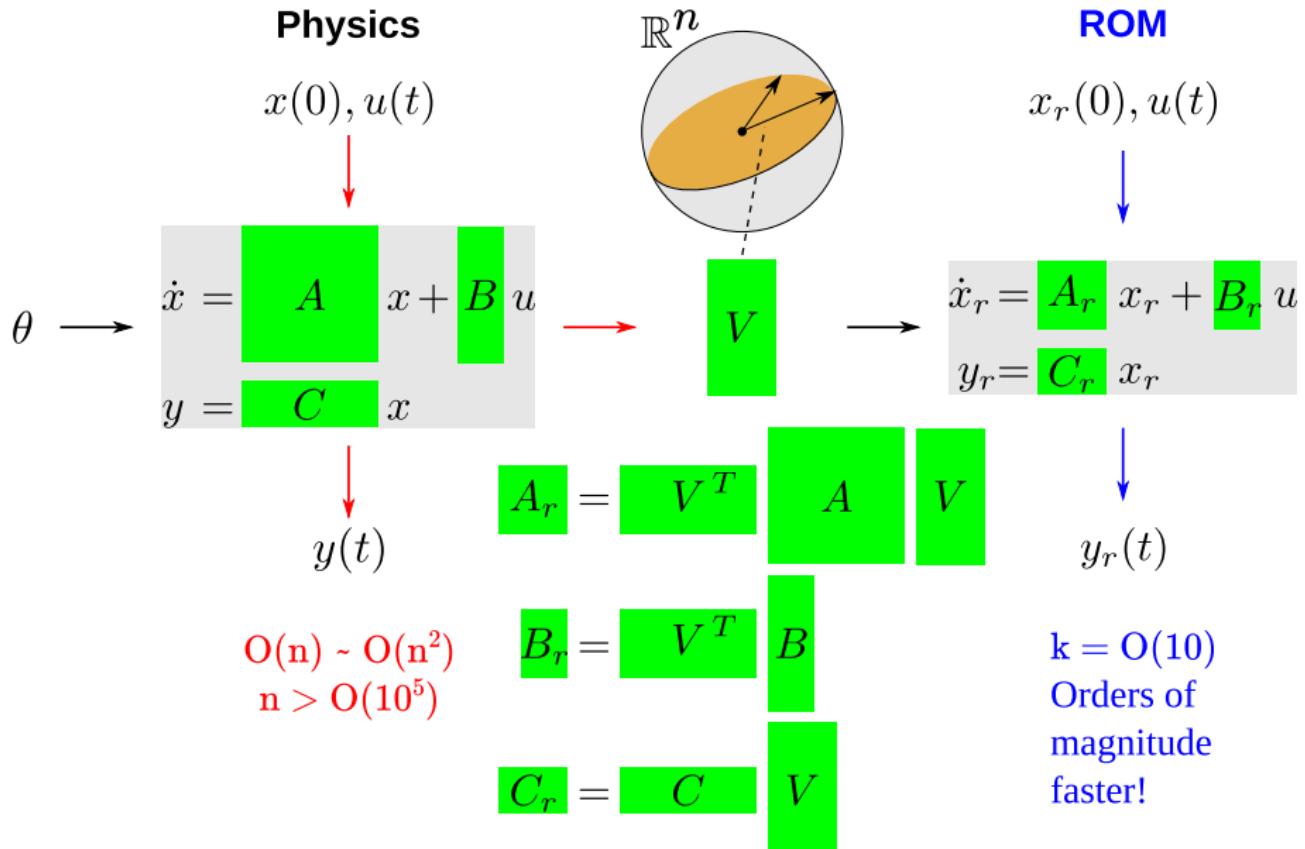
$O(n) \sim O(n^2)$
 $n > O(10^5)$

$k = O(10)$
Orders of
magnitude
faster!

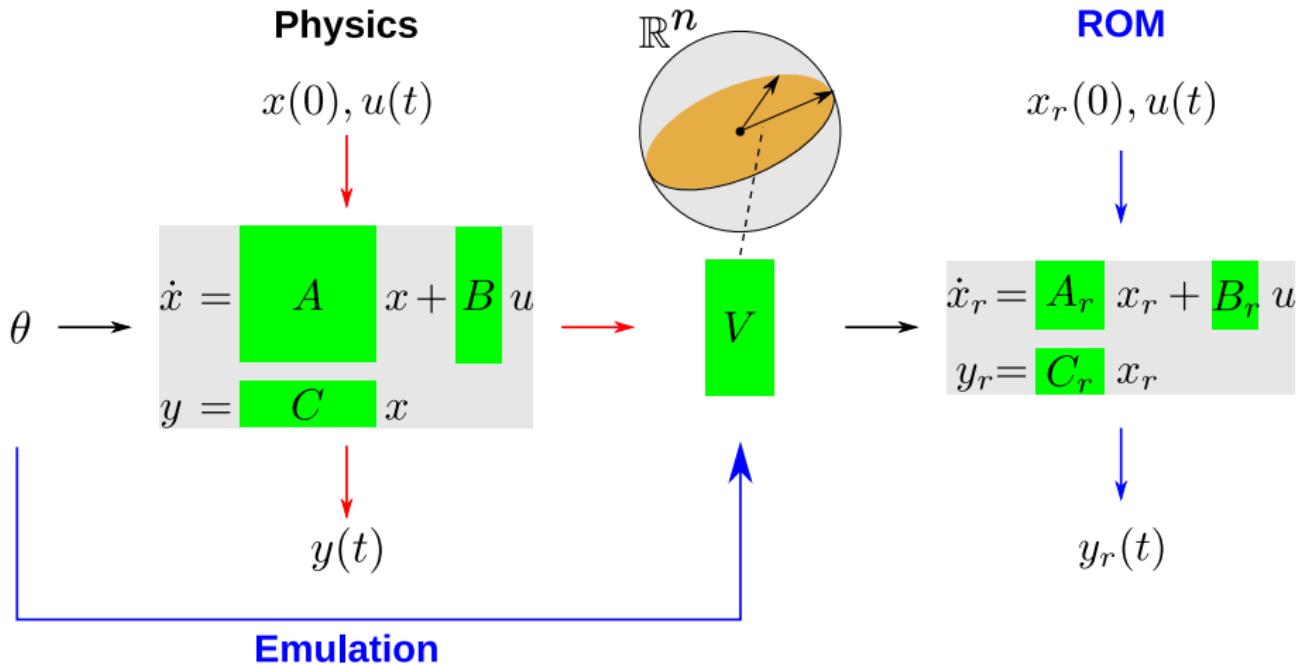
Reduced Order Modeling via Projection to Subspaces



Reduced Order Modeling via Projection to Subspaces



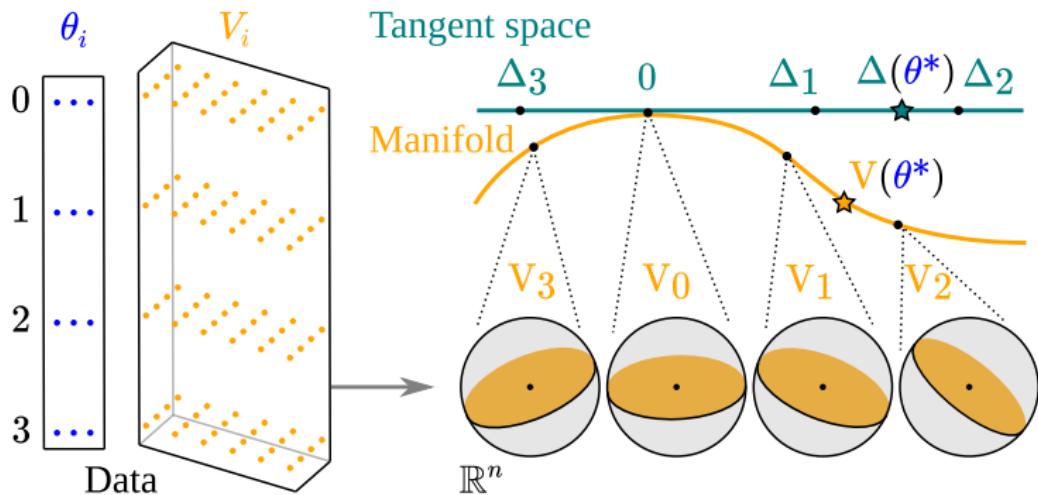
Reduced Order Modeling via Projection to Subspaces



Emulator + ROM \Rightarrow Parametric ROM

Problem: Emulating Subspace-valued Functions

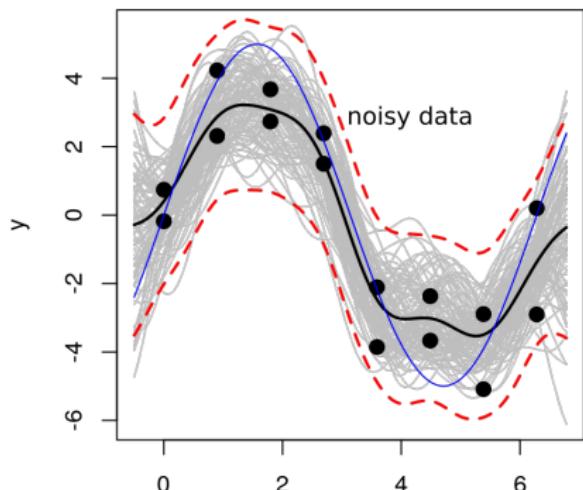
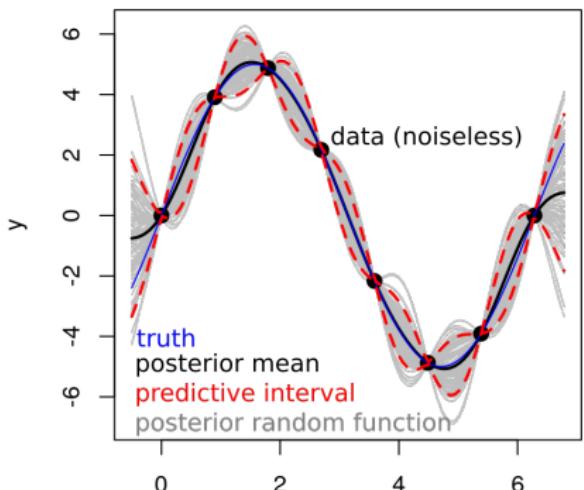
- ▶ Old idea: interpolate on tangent spaces.
- ▶ Pros: ↗ 7–45 times speedup than direct ROM.
- ▶ Cons: ↘ inflexible, ↘ extrinsic, ↘ no UQ.



D. Amsallem, C. Farhat. Interpolation method for adapting ROMs and application to aeroelasticity. *AIAA Journal*, 46(7):1803–1813, July 2008. (587 cites; 78 in 2020.)

Gaussian Process (GP) Models: Basics

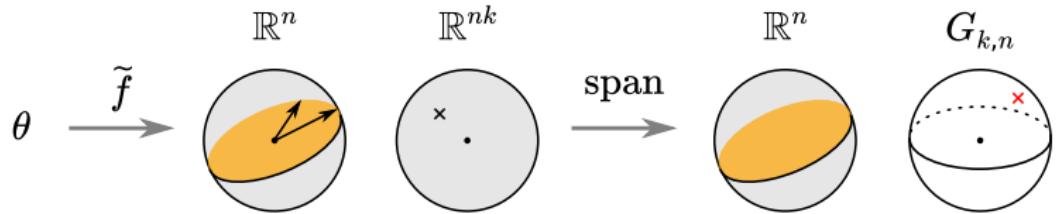
- ▶ Unknown function: $f : \Theta \mapsto \mathbb{R}$, **real-valued**.
- ▶ GP prior process: $f \sim \mathcal{GP}(\mu(x), k(x, x'; \psi))$, hyper-parameters ψ .
- ▶ Likelihood: $p(y | f) \sim N(f, \sigma^2)$, **non-singular** Gaussian.
- ▶ Posterior: $p(f | y, x) \propto p(f | x) p(y | f)$.
- ▶ Noiseless data \Rightarrow conditional, $p(f_* | f, x_*, x)$.
- ▶ Noisy data $\Rightarrow p(f_* | y, x_*, x) = \int p(f_* | f, x_*, x) p(f | y, x) df$.



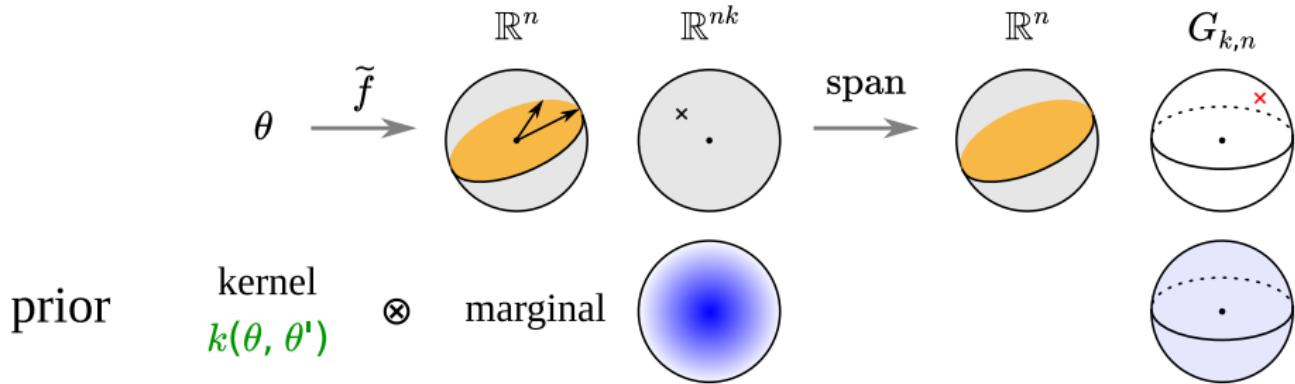
Gaussian Process Subspace (GPS) Model



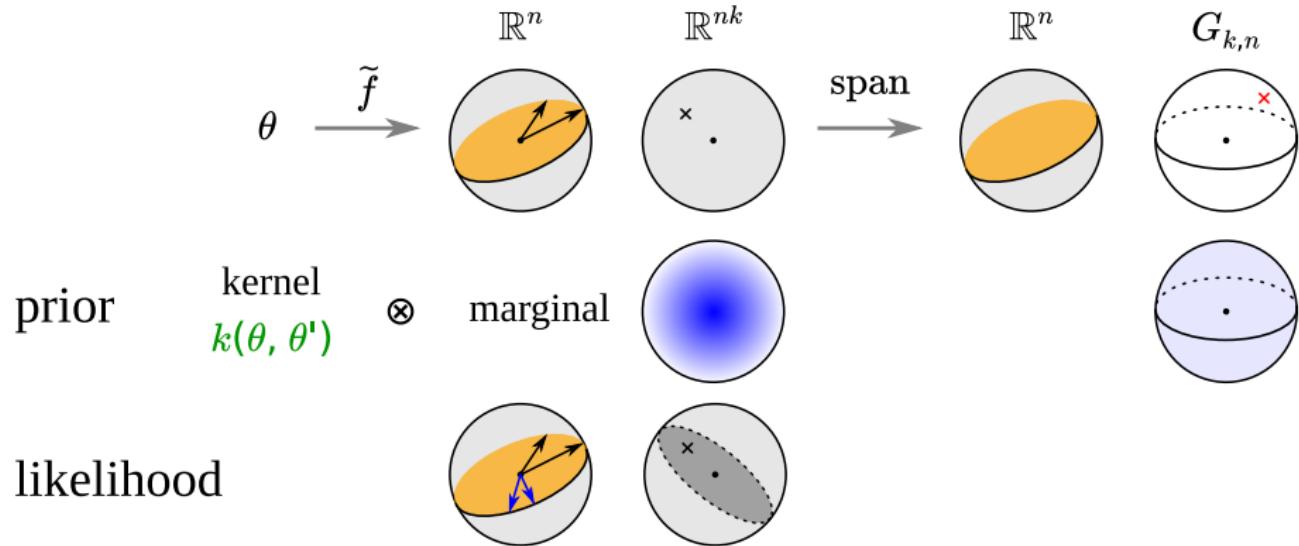
Gaussian Process Subspace (GPS) Model



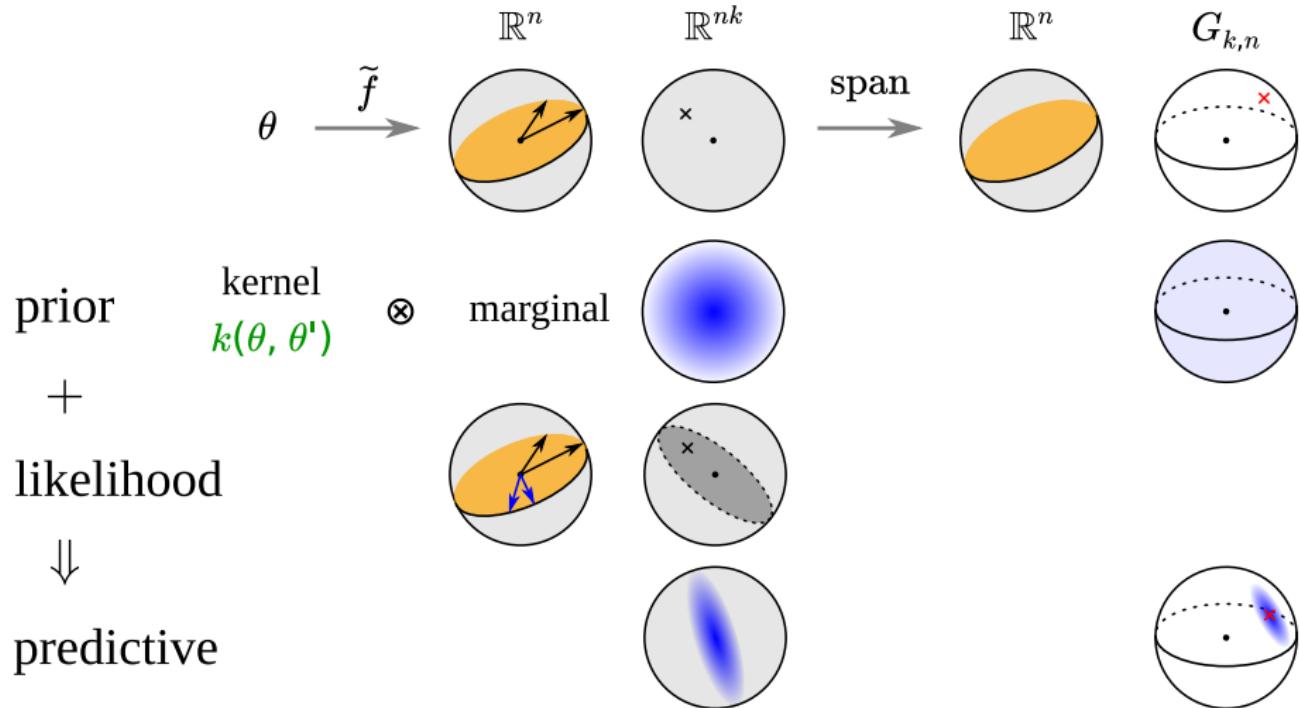
Gaussian Process Subspace (GPS) Model



Gaussian Process Subspace (GPS) Model



Gaussian Process Subspace (GPS) Model



The GPS Model

Prior process: $\tilde{f} \sim \mathcal{GP}(0, k \otimes \mathbf{I}_{nk})$.

$$(\mathbf{m}_*, \mathbf{m}) \sim N_{nk(l+1)}(0, \mathbf{K}_{l+1} \otimes \mathbf{I}_{nk})$$

Equal likelihood to all bases of a subspace:

$$L(\mathbf{m}_i | \mathfrak{X}_i) = \mathbf{1}(\mathbf{m}_i \in [\mathbf{x}_i])$$

$$[\mathbf{x}_i] = \{\text{vec}(\mathbf{X}_i \mathbf{A}) : \mathbf{A} \in \text{GL}_k\}$$

Predictive distribution has an analytical form:

$$\mathbf{m}_* | \mathfrak{X} \sim N_{nk}(0, \mathbf{I}_k \otimes \Sigma)$$

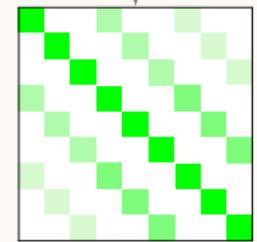
$$\Sigma(\theta) = \varepsilon^2 \mathbf{I}_n + \mathbf{X} [\mathbb{X}^T (\tilde{\mathbf{K}}_l \otimes \mathbf{I}_n) \mathbb{X}]^{-1} \mathbf{X}^T$$

Covariance structure:

$$n \begin{array}{|c|c|c|} \hline & & k \\ \hline & \text{---} & \text{---} \\ \hline & \text{---} & \text{---} \\ \hline \end{array} \sim N_{nk}(0, I_{nk})$$

$$(\begin{array}{|c|c|c|} \hline & & l \\ \hline & \text{---} & \text{---} \\ \hline & \text{---} & \text{---} \\ \hline \end{array})_{i=1}^l \sim N_{nkl}(0, \mathbf{K}_l \otimes I_{nk})$$

$$l \quad \begin{array}{|c|c|c|} \hline & & nk \\ \hline & \text{---} & \text{---} \\ \hline & \text{---} & \text{---} \\ \hline \end{array} \quad nk$$

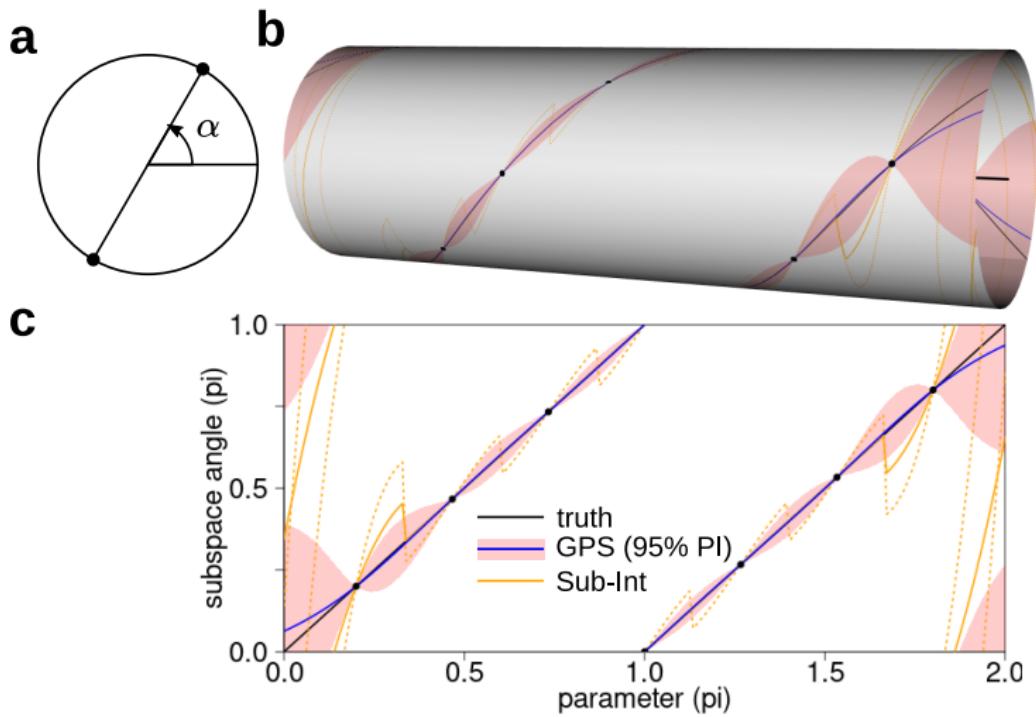
$$nkl \quad \begin{array}{|c|c|c|} \hline & & \\ \hline & \text{---} & \\ \hline & \text{---} & \\ \hline \end{array}$$


Theorem (Zhang, Mak, Dunson, 2021+)

Subspace prediction by the GPS has a matrix angular central Gaussian distribution $\mathfrak{M}_* | \mathfrak{X} \sim MACG(\Sigma)$, which admits easy sampling and inference.

Visualization of the GPS posterior process

- ▶ Target function $f : \mathbb{R} \mapsto G_{1,2}$
- ▶ Constant-speed rotation of lines in the plane.



Computational Complexity of GPS

Table: Interpolatory PROM methods: floating point operations.

	predict subspace	compute ROM
GPS	k^3l^3	$2k^3l^2$
Subspace-Int	$8nk^2$	$2nk^2$
Matrix-Int ^[1]	-	$2k^2l$
Manifold-Int ^[2]	-	$\mathcal{O}(k^3l)$

- Subspace-Int: most used; ↘ scale with $n > \mathcal{O}(10^5)$.
- Matrix-/Manifold-Int: ↗ faster online computation; ↘ less accurate.

[1] H. Panzer, J. Mohring, R. Eid, B. Lohmann. Parametric model order reduction by matrix interpolation. *at - Automatisierungstechnik*, 58 (8): 475–484, Aug. 2010.

[2] D. Amsallem, C. Farhat. An online method for interpolating linear parametric reduced-order models. *SIAM Journal on Scientific Computing*, 33 (5): 2169–2198, Jan. 2011.

Prediction Accuracy of GPS: Anemometer

Setup:

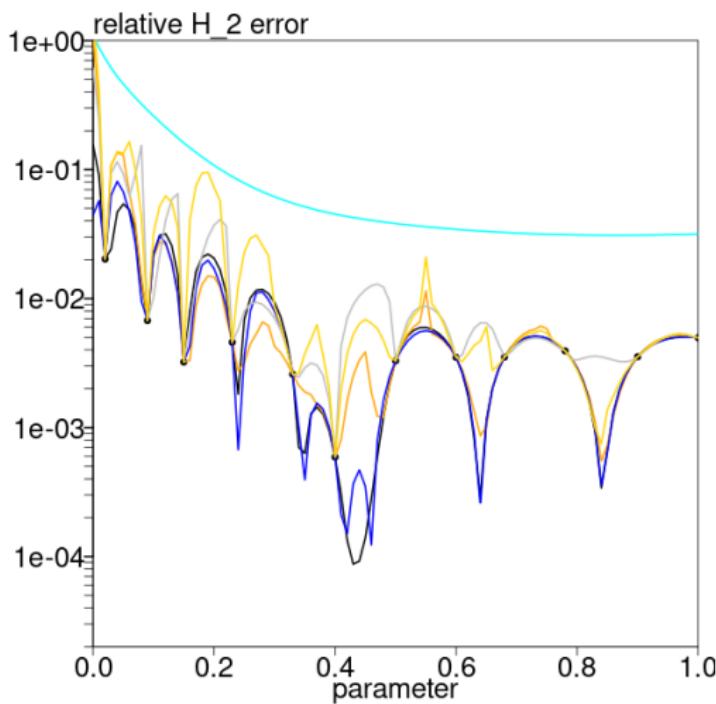
- ▶ subspace dim: $k = 20$
- ▶ sample size: $l = 12$

References: (lower is better)

- ▶ local bases (lower bound)
- ▶ global basis (upper bound)

Methods:

- ▶ GPS (rel. speedup: 6.7)
- ▶ Subspace interpolation
- ▶ Manifold interpolation
- ▶ Matrix interpolation



Prediction Accuracy of GPS: Anemometer

Setup:

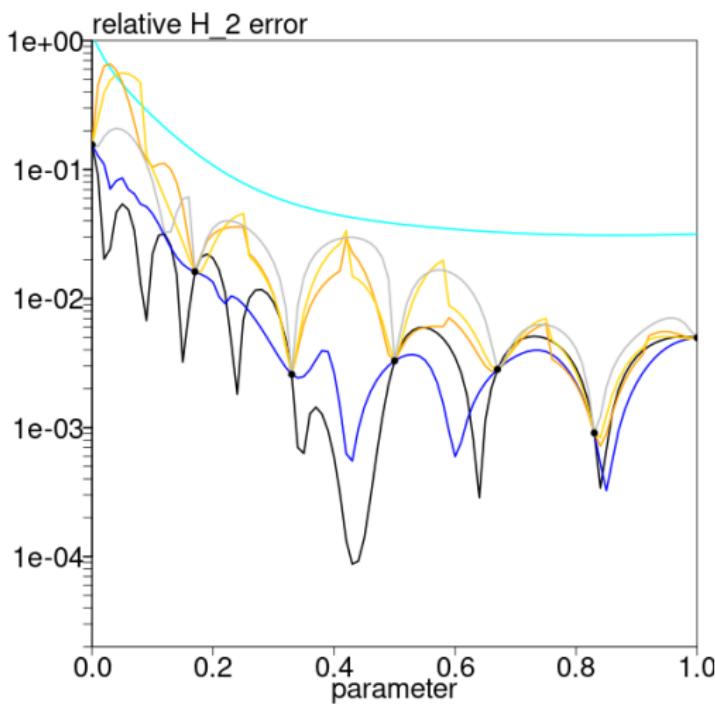
- ▶ subspace dim: $k = 20$
- ▶ sample size: $l = 7$

References: (lower is better)

- ▶ local bases (lower bound)
- ▶ global basis (upper bound)

Methods:

- ▶ GPS (rel. speedup: 34)
- ▶ Subspace interpolation
- ▶ Manifold interpolation
- ▶ Matrix interpolation



Prediction Accuracy of GPS: Anemometer

Setup:

- ▶ subspace dim: $k = 40$
- ▶ sample size: $l = 11$

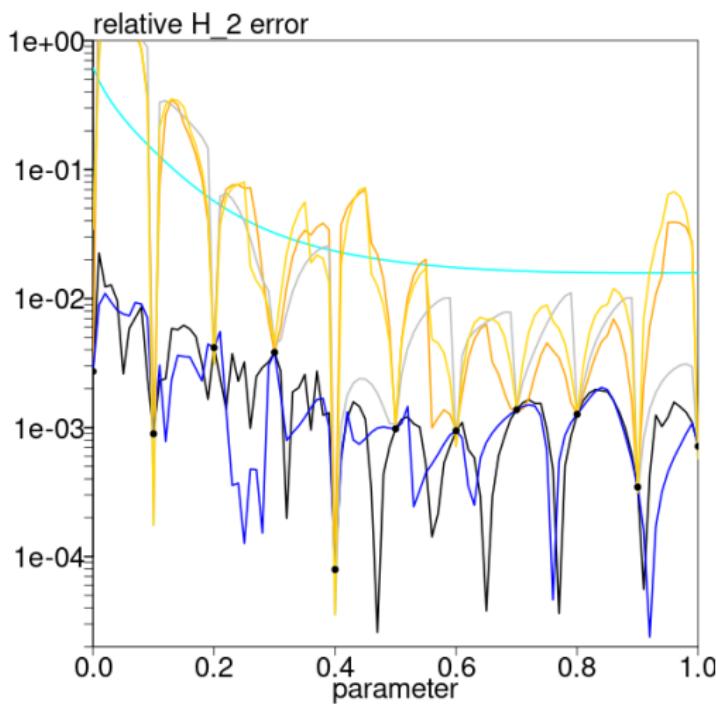
References: (lower is better)

- ▶ local bases (lower bound)
- ▶ global basis (upper bound)

Methods:

- ▶ GPS (rel. speedup: 4.4)
- ▶ Subspace interpolation
- ▶ Manifold interpolation
- ▶ Matrix interpolation

GPS retains accuracy
of local reduced bases!



Summary

- ▶ Problem: approximate subspace-valued functions
- ▶ Use: parametric reduced order modeling (PROM)
- ▶ State of the art: interpolate on tangent spaces
 - ↘ inflexible, extrinsic, deterministic
 - ↘ slow for large-scale systems, $n \gg 1$
- ▶ New idea: GP for subspace prediction
 - high-dim ($> 10^5$), non-vector output
 - ↗ accurate (small sample, high dim) + UQ \Leftarrow Bayesian nonparametric
 - ↗ fast: suitable for online computation
 - data-driven (GP) + physics-based (ROM) method for surrogate modeling of engineering systems.
 - future directions: prior, kernel, etc.



RZ, S. Mak, D. Dunson. Gaussian Process Subspace Regression for Model Reduction. arXiv, 2021. <https://arxiv.org/abs/2107.04668>.

Software

R package for GPS: <https://github.com/rudazhang/gpsr>

The screenshot shows the GitHub repository page for 'rudazhang / gprsr'. The 'Code' tab is active. In the file list, the 'inst/script' folder is circled in red. Other visible files include 'R', 'data', 'data-raw', 'man', '.Rbuildignore', and '.citationore'. The 'About' section describes it as a Gaussian process subspace regression R package. It includes links for 'r' and 'r-package', a 'Readme' file, and a 'GPL-3.0 License'. A 'Releases' section shows one release at version 0.0.0.9000 (Latest, released on Jun 29).

File/Folder	Description	Last Commit
R	Update LOOCV error gradient: option for trun...	25 days ago
data-raw	Update thermal model matrices, 3p anemom...	25 days ago
data	Update thermal model matrices, 3p anemom...	25 days ago
inst/script	First build: documentation, working script.	3 months ago
man	First build: documentation, working script.	3 months ago
.Rbuildignore	First build: documentation, working script.	3 months ago
.citationore	First build: documentation, working script.	3 months ago

Scripts for replicating the numerical examples in paper.