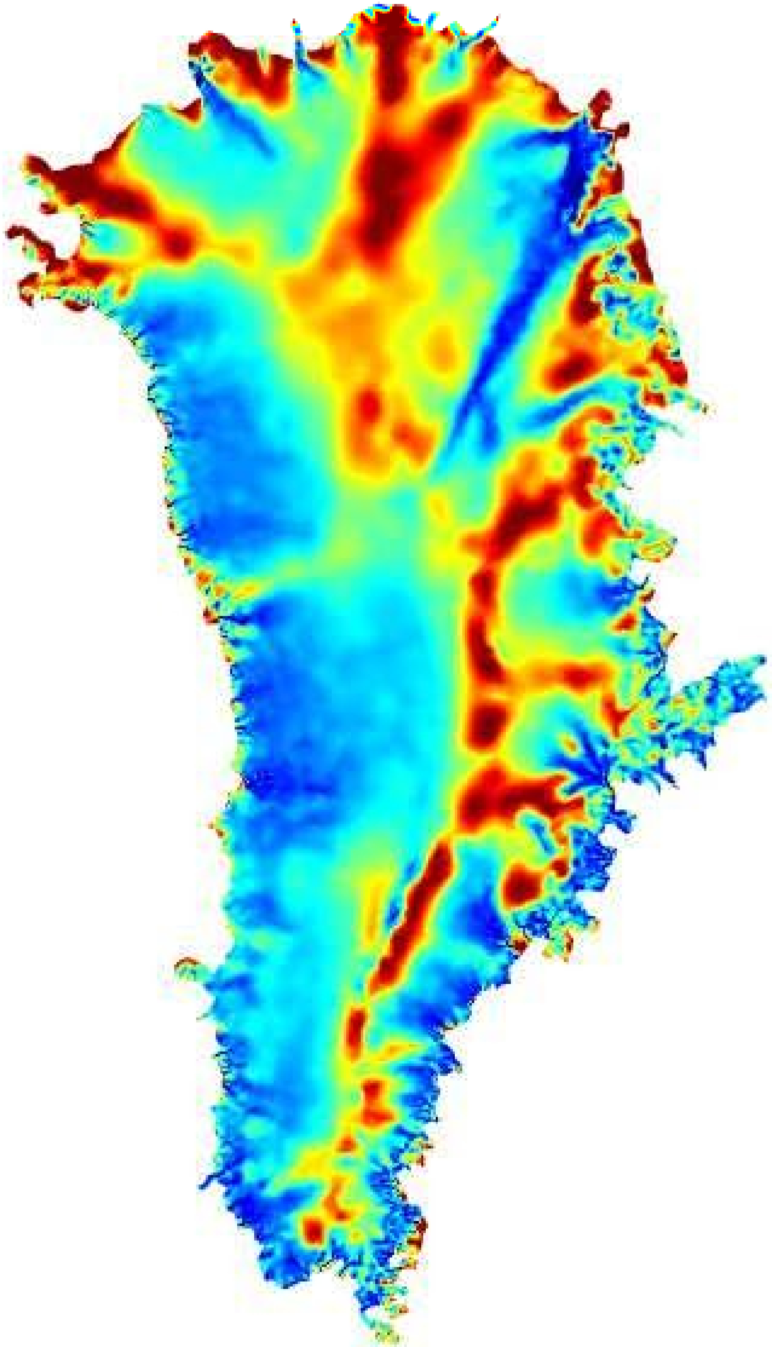


# Multifidelity Proper Orthogonal Decomposition

## Workshop on Learning Models from Data for Multi-Fidelity Fusion Plasma Physics

## Institute for Pure and Applied Mathematics
## April 15, 2026

**Professor Karen E. Willcox**
Director, Oden Institute for Computational Engineering & Sciences
Professor of Aerospace Engineering & Engineering Mechanics
University of Texas at Austin

ODEN

1973

2026

# 50+ Years of Leadership

in Interdisciplinary Research & Education in Computational Engineering & Sciences

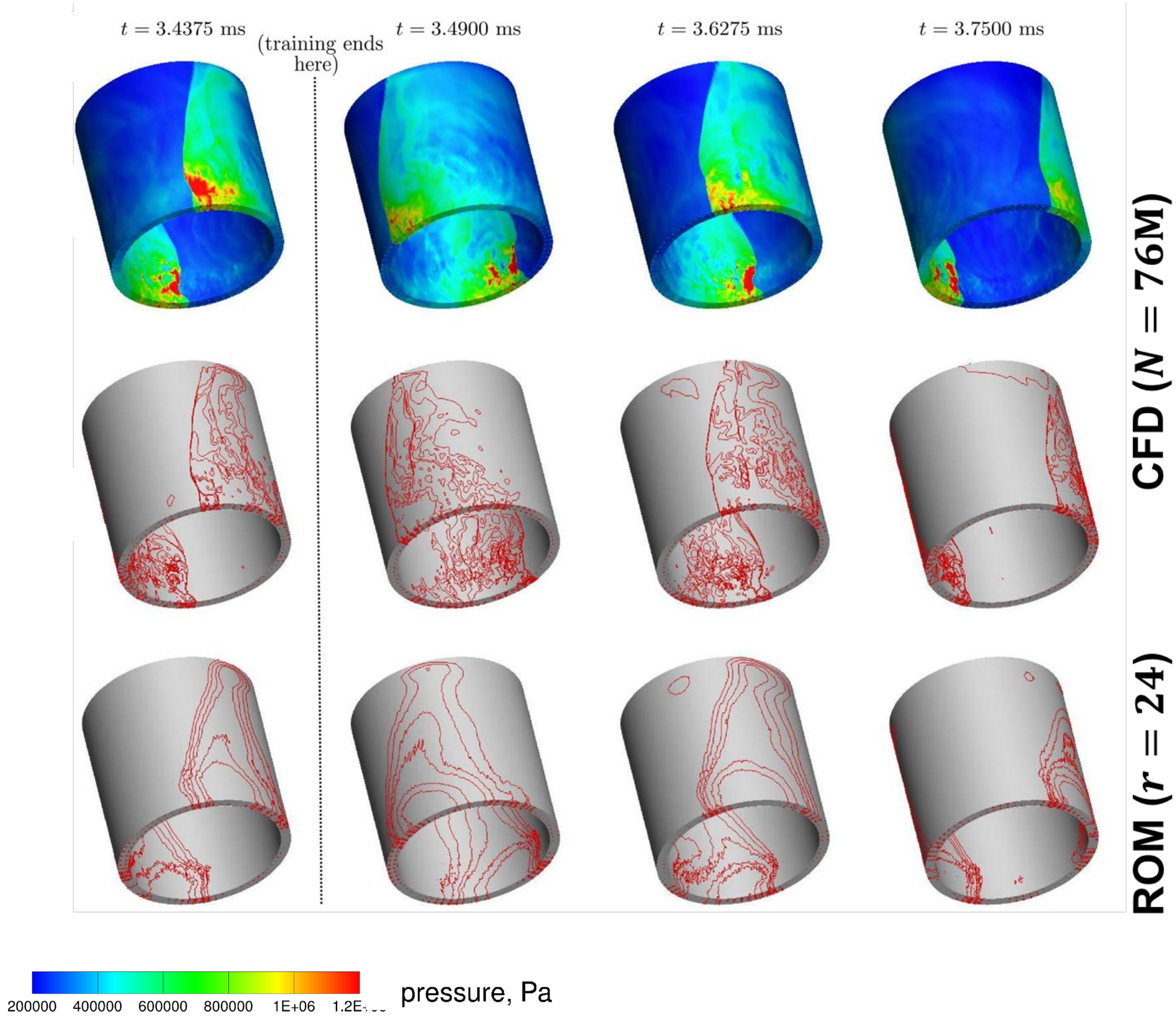
- Mathematical & Statistical Foundations
- High Performance Computing
- Scientific Machine Learning
- Computational Biology
- Artificial Intelligence
- Autonomy
- Computational Geosciences
- Computational Physics
- Optimization, Inversion & UQ
- Computational Medicine
- Computational Chemistry
- Numerical Analysis & Numerical PDEs
- Computational Materials
- Computational Engineering

| | |
|---|---|
| 25 | Research Centers/Groups |
| 27 | Principal Faculty |
| 127 | Affiliated & Core Faculty |
| 106 | Research Staff/Postdocs |
| 33 | Administrative & IT Staff |
| 115 | CSEM Students |
| 195 | GRAs (83 CSEM) |
| 22 | Visiting Faculty Fellows |
| $128M | Active Research |

# Motivation

**Proper orthogonal decomposition (POD) is a workhorse of dimension reduction and reduced-order modeling.**

**But the computational cost of generating training snapshots is often prohibitive.**

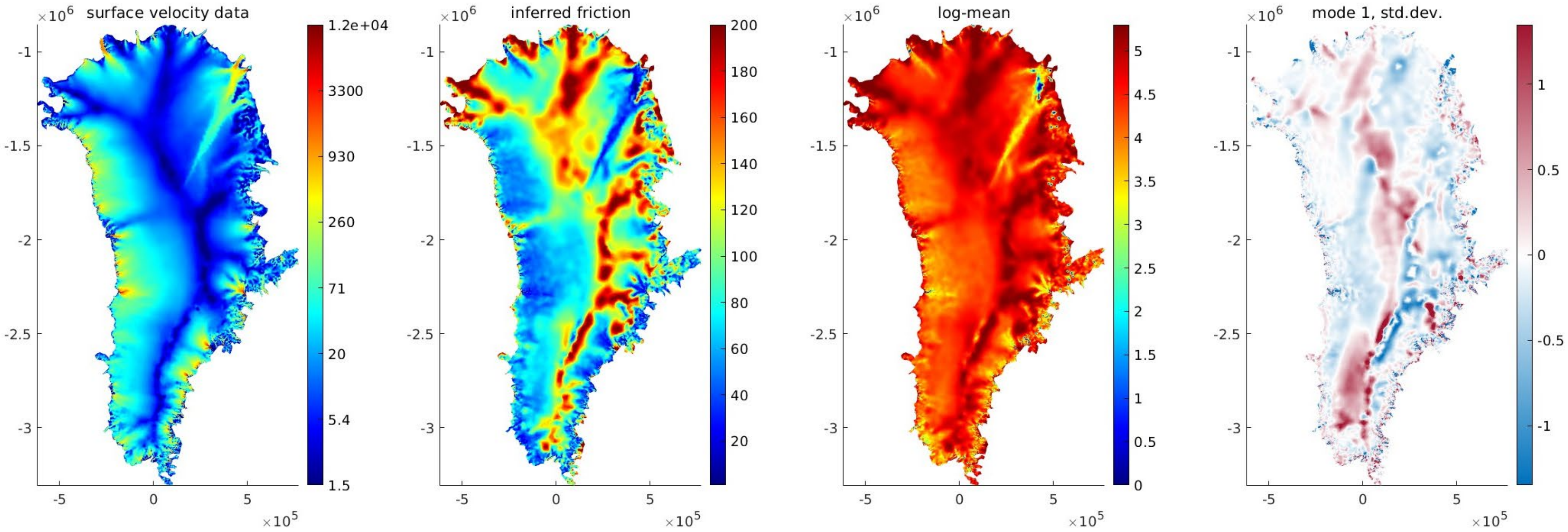


# Multifidelity POD

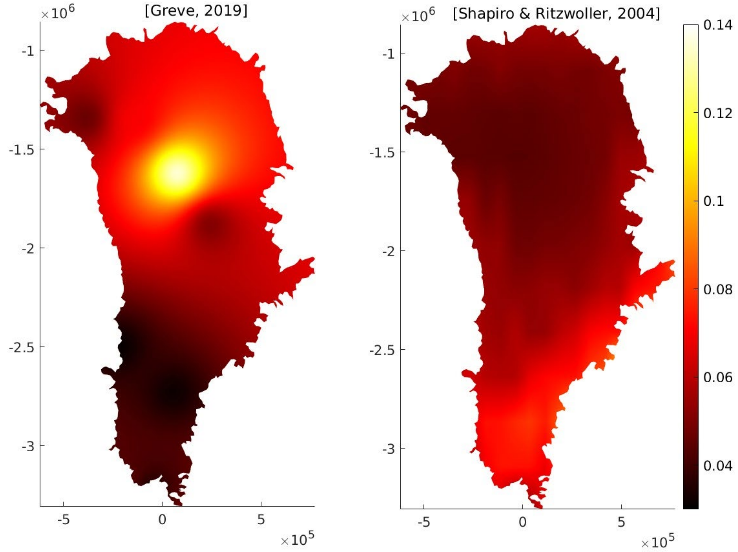
**We draw inspiration from multifidelity UQ methods, which leverage approximate models to reduce the cost of sampling.**

Aretz, Gunzburger, Morlighem & W. Multifidelity UQ for ice sheet simulations. *Computational Geosciences*, 2025.

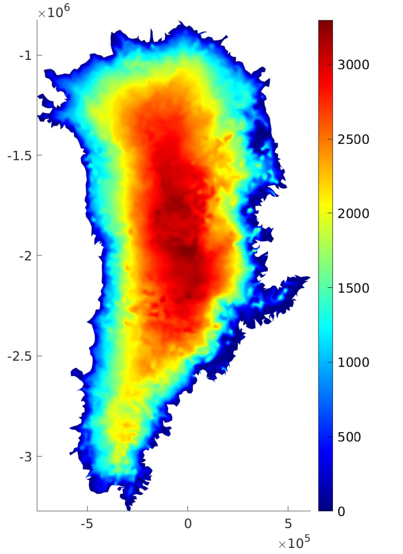
## Uncertainty in basal friction field



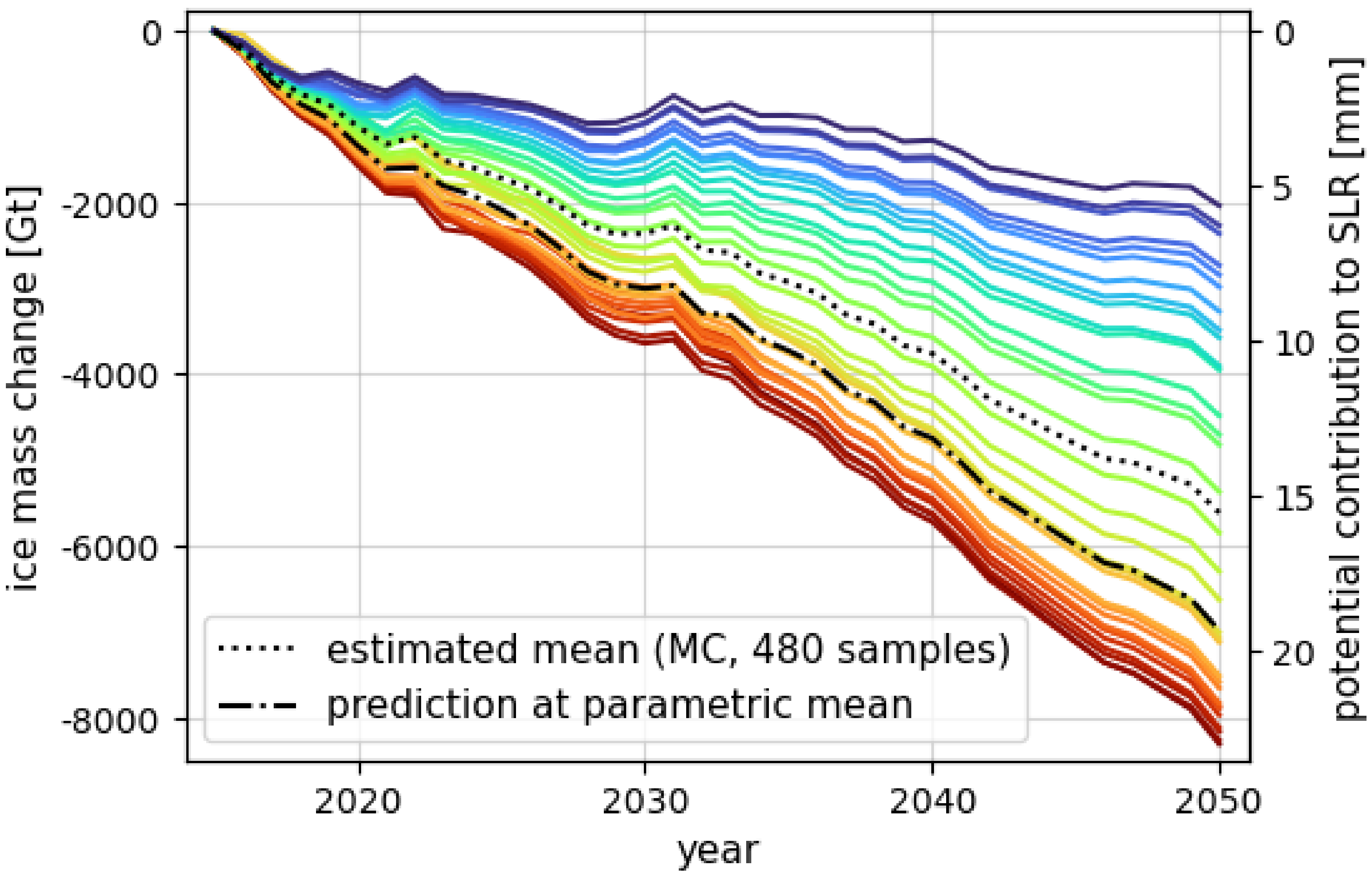

## Uncertainty in geothermal heat flux




## Uncertainty in estimated ice thickness



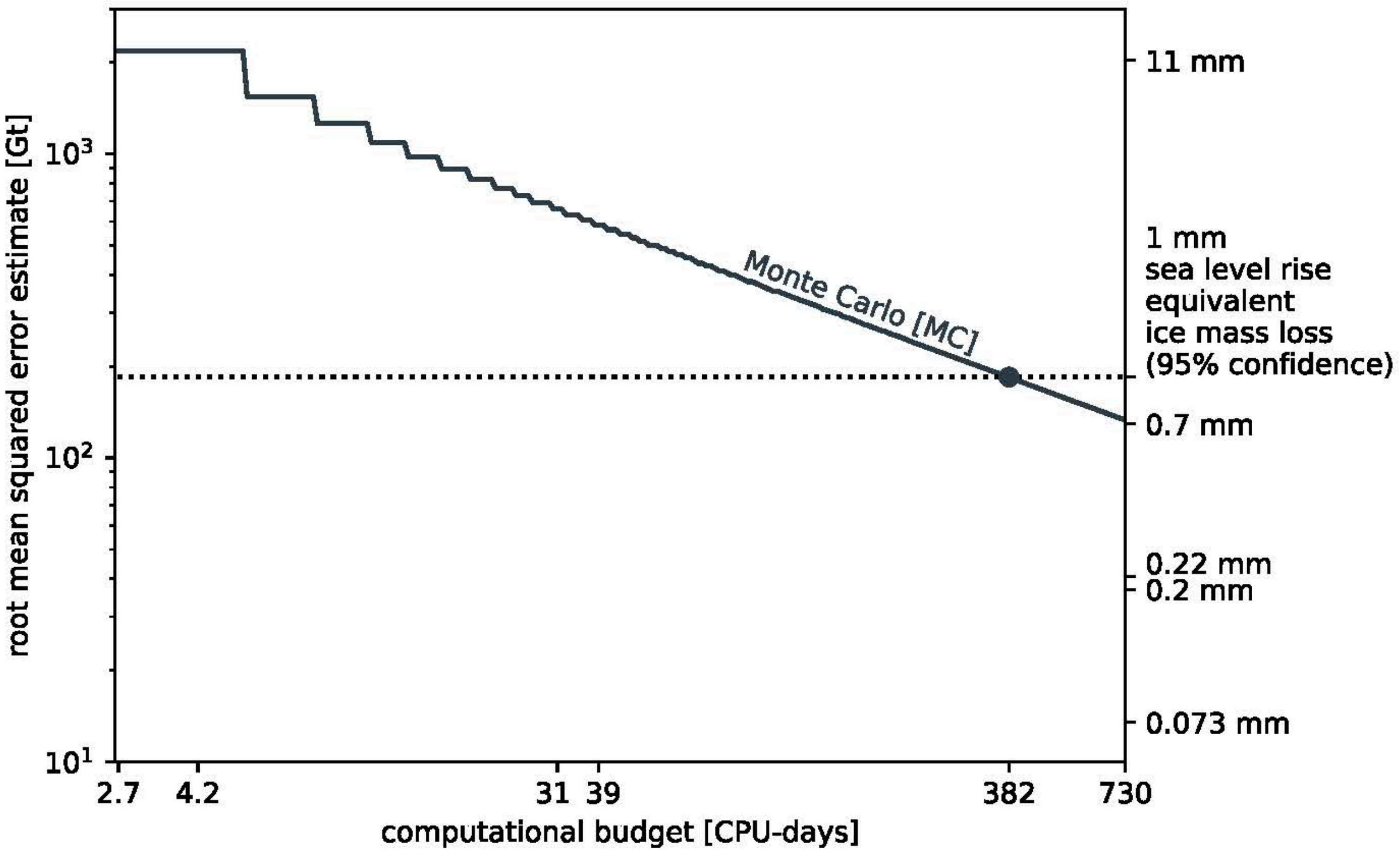
# Leads to significant uncertainty in prediction quantities of interest (ice mass change, contribution to sea level rise)



To achieve estimate accuracy of $\pm 1mm$ sea level rise (95% confidence)

Monte Carlo sampling would require

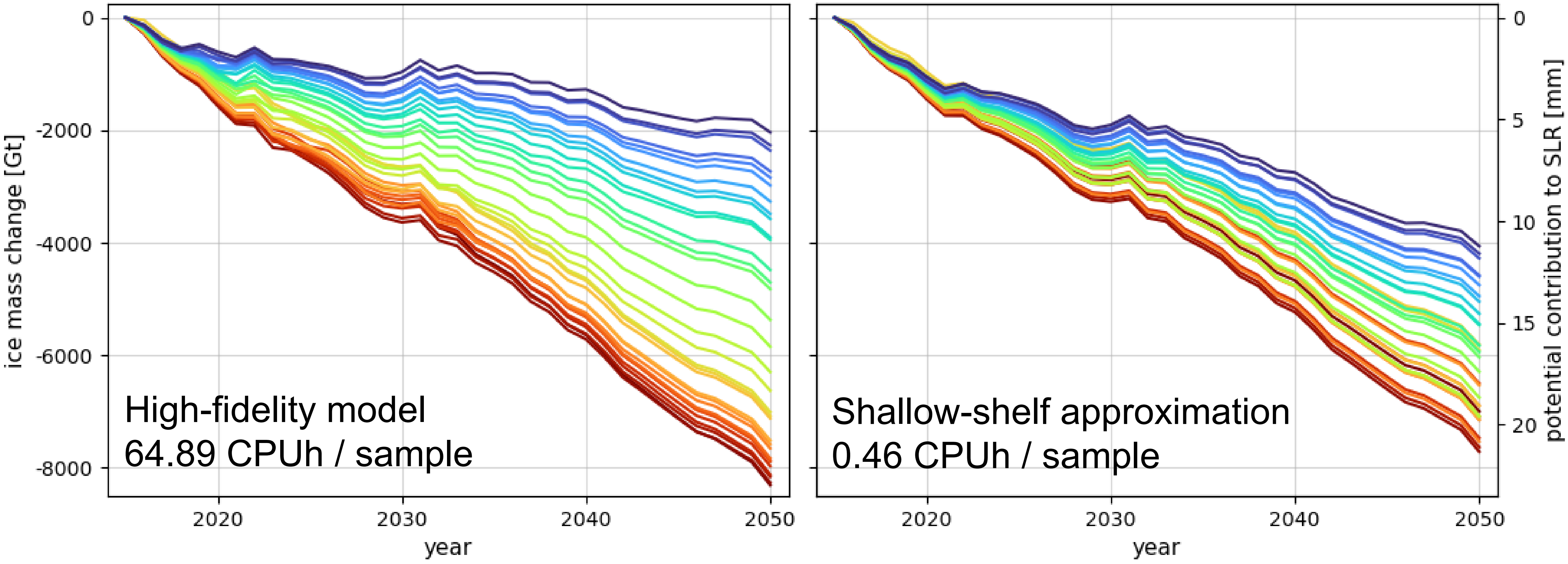
- 141 samples
- 382 CPU-days



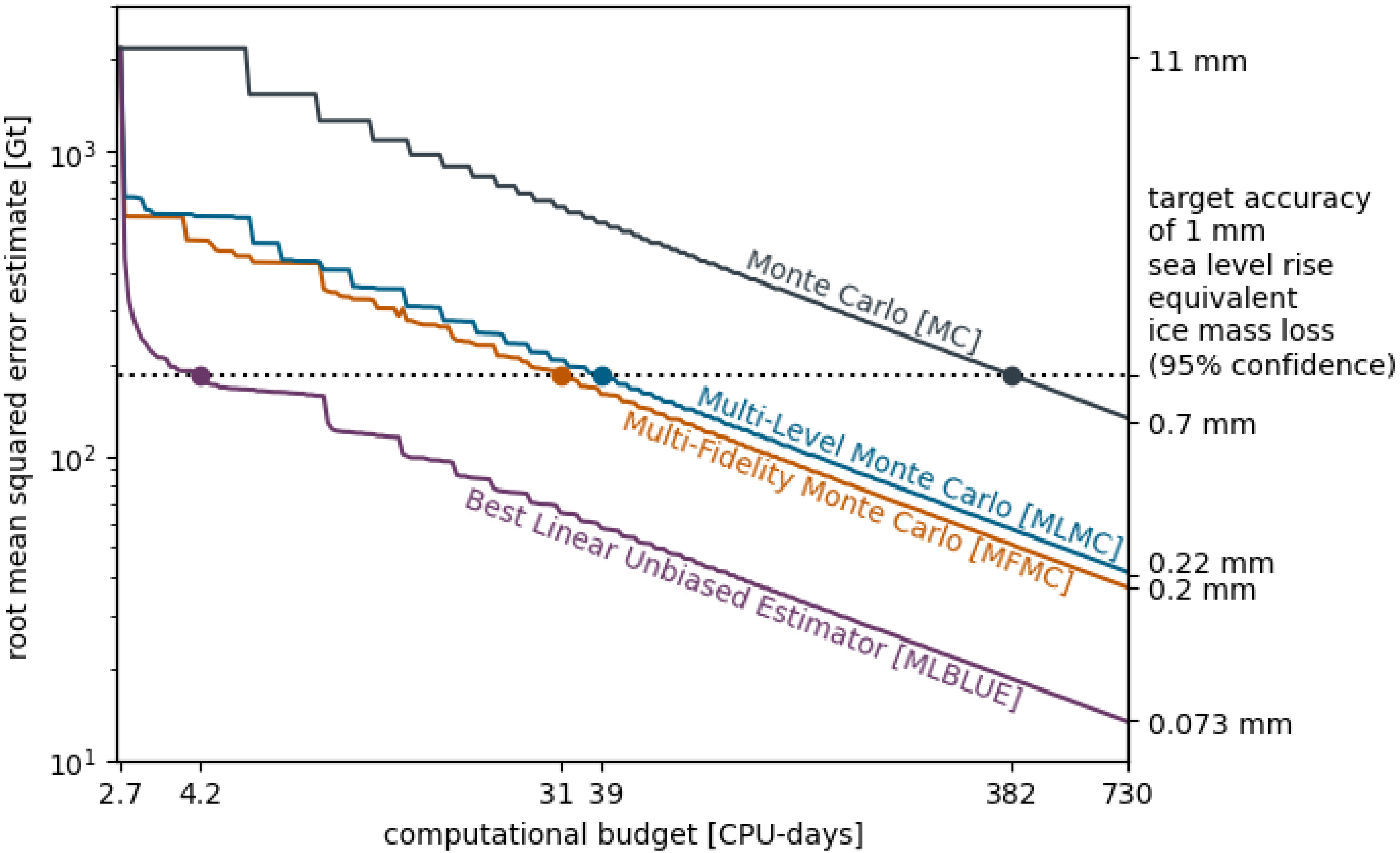

$$\mathbb{E}[S_0] \approx \frac{1}{N_{MC}} \sum_{n=1}^{N_{MC}} S_0(\omega_n) \;=:\; \mathbb{MC}\,[S_0; N_{MC}]$$

$$\mathrm{MSE} = \mathbb{E}[(\,\mathbb{E}[S_0] - \mathbb{MC}\,[S_0; N_{MC}]\,)^2] = \frac{\mathbb{V}ar[S_0]}{N_{MC}^2}$$

# Why not just replace the high-fidelity model?
# → model bias of surrogate models



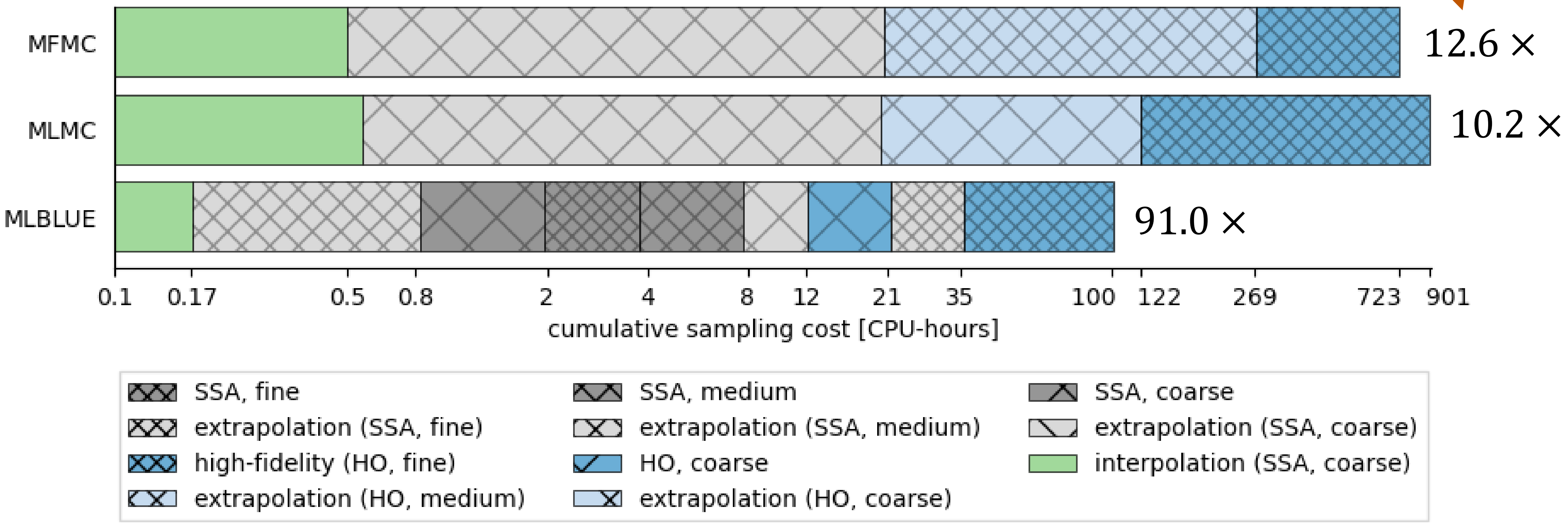
# Multifidelity methods achieve target accuracy with reduced computational budget



# Multifidelity Uncertainty Quantification

- Surrogate models are essential for speed-up
- Each multifidelity method employs surrogates differently

**Speed-up** compared to MC sampling for target accuracy $\pm 1mm$ SLR



# Why use a multifidelity formulation?

Reduced-order or surrogate model
(approximate)

Full model
(“truth”)

# Why use a multifidelity formulation?

Reduced-order or surrogate model
(approximate)

Full model
("truth")

Computationally
cheaper

Computationally
expensive

# Why use a multifidelity formulation?

Reduced-order or surrogate model
(approximate)

Full model
("truth")

Certified?

yes

- Replace full model with reduced model and solve {opt, UQ, inverse}
- Propagate error estimates on forward predictions to determine error in {opt, UQ, inverse} solutions (may be non-trivial)

# Why use a multifidelity formulation?

Reduced-order or surrogate model
(approximate)

Full model
("truth")

Certified?

no

- Replace full model with reduced model and solve {opt, UQ, inverse}
- Hope for the best

# Why use a multifidelity formulation?

Reduced-order or surrogate model
(approximate)

Full model
("truth")

Certified?

no

- Use a multifidelity formulation that invokes both the reduced model and the full model
- Trade computational cost for the ability to place guarantees on the solution of {opt, UQ, inverse}

# Why use a multifidelity formulation?

Reduced-order or surrogate model
(approximate)

Full model
("truth")

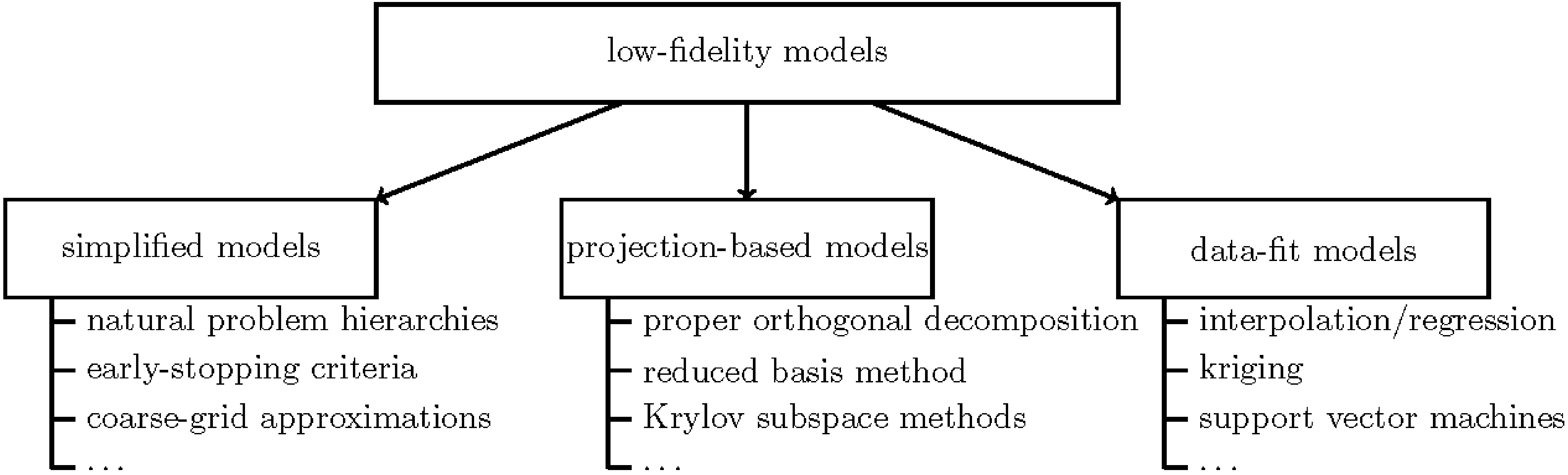
Certified?

no

- Use a multifidelity formulation that invokes both the reduced model and the full model
- Trade computational cost for the ability to place guarantees on the solution of {opt, UQ, inverse}
- Certify the solution of {opt, UQ, inverse} even in the absence of guarantees on the reduced model itself

# Multifidelity models come in different forms

Covering a range of different resolutions, scales, reduced order, modeling assumptions, etc.




From Peherstorfer, W, Gunzburger. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. *SIAM Review*, 2018

# Multifidelity Monte Carlo

leveraging multiple approximate models to estimate statistics of the high-fidelity model

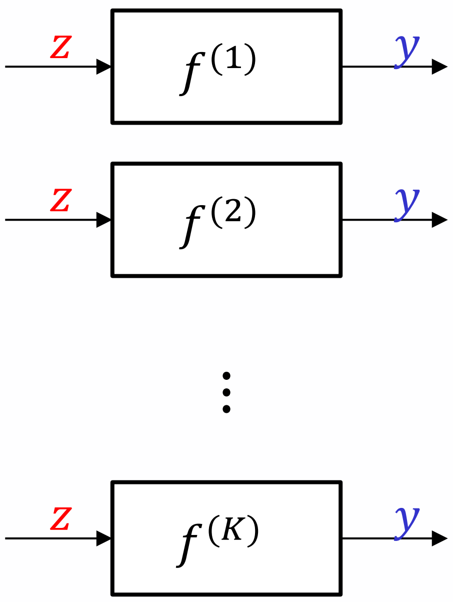
Ng & W. AIAA 2012, IJNME 2014
Peherstorfer, W, Gunzburger, SISC 2016

- high-fidelity model

  $f^{(1)}: \mathcal{Z} \rightarrow \mathcal{Y}$ ("truth")

- *K – 1* surrogate models

  $$f^{(2)}, \ldots, f^{(K)}: \mathcal{Z} \rightarrow \mathcal{Y}$$




- model $f^{(i)}$ has cost $w_i$

- $m_i$ evaluations for model $i$, with

  $$m_1 \leq m_2 \leq \ldots \leq m_K$$

- Models do not necessarily form a hierarchy (cf. multi-level Monte Carlo)
  - How to combine models?
  - How to balance evaluations among them?

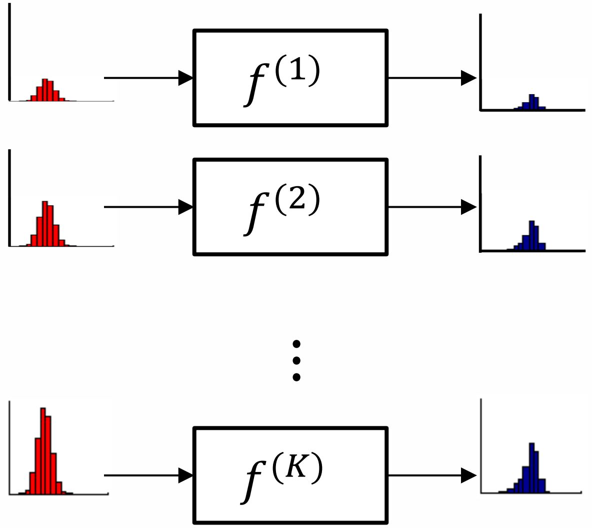
# Multifidelity Monte Carlo

leveraging multiple approximate models to estimate statistics of the high-fidelity model

Ng & W. AIAA 2012, IJNME 2014

- Draw $m_k$ realizations $z_1, \dots, z_{m_K}$ of $Z$ and evaluate $f^{(i)}$:

  $f^{(i)}(z_1), \dots, f^{(i)}(z_{m_i})$




- Compute mean estimators

  $\bar{y}^{(1)}_{m_1}, \dots, \bar{y}^{(k)}_{m_k}$ and $\bar{y}^{(2)}_{m_1}, \dots, \bar{y}^{(k)}_{m_{k-1}}$

- MFMC estimator:

$$\hat{s} = \bar{y}^{(1)}_{m_1} + \sum_{i=2}^{K} \alpha_i \left( \bar{y}^{(i)}_{m_i} - \bar{y}^{(i)}_{m_{i-1}} \right)$$

- $\hat{s}$: MFMC estimate for the mean
- $\bar{y}^{(1)}_{m_1}$: mean estimate using $m_1$ evaluations of truth model
- $\bar{y}^{(i)}_{m_i}$: mean estimate using $m_i$ evaluations of model $i$
- $\bar{y}^{(i)}_{m_{i-1}}$: mean estimate using $m_{i-1}$ evaluations of model $i$

# Multifidelity Monte Carlo

optimally allocate computational budget across *K* models

Peherstorfer, W., Gunzburger, *SISC*, 2016

- MFMC estimator

$$\hat{s} = \bar{y}^{(1)}_{m_1} + \sum_{i=2}^{K} \alpha_i \left( \bar{y}^{(i)}_{m_i} - \bar{y}^{(i)}_{m_{i-1}} \right)$$

$\hat{s}$: MFMC estimate for the mean

$\bar{y}^{(1)}_{m_1}$: mean estimate using $m_1$ evaluations of truth model

$\bar{y}^{(i)}_{m_i}$: mean estimate using $m_i$ evaluations of model $i$

$\bar{y}^{(i)}_{m_{i-1}}$: mean estimate using $m_{i-1}$ evaluations of model $i$

- MFMC estimator is **unbiased**, even with no error bounds for surrogates: $E[\hat{s}] = s$

- MSE is given by $\mathrm{Var}[\hat{s}]$

$$\mathrm{Var}[\hat{s}] = \frac{\sigma_1^2}{m_1} + \sum_{i=2}^{K} \left( \frac{1}{m_{i-1}} - \frac{1}{m_i} \right) \left( \alpha_i^2 \sigma_i^2 - 2\alpha_i \rho_i \sigma_i \sigma_1 \right)$$

  - $\sigma_i^2$ is variance of $f^{(i)}(Z)$
  - $\rho_i$ is correlation coefficient between $f^{(1)}(Z)$ and $f^{(i)}(Z)$

# Multifidelity Monte Carlo

We optimally balance the number of model evaluations to obtain the best multifidelity estimator given the computational budget

- The costs of the MFMC estimator are: $c(\hat{s}) = \sum_{i=1}^{K} w_i m_i$

- Minimize the MSE of the MFMC estimator for a given computational budget $p$. Leads to optimization problem

$$\begin{aligned} \min_{m \in \mathbb{R}^K, \alpha_2, \dots, \alpha_K \in \mathbb{R}} \quad & \mathrm{Var}[\hat{s}] \\ \text{such that} \quad & m_{i-1} \leq m_i, i = 2, \dots, K \\ & 0 \leq m_1 \\ & c(\hat{s}) = w^T m = p \end{aligned}$$

- Distinguishing features of MFMC method:
  - optimal selection of the number of model evaluations $m_1 \leq m_2 \leq \dots \leq m_K$ and of coefficients $\alpha_2, \dots, \alpha_K$
  - applicable to general information sources (e.g., any type of surrogate model, database curve fits, etc.)

# Let’s come back to POD

# Proper Orthogonal Decomposition

Given a set of snapshots, find the POD basis and the associated POD singular values

- Consider $m$ **snapshots** $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m \in \mathbb{R}^N$ (solutions at selected times or parameter values)

- Form the snapshot matrix $\mathbf{X} = [\mathbf{x}_1 \ \ \mathbf{x}_2 \ \dots \ \mathbf{x}_m]$

- Choose the $r$ basis vectors $\mathbf{V} = [\mathbf{V}_1 \ \ \mathbf{V}_2 \ \dots \ \mathbf{V}_r]$ to be left **singular vectors** of the snapshot matrix, with singular values

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r \geq \sigma_{r+1} \geq \cdots \geq \sigma_m \geq 0$$

- This is the optimal projection in a least squares sense:

$$\min_{\mathbf{V}} \sum_{i=1}^{m} ||\mathbf{x}_i - \mathbf{V}\mathbf{V}^T\mathbf{x}_i||_2^2 = \sum_{i=r+1}^{m} \sigma_i^2$$

# Proper Orthogonal Decomposition

Given a set of snapshots, find the POD basis and the associated POD singular values

POD basis vectors are

- the left singular vectors of the snapshot matrix $\mathbf{X} = [\mathbf{x}_1 \;\; \mathbf{x}_2 \;\ldots\; \mathbf{x}_m]$
- the eigenvectors of $\mathbf{X}\mathbf{X}^\top$

$$\mathbf{X}\mathbf{X}^\top = \sum_{i=1}^{m} \mathbf{x}_i \mathbf{x}_i^\top \qquad (N \times N)$$

- linear combinations of the snapshots $\mathbf{U}\mathbf{X}$, where $\mathbf{U}$ are the eigenvectors of $\mathbf{X}^\top\mathbf{X}$

$$\mathbf{X}^\top\mathbf{X} = \sum_{i=1}^{m} \mathbf{x}_i^\top \mathbf{x}_i \qquad (m \times m)$$

# Multifidelity Proper Orthogonal Decomposition

Aretz & W.
Multifidelity Proper Orthogonal Decomposition.
*In preparation.*

Consider two fidelity levels sampled over parameters $\boldsymbol{\theta}$:

- High-fidelity snapshots

$$\mathbf{X}_{\mathrm{hi}} = \left[\mathbf{x}(\boldsymbol{\theta}_1) \quad \mathbf{x}(\boldsymbol{\theta}_2) \ \dots \ \mathbf{x}(\boldsymbol{\theta}_{m_1})\right]$$

- Low-fidelity snapshots (mapped to common reference)

$$\mathbf{X}_{\mathrm{lo}} = \left[\tilde{\boldsymbol{x}}(\boldsymbol{\theta}_1) \ \tilde{\boldsymbol{x}}(\boldsymbol{\theta}_2) \ \dots \ \tilde{\boldsymbol{x}}(\boldsymbol{\theta}_{m_1})\right]$$

$$\mathbf{X}_{\mathrm{lo}}^{+} = \left[\tilde{\boldsymbol{x}}(\boldsymbol{\theta}_{m_1+1}) \ \tilde{\boldsymbol{x}}(\boldsymbol{\theta}_{m_1+2}) \ \dots \ \tilde{\boldsymbol{x}}(\boldsymbol{\theta}_{m_2})\right]$$

Multifidelity equivalent of $\mathcal{J}_{\mathrm{POD}}(\mathbf{V}) = \frac{1}{m}\|\mathbf{X} - \mathbf{V}\mathbf{V}^{\top}\mathbf{X}\|_2^2$ is

$$\mathcal{J}_{mf}(\mathbf{V}, \alpha) = \frac{1}{m_1}\|\mathbf{X}_{\mathrm{hi}} - \mathbf{V}\mathbf{V}^{\top}\mathbf{X}_{\mathrm{hi}}\|_2^2 + \frac{\alpha}{m_2}\left\|[\mathbf{X}_{\mathrm{lo}}\ \mathbf{X}_{\mathrm{lo}}^{+}] - \mathbf{V}\mathbf{V}^{\top}[\mathbf{X}_{\mathrm{lo}}\ \mathbf{X}_{\mathrm{lo}}^{+}]\right\|_2^2 - \frac{\alpha}{m_1}\|\mathbf{X}_{\mathrm{lo}} - \mathbf{V}\mathbf{V}^{\top}\mathbf{X}_{\mathrm{lo}}\|_2^2$$

# Multifidelity Proper Orthogonal Decomposition

Aretz & W.
Multifidelity Proper Orthogonal Decomposition.
*In preparation.*

$$\mathcal{J}_{mf}(\mathbf{V},\alpha) = \frac{1}{m_1}\sum_{i=1}^{m_1}\|\mathbf{x}(\boldsymbol{\theta}_i) - \mathbf{V}\mathbf{V}^\top\mathbf{x}(\boldsymbol{\theta}_i)\|_2^2$$
$$+\frac{\alpha}{m_2}\sum_{i=1}^{m_2}\|\tilde{\mathbf{x}}(\boldsymbol{\theta}_i) - \mathbf{V}\mathbf{V}^\top\tilde{\mathbf{x}}(\boldsymbol{\theta}_i)\|_2^2 - \frac{\alpha}{m_1}\sum_{i=1}^{m_1}\|\tilde{\mathbf{x}}(\boldsymbol{\theta}_i) - \mathbf{V}\mathbf{V}^\top\tilde{\mathbf{x}}(\boldsymbol{\theta}_i)\|_2^2$$

We compute the mfPOD basis as the eigenvectors of

$$\mathcal{J}_{mf}\mathbf{v}_i = \lambda_i\mathbf{v}_i$$

Glossing over important details:

- Choosing $\alpha$ (control variate coefficients)
- Choosing $m_1, m_2$ (optimal sample allocations across models)
- Mapping $\tilde{\mathbf{x}}$ and $\mathbf{x}$ to have the same representation
- $\mathcal{J}_{mf}$ not guaranteed to be positive semidefinite

# Illustrative Example: 1D advection diffusion

$$-\frac{1}{\theta}\Delta u - u_x = 1 \qquad \text{on } \Omega := (0,1)$$

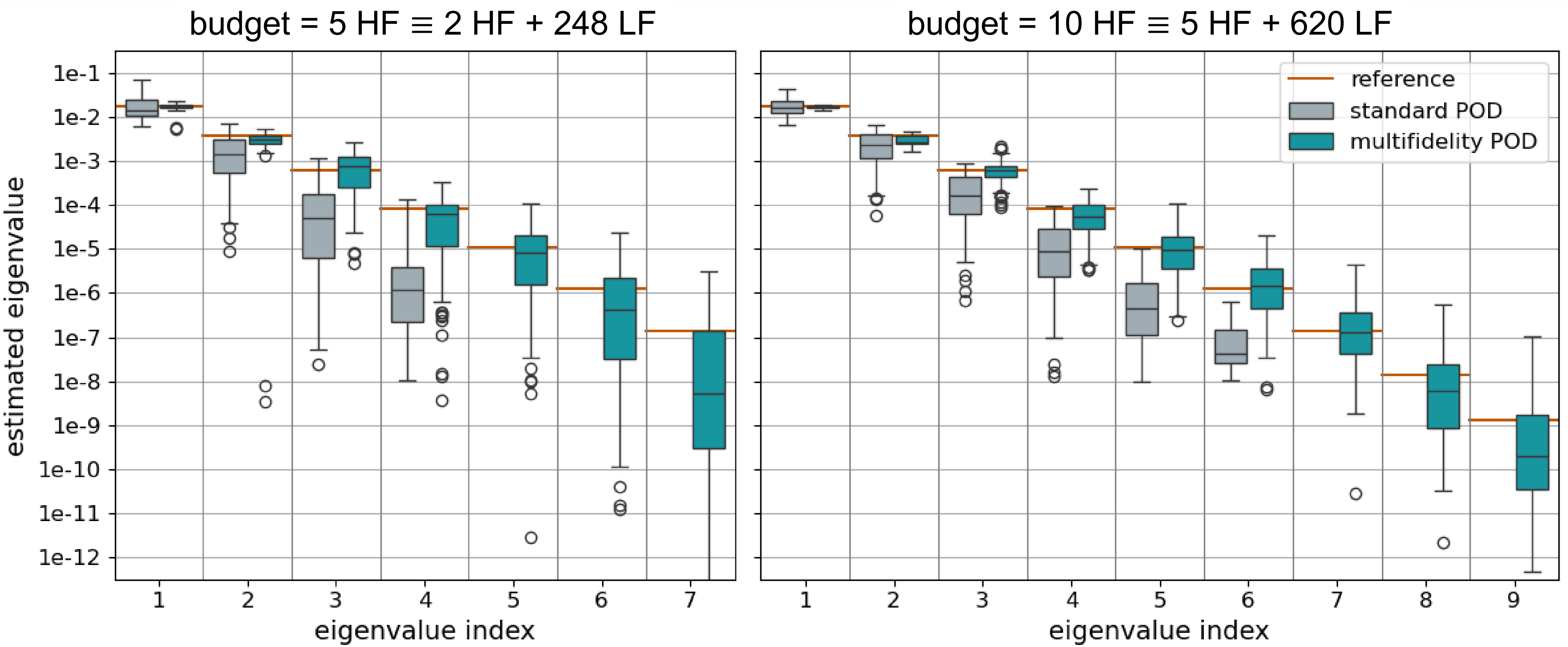
$$u(0) = 1, u(1) = 0$$

High fidelity FEM: 4097 dof

Low fidelity FEM: 33 dof

$\theta \sim \mathcal{U}(1,100)$

Reference: 100,000 snapshots

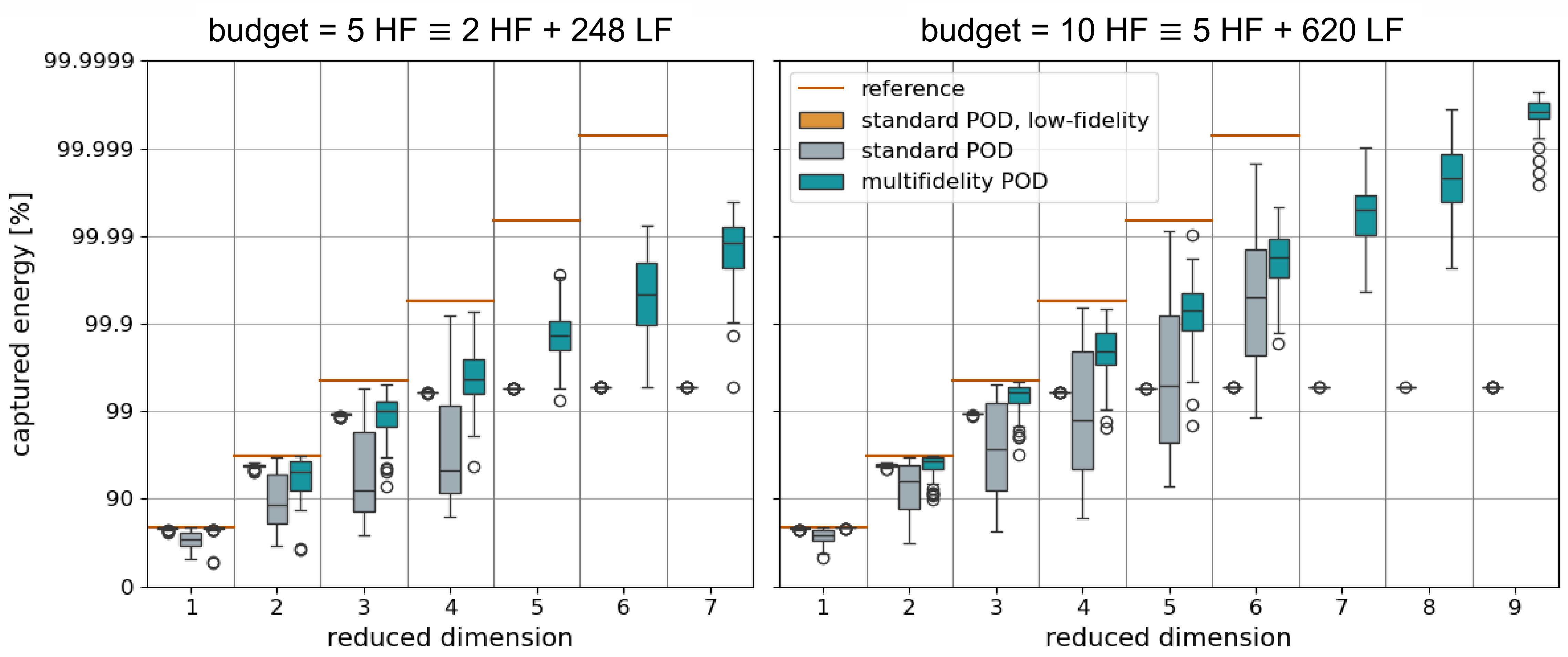


# Illustrative Example: 1D advection diffusion

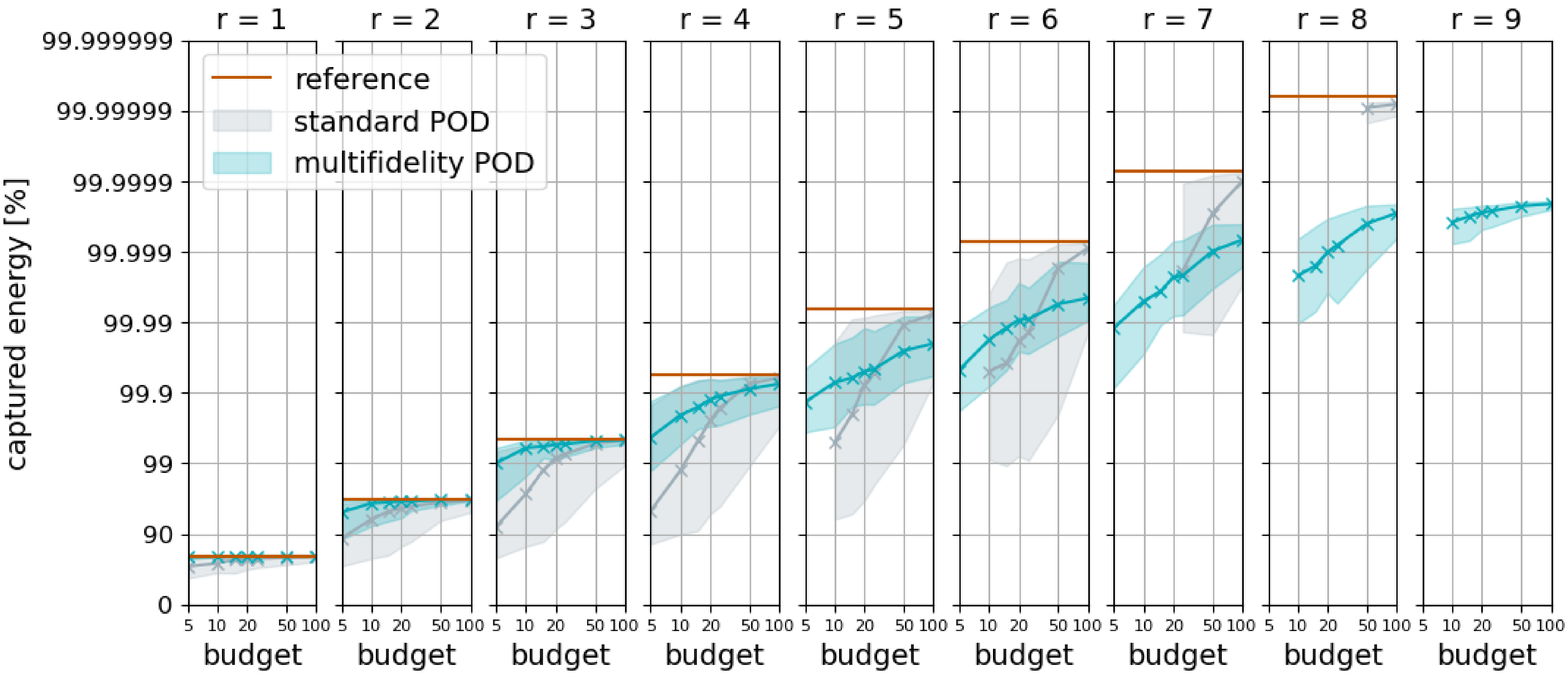
$$-\frac{1}{\theta}\Delta u - u_x = 1 \qquad \text{on } \Omega := (0,1)$$

High fidelity FEM: 4097 dof
Low fidelity FEM: 33 dof

$\theta \sim \mathcal{U}(1{,}100)$
Reference: 100,000 snapshots



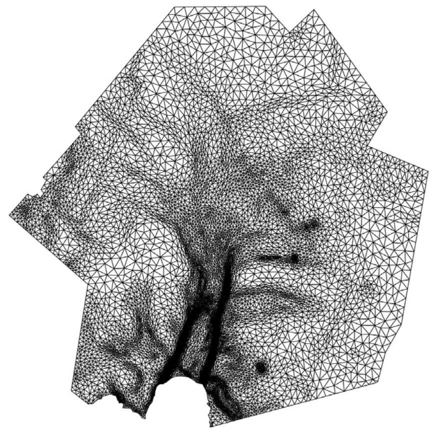
# Illustrative Example: 1D advection diffusion



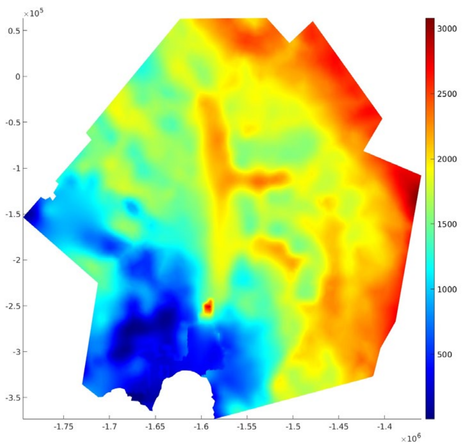
# Example: Ice sheet dynamics of Pine Island Glacier

- Ice thickness equation models the dynamics of ice thickness $h(\boldsymbol{x}, t)$ over the domain

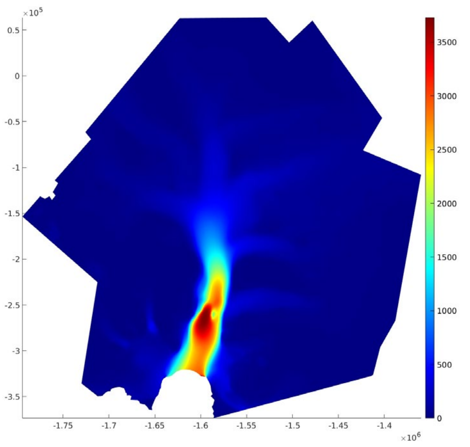
$$\frac{\mathrm{d}h}{\mathrm{d}t} = -\nabla \cdot h\boldsymbol{v} + \nabla \cdot (\boldsymbol{D}\, \nabla h) + m_s - m_b(h, \alpha)$$



- Shallow shelf approximation for depth-averaged velocity (two velocity components)
- Varying ice melt rate parameter (1 to 100 m / yr)
- High-fidelity: 10,635 dofs / variable (total 31,905 per time step)
- Surrogate: 2,716 dofs / variable (total 8,148 per time step), 9x faster



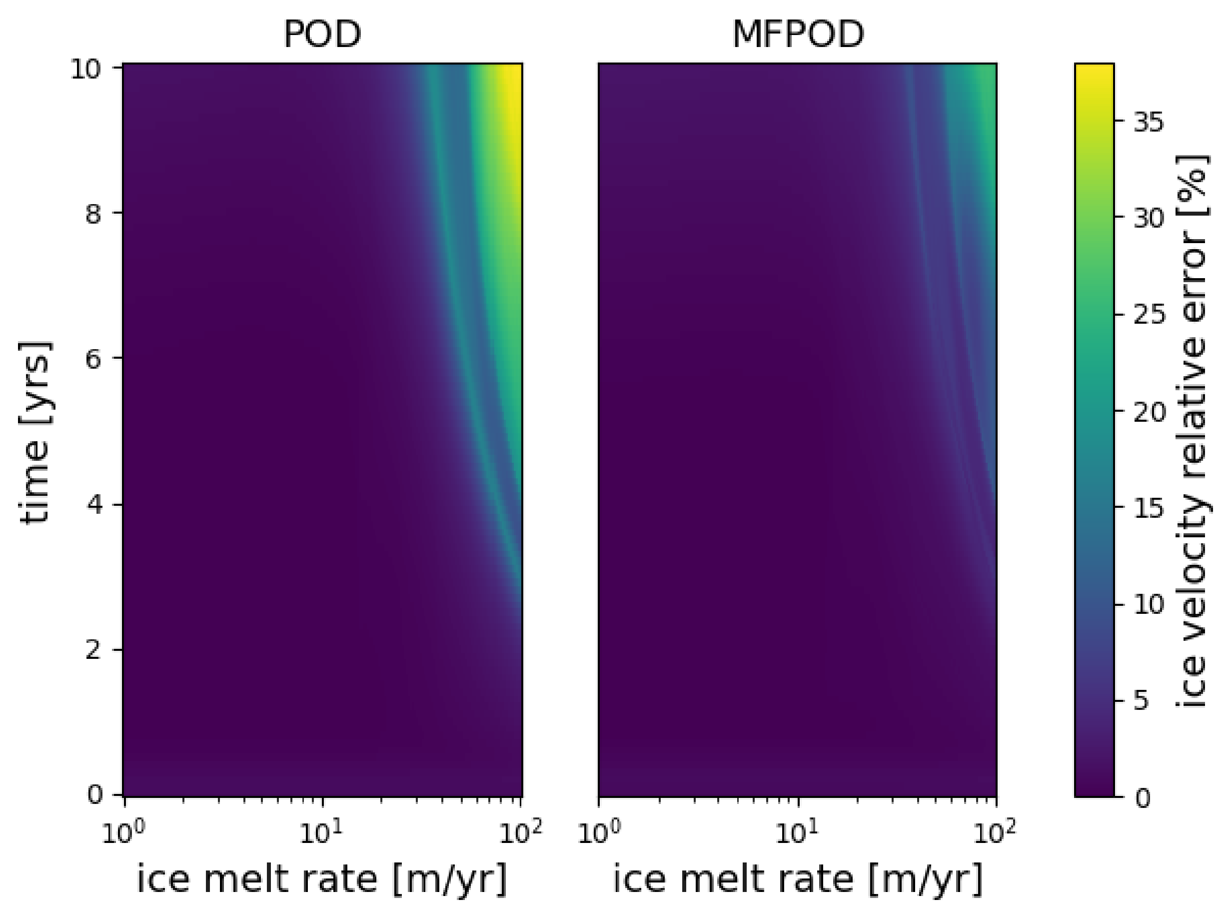
initial ice thickness



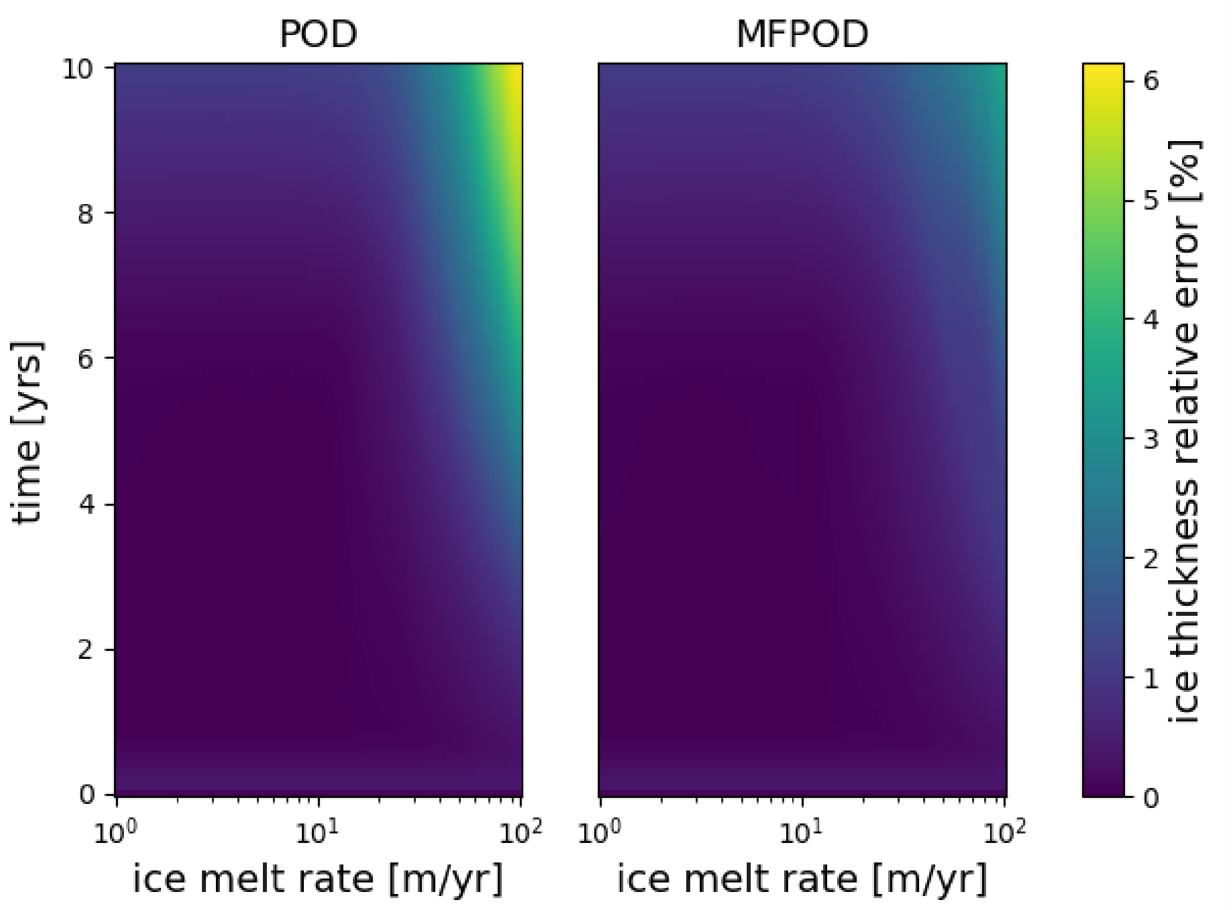
initial velocity

# Example: Ice sheet dynamics of Pine Island Glacier

POD trained on 5 random parameter samples and time t = 0, 1, 2, 3, 4, 5 (30 data points total)



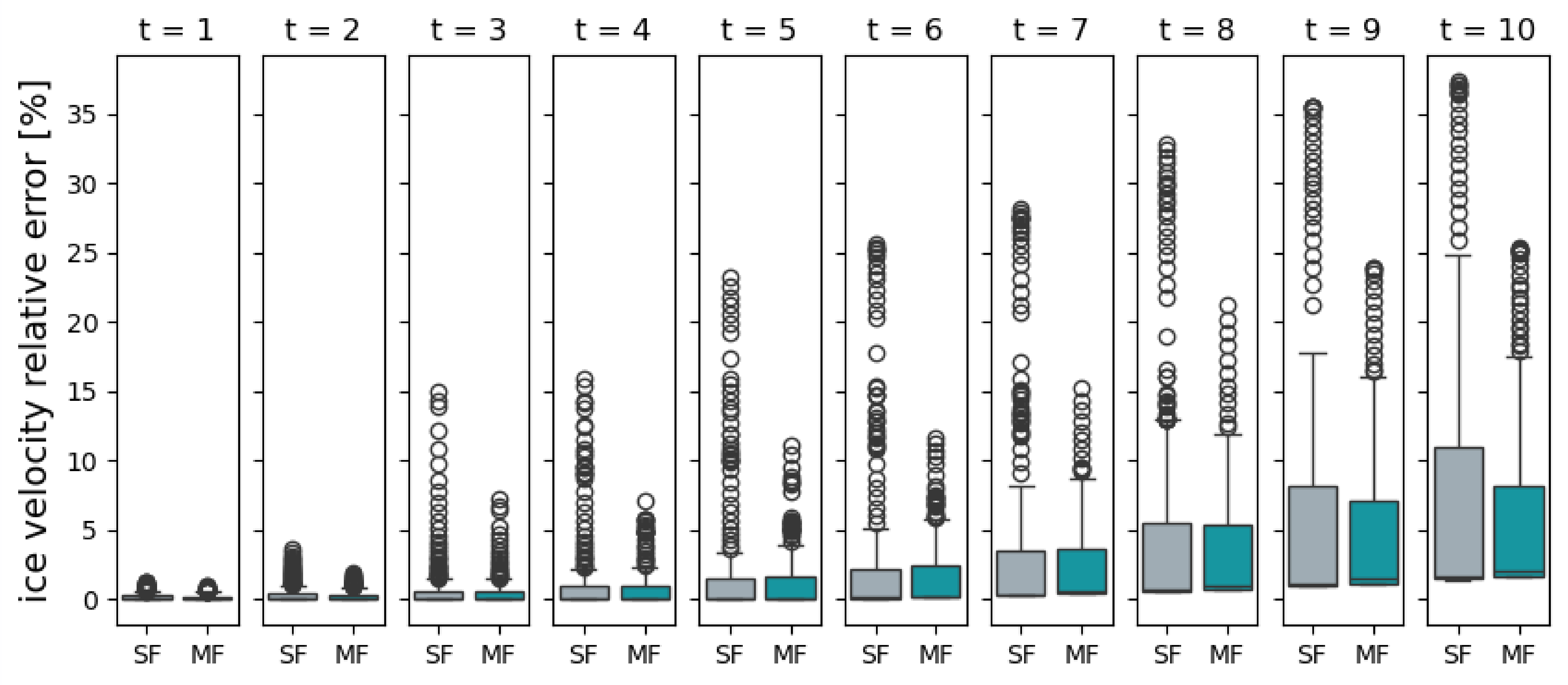

MFPOD trained on 3 high-fidelity samples and 18 low-fidelity samples



# Example: Ice sheet dynamics of Pine Island Glacier

POD trained on 5 random parameter samples and time t = 0, 1, 2, 3, 4, 5 (30 data points total)

MFPOD trained on 3 high-fidelity samples and 18 low-fidelity samples

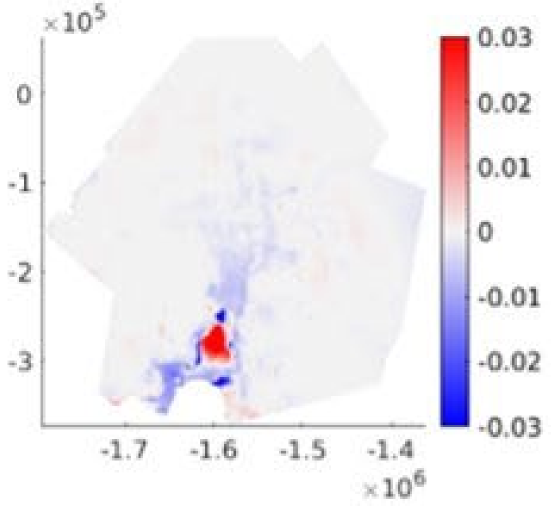


# Example: Ice sheet dynamics of Pine Island Glacier

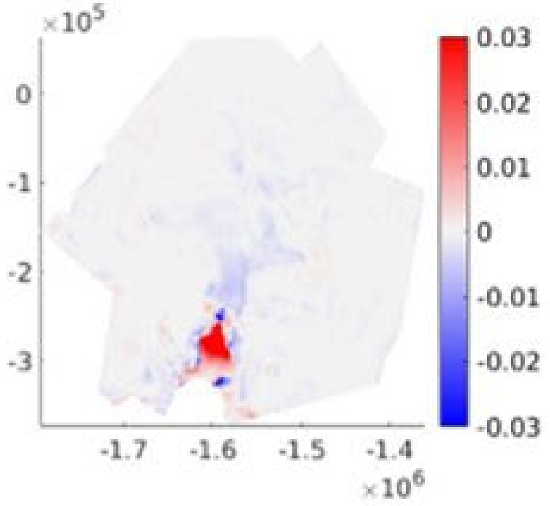
POD trained on 5 random parameter samples and time t = 0, 1, 2, 3, 4, 5 (30 data points total)

MFPOD trained on 3 high-fidelity samples and 18 low-fidelity samples

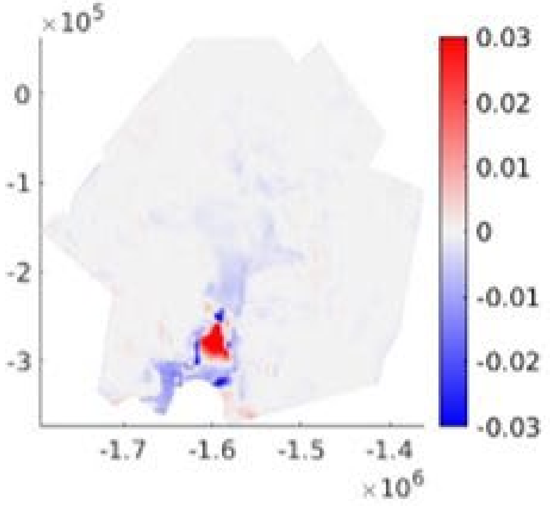
## POD modes: Ice thickness



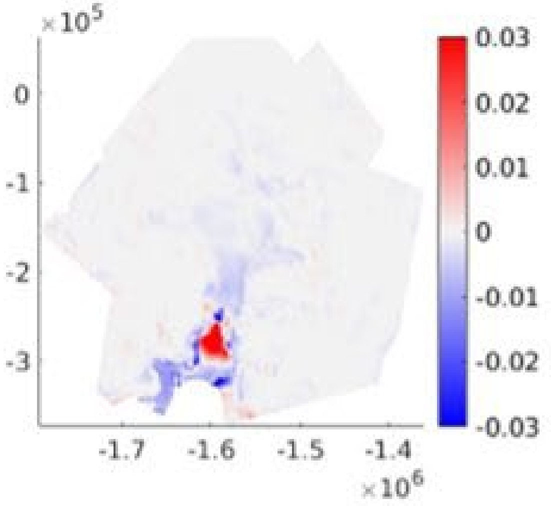
low-fidelity



high-fidelity

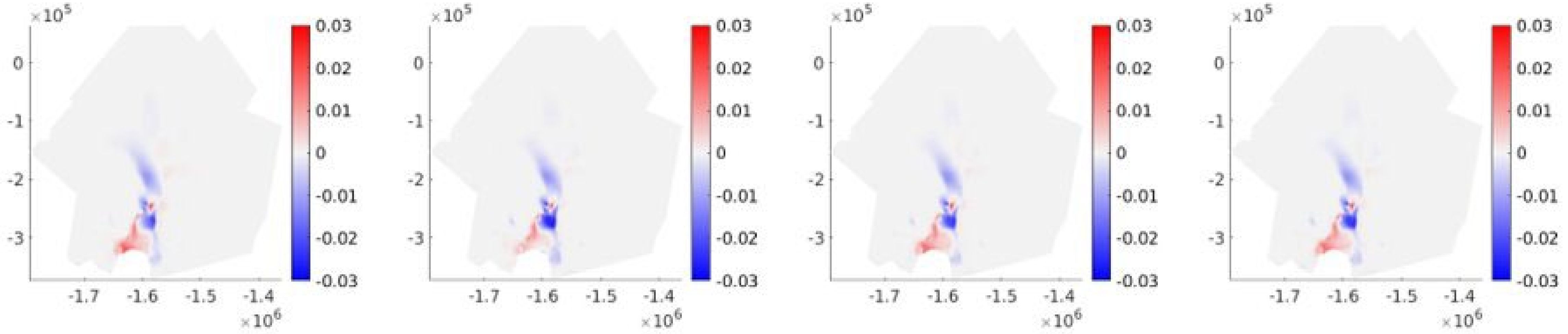


multifidelity



high-fidelity reference

# POD modes: Velocity



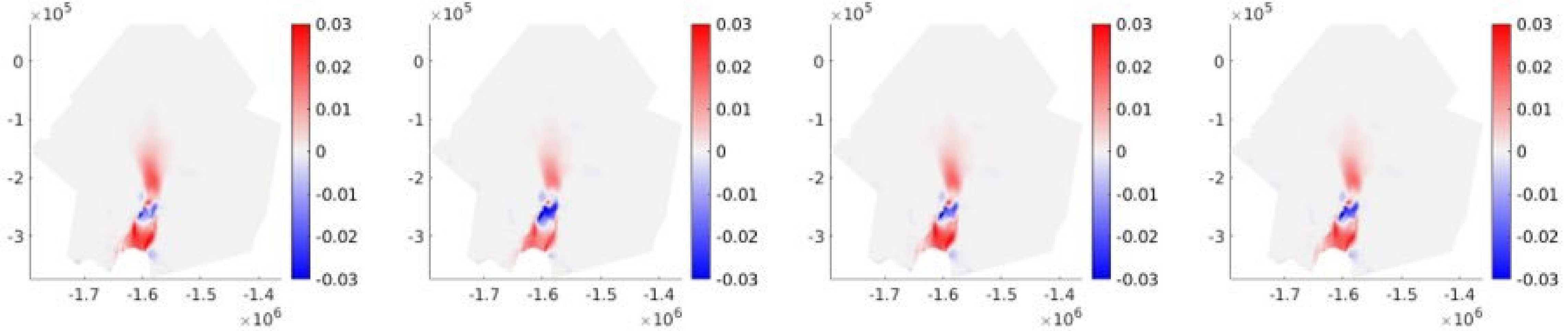

x-coord

low-fidelity

high-fidelity

multifidelity

high-fidelity reference




y-coord

# Summary

- Multifidelity POD addresses the expense of generating training data.
- It has the potential provide orders of magnitude in computational speedups together with theoretical guarantees of accuracy.
- We would love to try it out on a challenging real-world fusion example...