SAND2022-13812C

Sandia National Laboratories

Hyper-differential sensitivity analysis with respect to model discrepancy

RISE: Robust Interpretable Scalable Efficient

Joseph Hart and Bart van Bloemen Waanders

Sandia National Laboratories

Scientific Machine Learning for Complex Systems: Beyond Forward Simulation to Inference and Optimization

October 10, 2022

U.S. DEPARTMENT OF ENERGY NNSA National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

SAND2022-XXXX

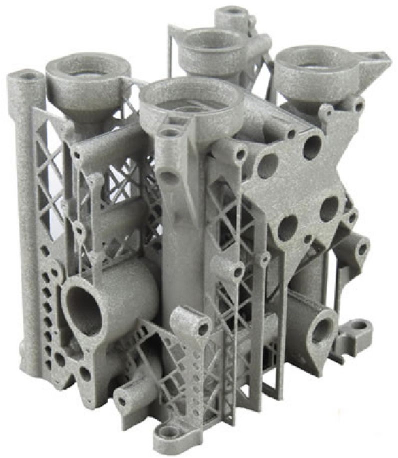
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Decision-making for complex systems

Goals
- Calibrate models
- Determine optimal designs/controllers
- Quantify uncertainty

Challenges:
- Computation cost
- Incomplete models and data
- Large decision and uncertainty spaces

Strategy:
- Fuse data, models, decisions with:
    - **R**obust performance to address uncertainties
    - **I**nterpretability to support decision-making
    - **S**calable methods to overcome high dimensionality
    - **E**fficient algorithms to mitigate cost

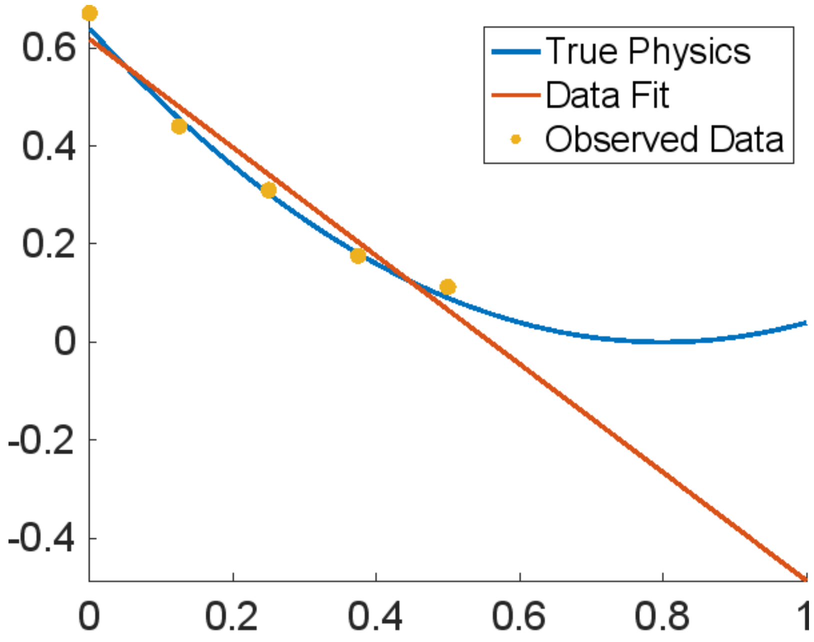


Scientific ML – Data, physics, and decisions

- Data summarizes what has happened (past)
- Data is noisy and sparse

- Physics describes what will happen (future)
- Models are approximate, uncertain, and incomplete

- Predictions are plagued by uncertainties
- Computational support for decisions is limited



$$\frac{\partial u}{\partial t} + \mathbf{v} \cdot \nabla u = f$$

$$\mathbf{v} = ???$$

- Accurate prediction requires fusing information sources

- SciML should synthesize physics, data, and decisions

PDE/ODE-constrained optimization foundation

$$\min_{u,z} \quad J(u,z)$$

$$\text{s.t.} \quad c(u,z) = 0$$

where

$J =$ objective function

$c :=$ state equations PDEs or ODEs

$u :=$ state variables

$z :=$ decision variables (design, control, inversion)

Challenges:

- Infinite dimensional state and possibly decision variables
- Matrix-free linear algebra
- Iterative solvers needed
- Preconditioning essential
- Parallelism necessary
- Large sensitivity requirements

PDE-constrained optimization solution

$$\min_{u,z} \quad J(u,z)$$

$$\text{s.t.} \quad c(u,z) = 0$$

Define Lagrangian function: $\mathcal{L}(u,z,\lambda) = J(u,z) + \langle \lambda, c(u,z) \rangle$

where λ is the Lagrange multiplier

Computation of the reduced gradient:

$c(u,z) = 0$	state equation
$c_u^*(u,z)\lambda = -J_u(u,z)$	adjoint equation
$\hat{J}_z(z) = c_z^*(u,z)\lambda + J_z(u,z)$	gradient equation

Newton step:

$$\nabla^2 \hat{J}(z_k)\delta z = -\nabla \hat{J}(z_k)$$

$$z_{k+1} := z_k + \delta z$$

Solution strategy:
- Reduced space
- Adjoints
- Newton-Krylov solvers
- MPI based communication
- Trust-region globalization

Optimization of Approximate Models

$$\min_{z \in \mathcal{Z}} J(\tilde{S}(z), z)$$

- J is the objective
- z is a design, control, or inversion parameter
- $\tilde{S}(z)$ is an approximate model

Our goals are:

- Use the limited high-fidelity evaluations to improve the solution
- Characterize uncertainty in the optimal solution due to $S - \tilde{S}$

Learning Optimal Solution Updates

Approximate Optimal Solution

$$\tilde{z} = \underset{z}{\operatorname{argmin}} \quad J(\tilde{S}(z), z)$$

High-fidelity Data

$$S(z_l) - \tilde{S}(z_l)$$

Learning Optimal Solution Updates

Discrepancy Parameterization

$$\delta(z,\theta) \approx S(z) - \tilde{S}(z)$$

Approximate Optimal Solution

$$\tilde{z}(\theta) = \underset{z}{\text{argmin}}\, J(\tilde{S}(z) + \delta(z,\theta), z)$$

Post-Optimality Sensitivity

$$\frac{\partial \tilde{z}}{\partial \theta}\Delta\theta = -H^{-1}B\Delta\theta$$

Optimal Solution Posterior

$$\Theta \sim \pi_{post} \Rightarrow z^* \approx \frac{\partial \tilde{z}}{\partial \theta}\Theta$$

High-fidelity Data

$$S(z_l) - \tilde{S}(z_l)$$

Bayesian Inversion

$$\pi_{post} \propto \pi_{likelihood}\, \pi_{prior}$$

Prior Discrepancy

π_{prior} = Laws, length scales, etc.

Illustrative Example

$$\min_z \frac{1}{2}\int_0^1 (\tilde{S}(z) - T(x))^2 dx + \frac{\beta}{2}\int_0^1 z\mathcal{E}z$$

where $\tilde{S}(z)$ is the solution operator for

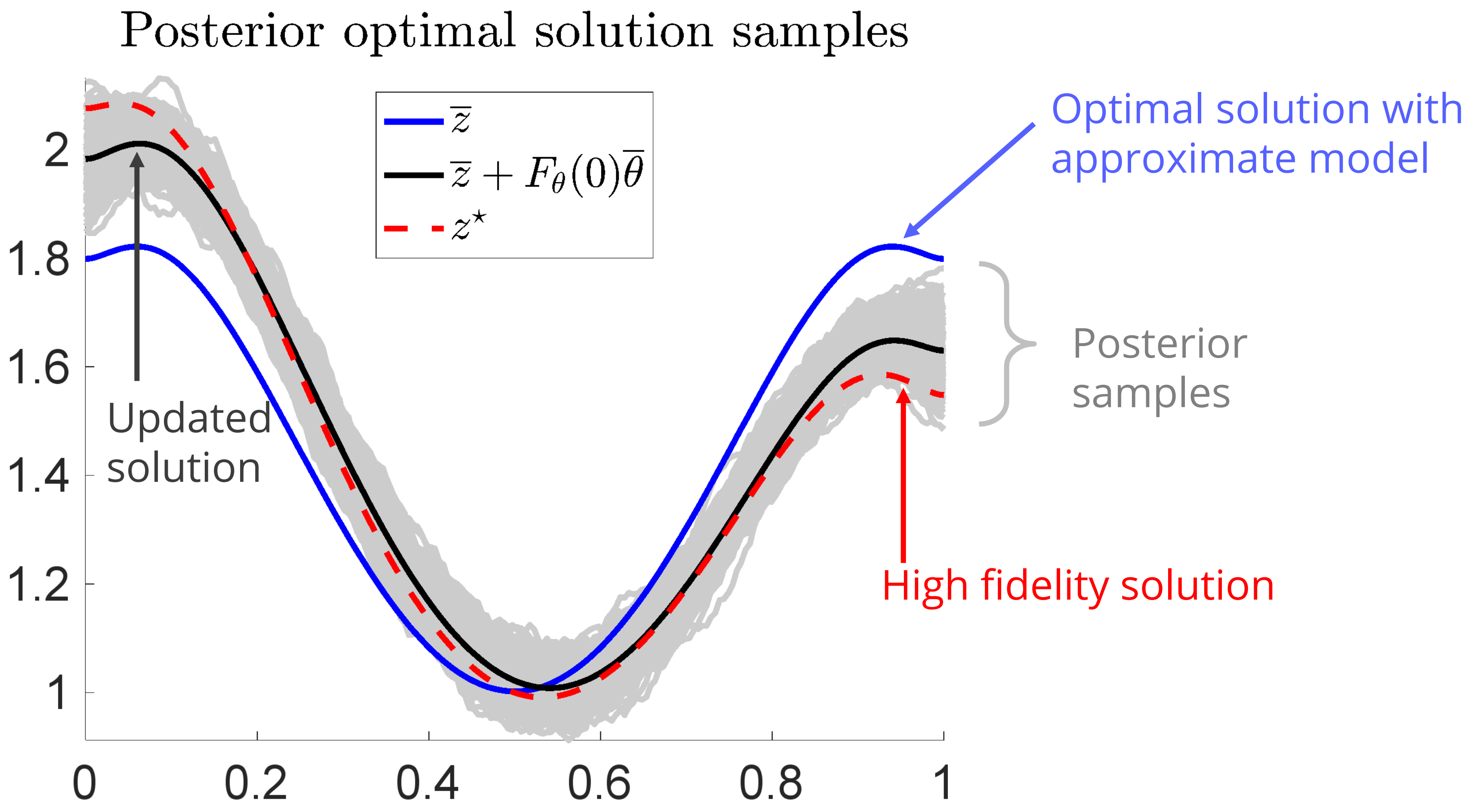
$$\begin{aligned} -\kappa u'' &= z && \text{on } (0,1) \\ \kappa u' &= hu && \text{on } \{0,1\} \end{aligned}$$

The high-fidelity model S solves

$$\begin{aligned} -\kappa u'' + vu' &= z && \text{on } (0,1) \\ \kappa u' &= hu && \text{on } \{0,1\} \end{aligned}$$

Given the high-fidelity solution $S(z)$ for 2 different source terms, improve and characterize uncertainty in the low-fidelity optimal source.

Optimal Solution Posterior



Learning Optimal Solution Updates

Discrepancy Parameterization

$$\delta(z,\theta) \approx S(z) - \tilde{S}(z)$$

Approximate Optimal Solution

$$\tilde{z}(\theta) = \underset{z}{\text{argmin}}\ J(\tilde{S}(z) + \delta(z,\theta), z)$$

Post-Optimality Sensitivity

$$\frac{\partial \tilde{z}}{\partial \theta}\Delta\theta = -H^{-1}B\Delta\theta$$

Optimal Solution Posterior

$$\Theta \sim \pi_{post} \Longrightarrow z^{*} \approx \frac{\partial \tilde{z}}{\partial \theta}\Theta$$

High-fidelity Data

$$S(z_l) - \tilde{S}(z_l)$$

Bayesian Inversion

$$\pi_{post} \propto \pi_{likelihood}\ \pi_{prior}$$

Prior Discrepancy

π_{prior} = Laws, length scales, etc.

Model Discrepancy Representation

- Since post-optimality analysis only depends on the mixed (z, θ) derivative, assume a bi-linear form

$$\delta(\mathbf{z}, \theta) = \left(\mathbf{I}_m \quad \mathbf{I}_m \otimes \mathbf{z}^T \mathbf{M}_z\right) \theta$$

- Discretized $\delta : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^m$ is parameterized by $\theta \in \mathbb{R}^p$
- $p = m(n+1)$ so the dimension of θ may be $\mathcal{O}(\text{mesh size}^2)$
- Evaluate $\delta(z, \theta)$ efficiently using Kronecker product
- $(M_z)_{i,j} = (\psi_i, \psi_j)_Z$ - mass matrix that defines the inner product on $\mathcal{Z}_h$

Learning Optimal Solution Updates

Discrepancy Parameterization

$$\delta(z,\theta) \approx S(z) - \tilde{S}(z)$$

Approximate Optimal Solution

$$\tilde{z}(\theta) = \underset{z}{\operatorname{argmin}}\ J(\tilde{S}(z) + \delta(z,\theta), z)$$

Post-Optimality Sensitivity

$$\frac{\partial \tilde{z}}{\partial \theta}\Delta\theta = -H^{-1}B\Delta\theta$$

Optimal Solution Posterior

$$\Theta \sim \pi_{post} \Longrightarrow z^* \approx \frac{\partial \tilde{z}}{\partial \theta}\Theta$$

High-fidelity Data

$$S(z_l) - \tilde{S}(z_l)$$

Bayesian Inversion

$$\pi_{post} \propto \pi_{likelihood}\ \pi_{prior}$$

Prior Discrepancy

π_{prior} = Laws, length scales, etc.

Post-optimality Sensitivities

$$\min_{\boldsymbol{z}} \mathbf{J}(\tilde{\mathbf{S}}(\boldsymbol{z}) + \boldsymbol{\delta}(\boldsymbol{z}, \theta), \boldsymbol{z}) \quad (1)$$

- $\tilde{\mathbf{z}}^\star$ solves (1) when $\delta(\mathbf{z}, \theta_0) = \mathbf{0}$, the problem solved in practice
- Under mild assumptions, applying the Implicit Function Theorem to

$$\nabla J(\tilde{\mathbf{z}}^\star, \theta_0) = \mathbf{0}$$

gives

$$\mathcal{F} : \mathcal{N}(\theta_0) \to \mathcal{N}(\tilde{\mathbf{z}}^\star)$$

such that $\mathcal{F}(\theta_0)$ solves (1) when $\theta = \theta_0$ and

$$\mathcal{F}'_\theta(\theta_0) = -\mathbf{H}^{-1}\mathbf{B}$$

is the sensitivity of the optimal solution with respect to model discrepancy

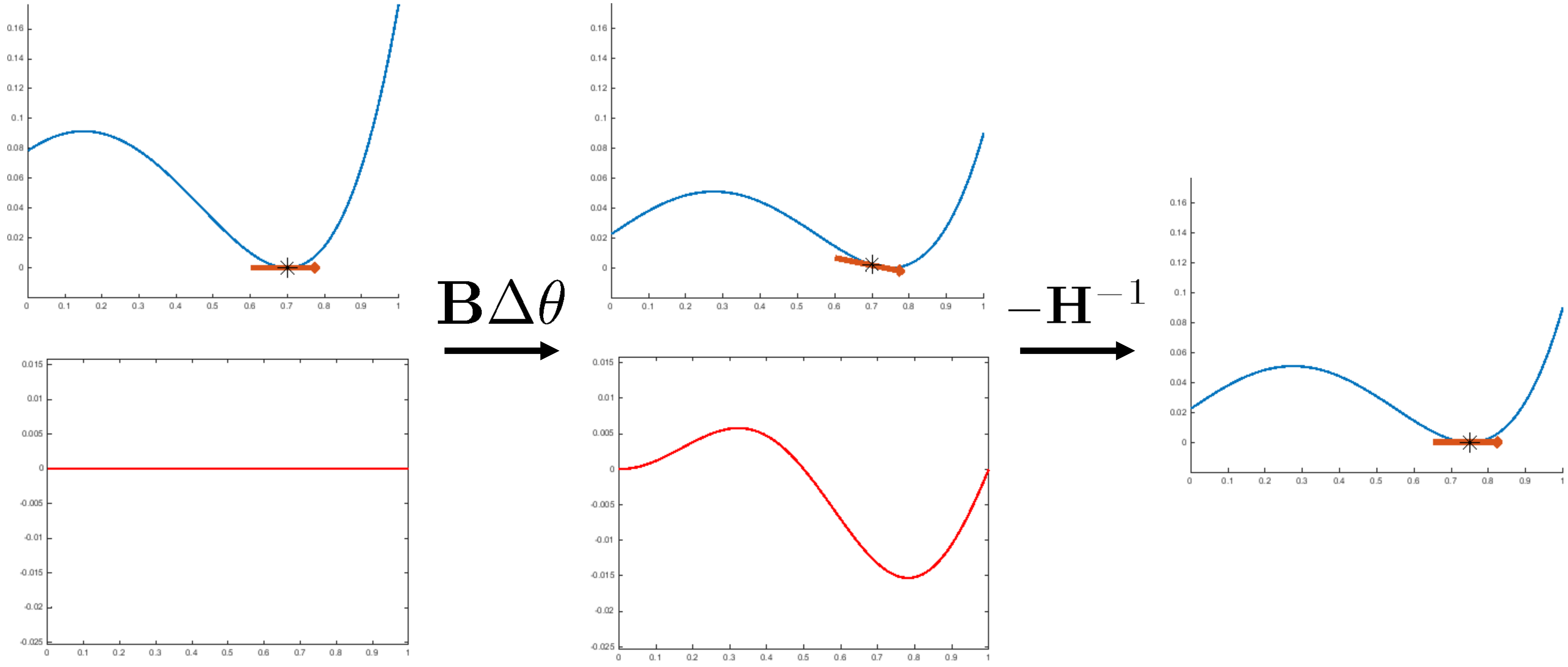
Post-optimality Sensitivities

$$\mathcal{F}'_\theta(\theta_0) = -\mathbf{H}^{-1}\mathbf{B}$$

- $\mathbf{H}$ is the Hessian of the objective function with respect to z
- $\mathbf{B}$ is the mixed second derivative of the objective with respect to z and θ

$\mathbf{J}(\tilde{\mathbf{S}}(\boldsymbol{z}) + \boldsymbol{\delta}(\boldsymbol{z}, \theta), \boldsymbol{z})$

$\delta(z, \theta)$



Discrepancy Parameterization

$$\delta(z,\theta) \approx S(z) - \tilde{S}(z)$$

Approximate Optimal Solution

$$\tilde{z}(\theta) = \underset{z}{\operatorname{argmin}}\, J(\tilde{S}(z) + \delta(z,\theta), z)$$

Post-Optimality Sensitivity

$$\frac{\partial \tilde{z}}{\partial \theta}\Delta\theta = -H^{-1}B\Delta\theta$$

Optimal Solution Posterior

$$\Theta \sim \pi_{post} \Longrightarrow z^* \approx \frac{\partial \tilde{z}}{\partial \theta}\Theta$$

High-fidelity Data

$$S(z_l) - \tilde{S}(z_l)$$

Bayesian Inversion

$$\pi_{post} \propto \pi_{likelihood}\, \pi_{prior}$$

Prior Discrepancy

π_{prior} = Laws, length scales, etc.

Bayesian Inverse Problem - Prior Discrepancy

- Measure size of $\boldsymbol{\delta}$:

$$\mathbb{E}_{\boldsymbol{z}}[||\boldsymbol{\delta}(\boldsymbol{z},\boldsymbol{\theta})||_{\boldsymbol{L}}^2] = \boldsymbol{\theta}^T \boldsymbol{M}_{\boldsymbol{\theta}} \boldsymbol{\theta}$$

where

$$\boldsymbol{M}_{\boldsymbol{\theta}} = \begin{pmatrix} \boldsymbol{L} & \boldsymbol{L} \otimes \overline{\boldsymbol{z}}^T \boldsymbol{M}_z \\ \boldsymbol{L} \otimes \boldsymbol{M}_z \overline{\boldsymbol{z}} & \boldsymbol{L} \otimes \boldsymbol{E} \end{pmatrix}$$

- $\boldsymbol{L}$ encodes known physics of the discrepancy - in our case a Laplacian like operator and $\boldsymbol{L}^{-1}$ represents the prior covariance
- $\boldsymbol{\Gamma}$ is a covariance matrix on the control space $\mathcal{Z}$
- Hence $\boldsymbol{M}_{\boldsymbol{\theta}}$ defines an inner product for $\boldsymbol{\theta}$ to measure the size of the model discrepancy $\boldsymbol{\delta}(\boldsymbol{z},\boldsymbol{\theta})$ according to our prior knowledge imposed in $\boldsymbol{L}$ and $\boldsymbol{\Gamma}$

Bayesian Inversion Problem

- Arrange data and discrepancy representation so that we seek θ such that

$$\mathbf{A}\theta \approx \mathbf{b}$$

- Given Gaussian prior and noise models, linearity of $\delta(\mathbf{z}, \theta)$ in θ, the posterior is Gaussian with a negative log probability density function

$$\frac{1}{2\alpha}\left(\mathbf{A}\theta - \mathbf{b}\right)^T\left(\mathbf{A}\theta - \mathbf{b}\right) + \frac{1}{2}\theta^T\mathbf{M}_\theta\theta.$$

- α balances the dependence of prior and data misfit

- The posterior mean is

$$\bar{\theta} = \frac{1}{\alpha}\mathbf{\Sigma}\mathbf{A}^T\mathbf{b}$$

  and the posterior covariance is

$$\mathbf{\Sigma} = \left(\mathbf{M}_\theta + \frac{1}{\alpha}\mathbf{A}^T\mathbf{A}\right)^{-1}.$$

Bayesian Inversion Problem – Enabling Sampling

- The goal is to sample from a Gaussian distribution which may be generated by multiplying a factor of the covariance matrix with a standard normal random vector and adding the mean
- But how do we invert the sum?

$$\mathbf{\Sigma} = \left(\mathbf{M}_\theta + \frac{1}{\alpha}\mathbf{A}^T\mathbf{A}\right)^{-1}$$

1. Factorize $\mathbf{A}$ to rewrite $\mathbf{M}_\theta + \frac{1}{\alpha}\mathbf{A}^T\mathbf{A}$
2. Invert $\mathbf{M}_\theta + \frac{1}{\alpha}\mathbf{A}^T\mathbf{A}$
3. Factorize $\mathbf{\Sigma}$
4. Compute matrix-vector products for posterior samples

Posterior Samples for Discrepancy

- Posterior samples take the form

$$\overline{\theta} + \hat{\theta} + \tilde{\theta}$$

where the mean is

$$\overline{\theta} = \frac{1}{\alpha} \sum_{\ell=1}^{N} \left[\binom{a_\ell u_\ell}{u_\ell \otimes \mathbf{M}_z^{-1} \mathbf{\Gamma}^{-1} (z_\ell - \overline{z})} - \sum_{i=1}^{N} b_{i,\ell} \binom{s_i \boldsymbol{u}_{i,\ell}}{\boldsymbol{u}_{i,\ell} \otimes \mathbf{M}_z^{-1} \mathbf{\Gamma}^{-1} \boldsymbol{w}_i} \right]$$

uncertainty in the data informed directions is

$$\hat{\theta} = \sqrt{\alpha} \sum_{i=1}^{N} \frac{1}{\sqrt{\lambda_i}} \binom{s_i \hat{u}_i}{\hat{u}_i \otimes \mathbf{M}_z^{-1} \mathbf{\Gamma}^{-1} \boldsymbol{w}_i}$$

and uncertainty in the data uninformed directions is

$$\tilde{\theta} = \sum_{k=1}^{n-N+1} \binom{\tilde{s}_k \tilde{\boldsymbol{u}}_k}{\tilde{u}_k \otimes \tilde{\boldsymbol{w}}_k}$$

Learning Optimal Solution Updates

Discrepancy Parameterization

$$\delta(z,\theta) \approx S(z) - \tilde{S}(z)$$

Approximate Optimal Solution

$$\tilde{z}(\theta) = \underset{z}{\text{argmin}}\, J(\tilde{S}(z) + \delta(z,\theta), z)$$

Post-Optimality Sensitivity

$$\frac{\partial \tilde{z}}{\partial \theta}\Delta\theta = -H^{-1}B\Delta\theta$$

Optimal Solution Posterior

$$\Theta \sim \pi_{post} \Longrightarrow z^{*} \approx \frac{\partial \tilde{z}}{\partial \theta}\Theta$$

High-fidelity Data

$$S(z_l) - \tilde{S}(z_l)$$

Bayesian Inversion

$$\pi_{post} \propto \pi_{likelihood}\, \pi_{prior}$$

Prior Discrepancy

π_{prior} = Laws, length scales, etc.

Propagating Samples Through Post-optimality Sensitivities

$$\mathcal{F}'_\theta(\theta_0)(\overline{\theta} + \hat{\theta} + \tilde{\theta}) = -\mathbf{H}^{-1}(\mathbf{B}\overline{\theta} + \mathbf{B}\hat{\theta} + \mathbf{B}\tilde{\theta})$$

$$\mathbf{B}\overline{\theta} = \frac{1}{\alpha}\tilde{\mathbf{S}}_z^T \nabla_{u,u}\mathbf{J}\left[\sum_{\ell=1}^{N}\left(u_\ell - \sum_{i=1}^{N} b_{i,\ell}(e^T \boldsymbol{g}_i)\boldsymbol{u}_{i,\ell}\right)\right] + \frac{1}{\alpha}\sum_{\ell=1}^{N}(\nabla_u J u_\ell)\boldsymbol{\Gamma}^{-1}(z_\ell - \overline{z}) - \frac{1}{\alpha}\sum_{\ell=1}^{N}\sum_{i=1}^{N} b_{i,\ell}(\nabla_u \mathbf{J}\boldsymbol{u}_{i,\ell})\boldsymbol{\Gamma}^{-1}\boldsymbol{w}_i$$

$$\mathbf{B}\hat{\theta} = \sqrt{\alpha}\tilde{\mathbf{S}}_z^T \nabla_{u,u}\mathbf{J}\left(\sum_{i=1}^{N}\frac{e^T \boldsymbol{g}_i}{\sqrt{\lambda_i}}\hat{u}_i\right) + \sqrt{\alpha}\sum_{i=1}^{N}\frac{\nabla_u \mathbf{J}\hat{u}_i}{\sqrt{\lambda_i}}\boldsymbol{\Gamma}^{-1}\boldsymbol{w}_i$$

$$\mathbf{B}\tilde{\theta} = \sum_{k=1}^{n-N+1}(\nabla_u \mathbf{J}\tilde{u}_k)\,\boldsymbol{\Gamma}^{-\frac{1}{2}}\,\tilde{z}_k.$$

High-fidelity data

Approximate model

Prior

Optimization

Propagating Samples Through Post-optimality Sensitivities

$$\mathcal{F}'_\theta(\theta_0)(\overline{\theta} + \hat{\theta} + \tilde{\theta}) = -\mathbf{H}^{-1}(\mathbf{B}\overline{\theta} + \mathbf{B}\hat{\theta} + \mathbf{B}\tilde{\theta})$$

$$\mathbf{B}\overline{\theta} = \frac{1}{\alpha}\tilde{\mathbf{S}}_z^T \nabla_{u,u}\mathbf{J}\left[\sum_{\ell=1}^{N}\left(u_\ell - \sum_{i=1}^{N} b_{i,\ell}(e^T\boldsymbol{g}_i)\boldsymbol{u}_{i,\ell}\right)\right] + \frac{1}{\alpha}\sum_{\ell=1}^{N}(\nabla_u\mathbf{J}u_\ell)\boldsymbol{\Gamma}^{-1}(z_\ell - \overline{z}) - \frac{1}{\alpha}\sum_{\ell=1}^{N}\sum_{i=1}^{N} b_{i,\ell}(\nabla_u\mathbf{J}\boldsymbol{u}_{i,\ell})\boldsymbol{\Gamma}^{-1}\boldsymbol{w}_i$$

$$\mathbf{B}\hat{\theta} = \sqrt{\alpha}\tilde{\mathbf{S}}_z^T \nabla_{u,u}\mathbf{J}\left(\sum_{i=1}^{N}\frac{e^T\boldsymbol{g}_i}{\sqrt{\lambda_i}}\hat{u}_i\right) + \sqrt{\alpha}\sum_{i=1}^{N}\frac{\nabla_u\mathbf{J}\hat{u}_i}{\sqrt{\lambda_i}}\boldsymbol{\Gamma}^{-1}\boldsymbol{w}_i$$

$$\mathbf{B}\tilde{\theta} = \sum_{k=1}^{n-N+1}(\nabla_u\mathbf{J}\tilde{u}_k)\,\boldsymbol{\Gamma}^{-\frac{1}{2}}\tilde{z}_k.$$

High-fidelity data

Approximate model

Prior

Optimization

Propagating Samples Through Post-optimality Sensitivities

$$\mathcal{F}'_\theta(\theta_0)(\overline{\theta} + \hat{\theta} + \tilde{\theta}) = -\mathbf{H}^{-1}(\mathbf{B}\overline{\theta} + \mathbf{B}\hat{\theta} + \mathbf{B}\tilde{\theta})$$

$$\mathbf{B}\overline{\theta} = \frac{1}{\alpha}\tilde{\mathbf{S}}_z^T \nabla_{u,u}\mathbf{J}\left[\sum_{\ell=1}^{N}\left(u_\ell - \sum_{i=1}^{N} b_{i,\ell}(e^T\boldsymbol{g}_i)\boldsymbol{u}_{i,\ell}\right)\right] + \frac{1}{\alpha}\sum_{\ell=1}^{N}(\nabla_u J u_\ell)\boldsymbol{\Gamma}^{-1}(z_\ell - \overline{z}) - \frac{1}{\alpha}\sum_{\ell=1}^{N}\sum_{i=1}^{N} b_{i,\ell}(\nabla_u \mathbf{J}\boldsymbol{u}_{i,\ell})\boldsymbol{\Gamma}^{-1}\boldsymbol{w}_i$$

$$\mathbf{B}\hat{\theta} = \sqrt{\alpha}\tilde{\mathbf{S}}_z^T \nabla_{u,u}\mathbf{J}\left(\sum_{i=1}^{N}\frac{e^T\boldsymbol{g}_i}{\sqrt{\lambda_i}}\hat{u}_i\right) + \sqrt{\alpha}\sum_{i=1}^{N}\frac{\nabla_u\mathbf{J}\hat{u}_i}{\sqrt{\lambda_i}}\boldsymbol{\Gamma}^{-1}\boldsymbol{w}_i$$

$$\mathbf{B}\tilde{\theta} = \sum_{k=1}^{n-N+1}(\nabla_u\mathbf{J}\tilde{u}_k)\boldsymbol{\Gamma}^{-\frac{1}{2}}\tilde{z}_k.$$

High-fidelity data

Approximate model

Prior

Optimization

Propagating Samples Through Post-optimality Sensitivities

$$\mathcal{F}_\theta'(\theta_0)(\overline{\theta} + \hat{\theta} + \tilde{\theta}) = -\mathbf{H}^{-1}(\mathbf{B}\overline{\theta} + \mathbf{B}\hat{\theta} + \mathbf{B}\tilde{\theta})$$

$$\mathbf{B}\overline{\theta} = \frac{1}{\alpha}\tilde{\mathbf{S}}_z^T \nabla_{u,u}\mathbf{J}\left[\sum_{\ell=1}^{N}\left(u_\ell - \sum_{i=1}^{N} b_{i,\ell}(e^T \boldsymbol{g}_i)\boldsymbol{u}_{i,\ell}\right)\right] + \frac{1}{\alpha}\sum_{\ell=1}^{N}(\nabla_u \mathbf{J} u_\ell \boldsymbol{\Gamma}^{-1}(z_\ell - \overline{z}) - \frac{1}{\alpha}\sum_{\ell=1}^{N}\sum_{i=1}^{N} b_{i,\ell}(\nabla_u \mathbf{J}\boldsymbol{u}_{i,\ell}\boldsymbol{\Gamma}^{-1}\boldsymbol{w}_i$$

$$\mathbf{B}\hat{\theta} = \sqrt{\alpha}\tilde{\mathbf{S}}_z^T \nabla_{u,u}\mathbf{J}\left(\sum_{i=1}^{N} \frac{e^T \boldsymbol{g}_i}{\sqrt{\lambda_i}} \hat{u}_i\right) + \sqrt{\alpha}\sum_{i=1}^{N} \frac{\nabla_u \mathbf{J}\hat{u}_i}{\sqrt{\lambda_i}} \boldsymbol{\Gamma}^{-1}\boldsymbol{w}_i$$

$$\mathbf{B}\tilde{\theta} = \sum_{k=1}^{n-N+1} (\nabla_u \mathbf{J}\tilde{u}_k \boldsymbol{\Gamma}^{-\frac{1}{2}} \xi_k.$$

High-fidelity data

Approximate model

Prior

Optimization

Propagating Samples Through Post-optimality Sensitivities

$$\mathcal{F}_\theta'(\theta_0)(\overline{\theta} + \hat{\theta} + \tilde{\theta}) = -\mathbf{H}^{-1}(\mathbf{B}\overline{\theta} + \mathbf{B}\hat{\theta} + \mathbf{B}\tilde{\theta})$$

$$\mathbf{B}\bar{\theta} = \frac{1}{\alpha}\tilde{\mathbf{S}}_z^T \nabla_{u,u}\mathbf{J}\left[\sum_{\ell=1}^{N}\left(u_\ell - \sum_{i=1}^{N} b_{i,\ell}(e^T \boldsymbol{g}_i)\boldsymbol{u}_{i,\ell}\right)\right] + \frac{1}{\alpha}\sum_{\ell=1}^{N}(\nabla_u J u_\ell)\boldsymbol{\Gamma}^{-1}(z_\ell - \bar{z}) - \frac{1}{\alpha}\sum_{\ell=1}^{N}\sum_{i=1}^{N} b_{i,\ell}(\nabla_u \mathbf{J} u_{i,\ell})\boldsymbol{\Gamma}^{-1}\boldsymbol{w}_i$$

$$\mathbf{B}\hat{\theta} = \sqrt{\alpha}\tilde{\mathbf{S}}_z^T \nabla_{u,u}\mathbf{J}\left(\sum_{i=1}^{N}\frac{e^T \boldsymbol{g}_i}{\sqrt{\lambda_i}}\hat{u}_i\right) + \sqrt{\alpha}\sum_{i=1}^{N}\frac{\nabla_u \mathbf{J}\hat{u}_i}{\sqrt{\lambda_i}}\boldsymbol{\Gamma}^{-1}\boldsymbol{w}_i$$

$$\mathbf{B}\tilde{\theta} = \sum_{k=1}^{n-N+1}(\nabla_u \mathbf{J}\tilde{u}_k)\boldsymbol{\Gamma}^{-\frac{1}{2}}\tilde{z}_k.$$

High-fidelity data

Approximate model

Prior

Optimization

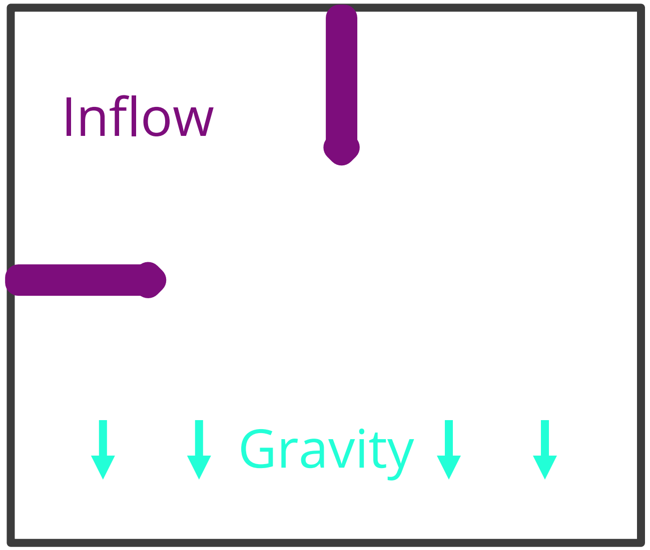
A Fluid Flow Example

Optimal design of a flow controller

$$\min_z \frac{1}{2}\int_\chi \mathbf{v}_y(z)^2 + \frac{\beta}{2}\int_\Omega ||\mathbf{z}||^2$$

constrained by the Stokes equations

$$\begin{aligned} -\mu\nabla\mathbf{v} + \nabla p &= \mathbf{g} + \mathbf{z} && \text{on } \Omega \\ \nabla\cdot\mathbf{v} &= 0 && \text{on } \Omega \end{aligned}$$

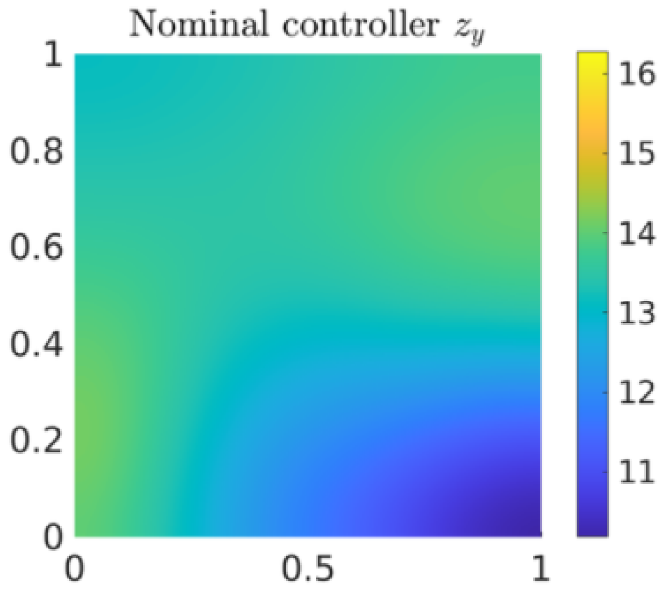


as a simplification of the Navier-Stokes equations

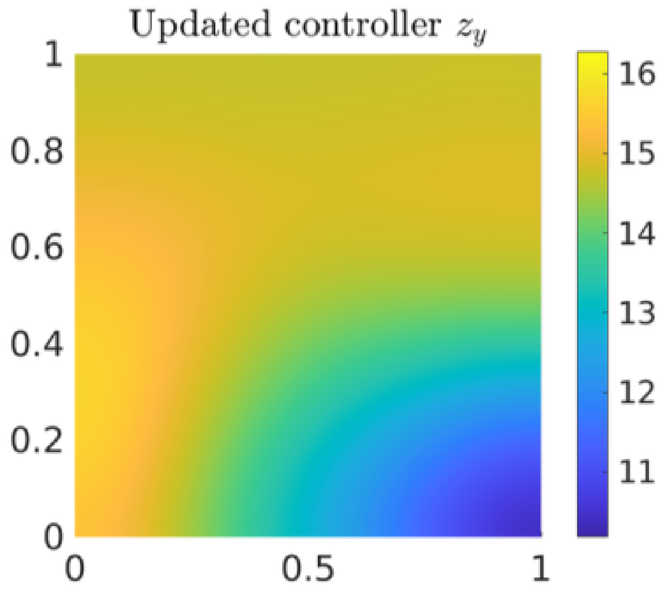
$$\begin{aligned} -\mu\nabla\mathbf{v} + (\mathbf{v}\cdot\nabla)\mathbf{v} + \nabla p &= \mathbf{g} + \mathbf{z} && \text{on } \Omega \\ \nabla\cdot\mathbf{v} &= 0 && \text{on } \Omega \end{aligned}$$

Comparison of Controllers

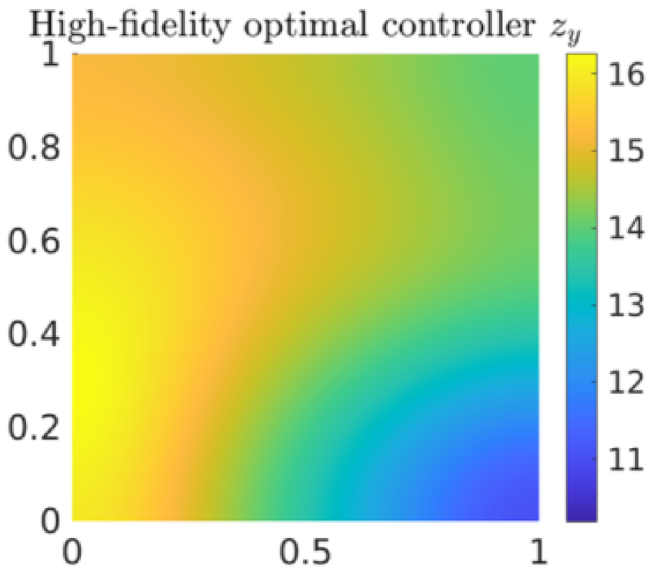
Using only Stokes



Using Stokes + 1 NS forward solve

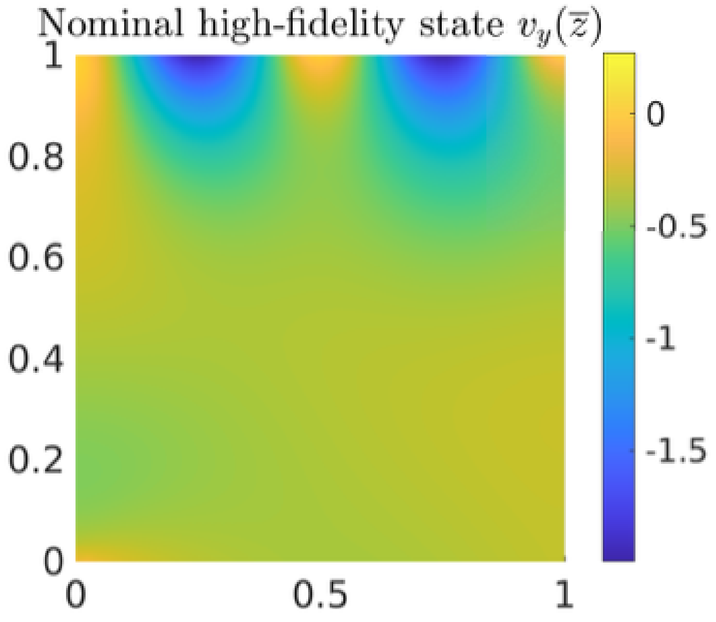


Using NS – "Ground Truth"



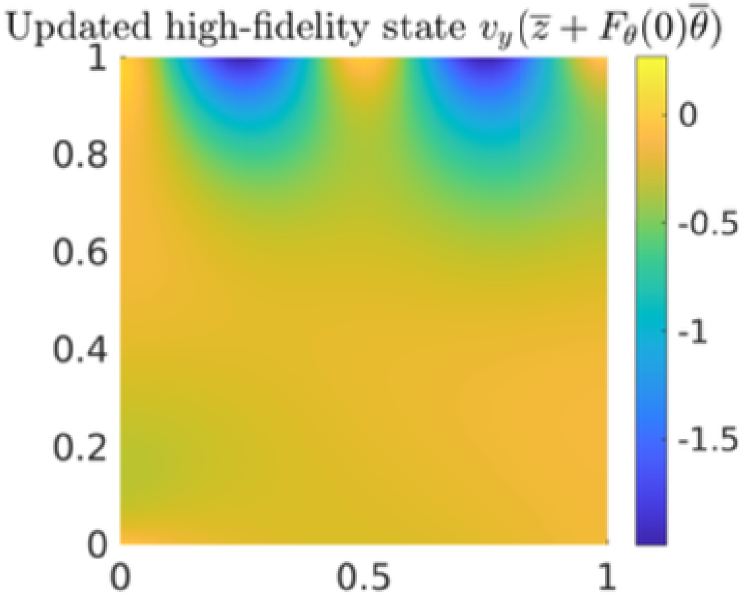
Comparison of States

Navier-Stokes solve
with nominal control



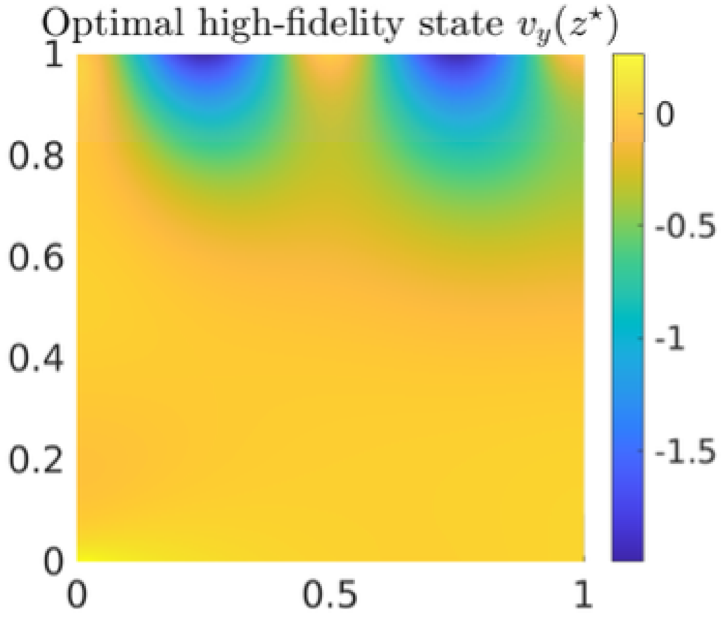
$J = 3 \times 10^0$

Navier-Stokes solve
with updated control



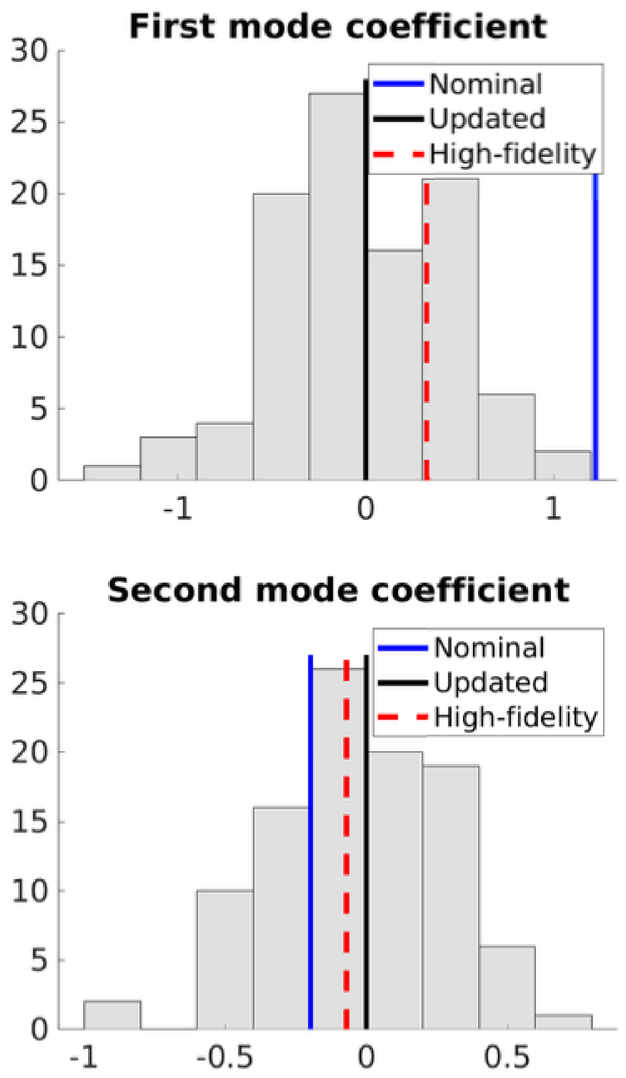
$J = 1 \times 10^{-2}$

Navier-Stokes solve
with optimal control



$J = 2 \times 10^{-3}$

Posterior Controller Uncertainty



- KL representation
- Histogram of posterior
- Goal is for updated to be as close as possible to high-fidelity

Conclusions

- Developed a framework to learn updates of low-fidelity optimal solutions using limited high-fidelity data
- Approach is non-intrusive to the high-fidelity data and hence applicable to wide range of applications
- Explore applications where high-fidelity data comes from experiments
- Future work to seek optimal data collection strategies
- Generalize to incorporate other modes and/or fidelities of the data

✓ Joseph Hart and Bart van Bloemen Waanders, “Hyper-Differential Sensitivity Analysis With Respect to Model Discrepancy: Mathematics and Computation” (in preparation)

✓ Joseph Hart and Bart van Bloemen Waanders, ”Hyper-differential sensitivity analysis with respect to model discrepancy: Calibration and Optimal Solution Updating” (in preparation)

- **R**obustness: UQ for optimal controller
- **I**nterpretability: prior, optimization, data, and physics in the controller update
- **S**calability: leveraging computational efficiency from PDECO methods
- **E**fficiency: Kronecker product and closed form solutions to controller updates