Developing Scalable and Multi-fidelity Approaches for Push-forward Based Inference

Tim Wildey[1], Troy Butler[2], John Jakeman[1]
Tan Bui-Thanh[3], Brad Marvin[3], Lukas Bruder[4]
[1]*Sandia National Laboratories, Albuquerque, NM*
[2]*University of Colorado - Denver, Denver, CO*
[3]*University of Texas-Austin, ICES. Austin, TX*
[4]*TU-Munich,, Germany*

Abstract

Computationally efficient, accurate and robust methods for producing **data-informed physics-based predictions** from multiscale/multiphysics models are critical to the mission of the DOE/ASCR. We build upon utilize a recently developed approach for solving stochastic inverse problems [1,2] based on a combination of measure theory and Bayes' rule. We utilize **scalable** approaches based on connections with deterministic optimization [3] and **multi-fidelity** approaches [4] to reduce the computational cost in solving the inverse problem.

Motivation

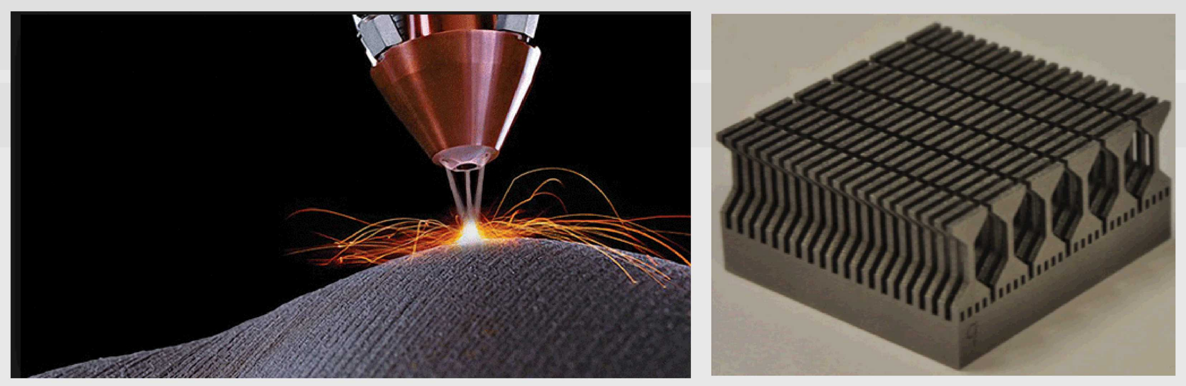
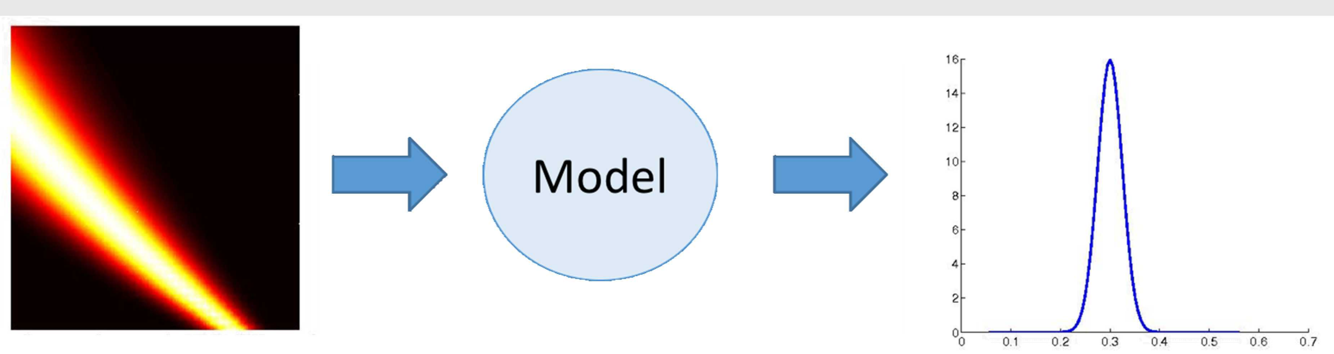


Figure: High-throughput manufacturing provides new data challenges.



An Inverse Problem

Given a probability density on observations, find a probability density on model inputs such that the push-forward matches the given density on the observed data.

Approach and Preliminary Results

Theorem (Butler, Jakeman, Wildey, SISC, 2018)

Given an initial probability measure, P_Λ^{init} on $(\Lambda, \mathcal{B}_\Lambda)$ and an observed probability measure, $P_\mathcal{D}^{obs}$, on $(\mathcal{D}, \mathcal{B}_\mathcal{D})$, the probability measure P_Λ^{up} on $(\Lambda, \mathcal{B}_\Lambda)$ defined by

$$P_\Lambda^{up}(A) = \int_\mathcal{D} \left(\int_{A \cap Q^{-1}(q)} \pi_\Lambda^{init}(\lambda) \frac{\pi_\mathcal{D}^{obs}(Q(\lambda))}{\pi_\mathcal{D}^{Q(init)}(Q(\lambda))} d\mu_{\Lambda,q}(\lambda) \right) d\mu_\mathcal{D}(q), \ \forall A \in \mathcal{B}_\Lambda$$

solves the stochastic inverse problem.

- $\square^{init}$ is the given initial probability density
- $\square_{\mathcal{D}}^{obs}$ is the given distribution of the data
- Computing $\square_{\mathcal{D}}^{Q(init)}$ requires a forward propagation of the initial.

Scalable methods based on optimization

- Consider a linear map: $A : \Lambda \to \mathcal{D}$ with $\pi_\Lambda^{init} \sim N(\bar{\lambda}, C_\Lambda)$ and $\pi_\mathcal{D}^{obs} \sim N(\bar{q}, C_\mathcal{D})$
- MAP minimizes the following functional with $C_A = A C_\Lambda A^T$

$$J(\lambda) = \underbrace{\frac{1}{2}\|C_\mathcal{D}^{-1/2}(A\lambda - \bar{q})\|_2^2}_{\text{Data misfit}} + \underbrace{\frac{1}{2}\|C_\Lambda^{-1/2}(\lambda - \bar{\lambda})\|_2^2}_{\text{Tikonov}} + \underbrace{\frac{1}{2}\|C_A^{-1/2}(A\lambda - A\bar{\lambda})\|_2^2}_{\text{Un-regularization}}$$

- The Hessian is given by:

$$H = A^T C_\mathcal{D}^{-1} A + C_\Lambda^{-1} - A^T \left(A C_\Lambda A^T\right)^{-1} A$$

- We utilize a randomized low-rank approximation method for scalability

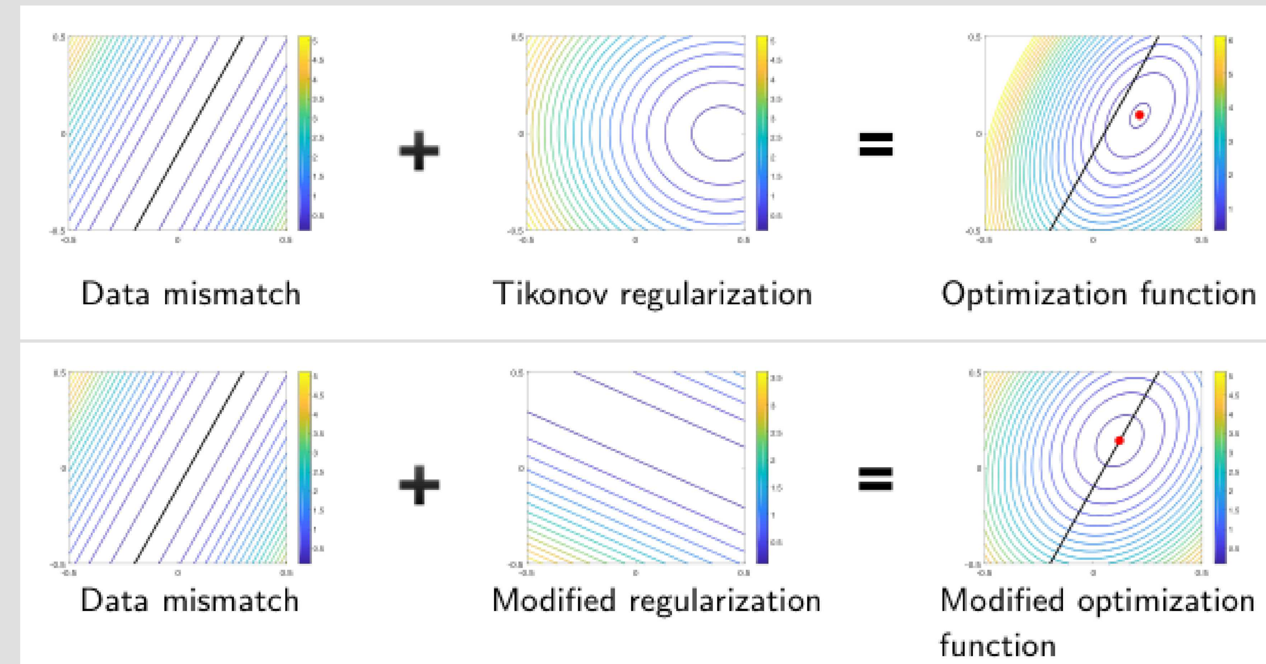


Figure: Illustration of Tikonov vs. modified regularization.

N	Classical Bayes				Push-forward Based			
	Iters	CG	Misfit	Time	Iters	CG	Misfit	Time
100	9.1	48.2	12.2	0.012	9.2	48.5	2.84e-13	0.012
500	9.2	55.1	60.7	0.026	9.2	55.6	3.96e-13	0.027
1000	10.6	61.7	121.9	0.109	10.7	61.9	3.06e-13	0.110
2000	11.0	62.9	242.8	0.549	11.0	63.9	2.76e-13	0.553
4000	11.2	65.3	484.1	2.88	11.3	66.8	3.75e-13	2.90
8000	12.4	71.4	965.6	19.7	12.6	73.0	8.59e-13	20.0

Table: Convergence of standard and modified functionals for a linear map with a trust region Newton-CG algorithm as dimension of parameter space increases.

Nonlinear hyperelastic example:

- Compressible Mooney-Rivlin solid
- Piecewise constant material property
- Dimension of parameter space: 36905
- Dimension of state space: 14487
- Dimension of data space: 2778

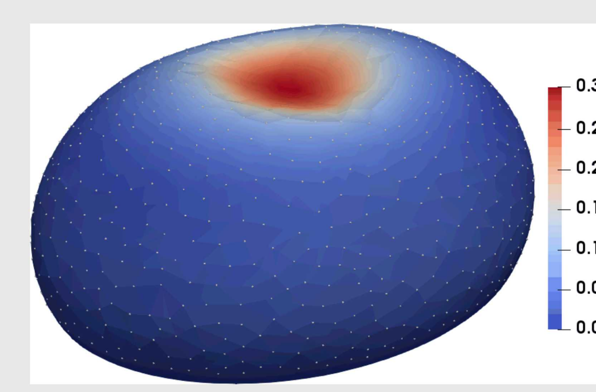


Figure: True displacement

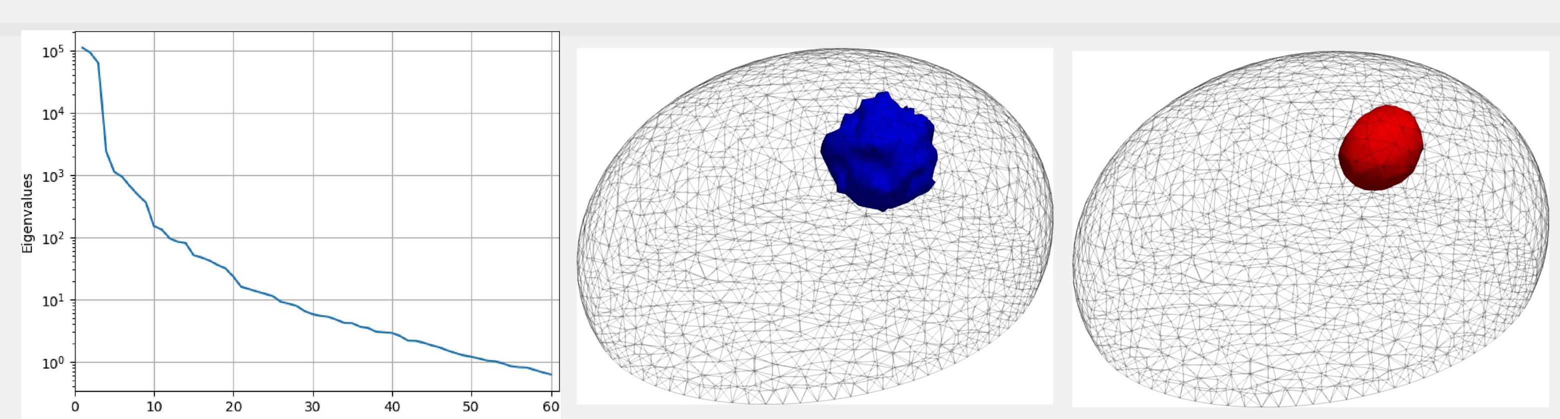


Figure: Decay of eigenvalues (left), true material coefficient (middle), MAP solution (right).

Multi-fidelity methods

- **Goal:** leverage correlations with low-fidelity models to reduce number of high-fidelity model evaluations
- Only desire ability to generate samples from updated density
- Let Q_H and Q_L denote the high- and low-fidelity models respectively

$$\pi_{\mathcal{D}_H}^{Q_H(\text{init})}(q) = \int_{\mathcal{D}_L} \pi(Q_H|Q_L)\, \pi_{\mathcal{D}_L}^{Q_L(\text{init})}(q) dq$$

- $\pi_{\mathcal{D}_L}^{Q_L(\text{init})}(q)$ can be approximated using Monte Carlo sampling and KDE
- The key is constructing $\pi(Q_H|Q_L)$
- We use Gaussian process regression with heteroscedastic noise model

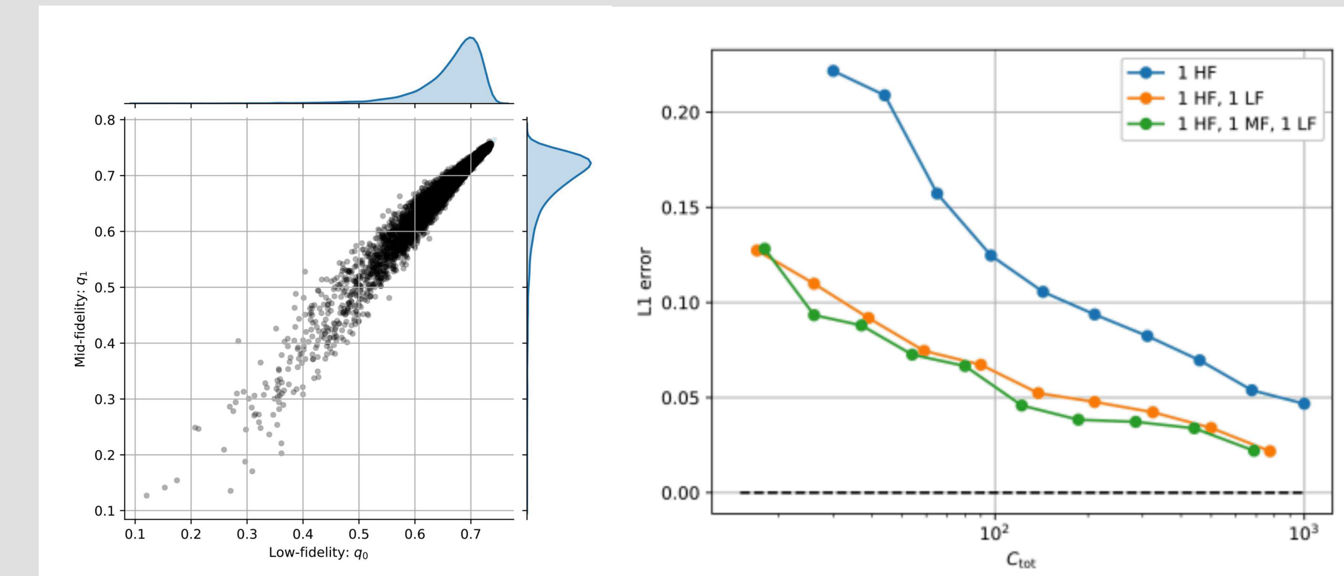


Figure: Samples from joint distribution and marginals (left) and convergence of the updated density (right) for the linear elliptic PDE.

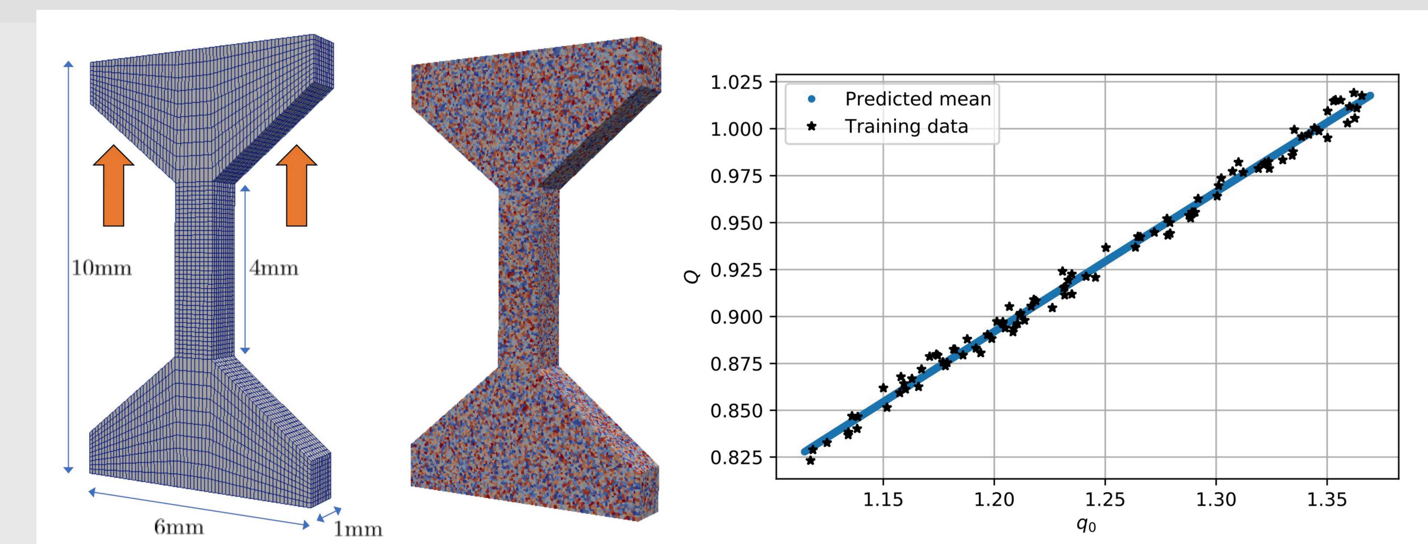


Figure: Low-fidelity LE model with 16,600 elements (left), high-fidelity CE model with 17,040,000 elements and 224,000 grains (middle), training data for regression model (right).

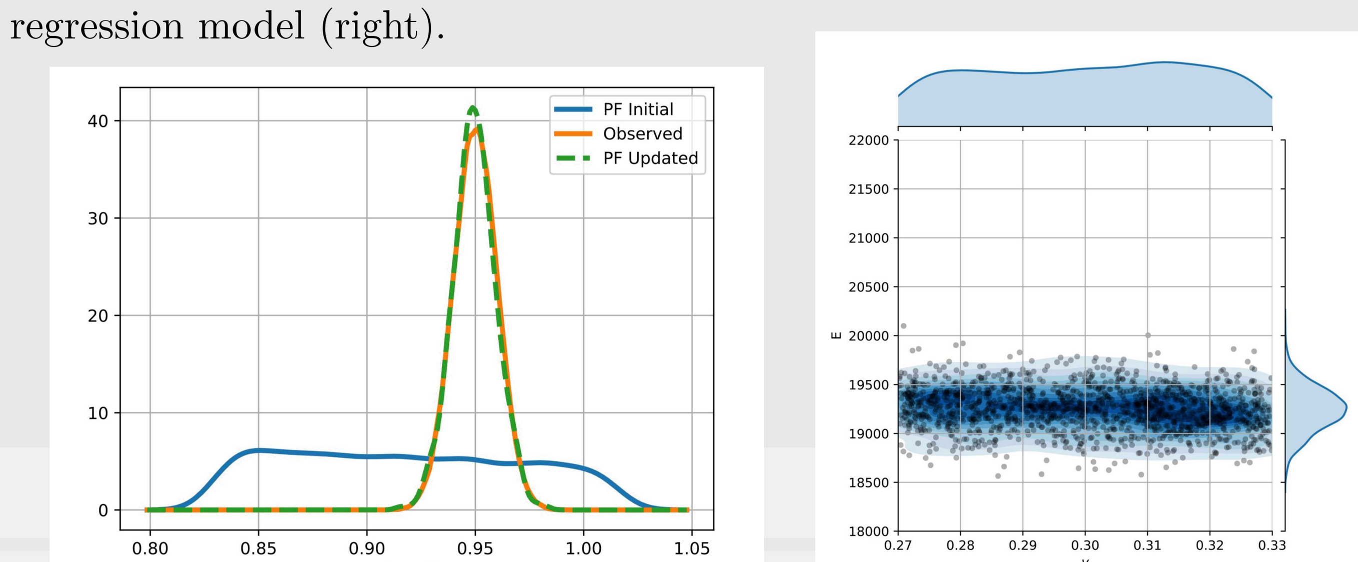


Figure: Push-forward of initial and updated densities with observed density (left) and samples from the updated density (right).

Potential Impact

- Enabling this form of stochastic inference on large-scale multiscale and multiphysics applications

Synergy

- Applications desiring data-informed physics-based predictions
- Robust/efficient/scalable solutions to stochastic inverse problems
- Optimal experimental design (for prediction)
- Extracting information from "big" data

References

[1] T. Butler, J. Jakeman, and T. Wildey, SISC, 40(2) 2018, A984–A1011.
[2] T. Butler, J. Jakeman, and T. Wildey, SISC, 40(5) 2018, A3523–A3548.
[3] B. Marvin, T. Wildey and T. Bui-Thanh, CCR Summer Research Program Proceedings, 2018.
[4] L. Bruder, M. W. Gee and T. Wildey, Submitted to IJUQ, 2019.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and 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-NA-0003525.