# Solving Stochastic Inverse Problems using Approximate Push-forward Densities based on a Multi-fidelity Monte Carlo Method

SAND2019-7564C

Tim Wildey[♣], Troy Butler[♦], John Jakeman[♣], Lukas Bruder[♠]

[♣] Sandia National Laboratories
Center for Computing Research
Optimization and UQ Group

[♦] University of Colorado Denver
Department of Mathematics and Statistics

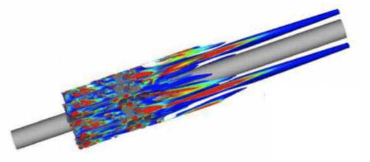
[♠] Technical University of Munich
Mechanics & High Performance Computing Group

9th International Congress on Industrial and Applied Mathematics
July 15-19, 2019

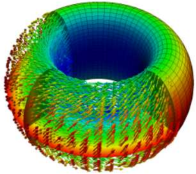
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. SAND2019-**** C

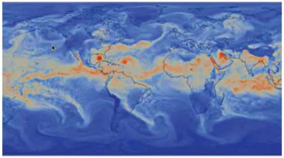
# Motivation

Flow in Nuclear Reactor (Turbulent CFD)

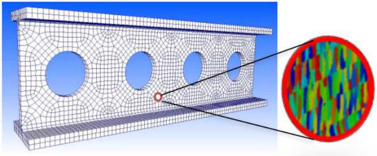


Tokamak Equilibrium (MHD)





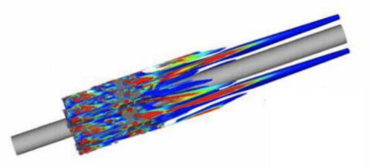
Climate Modeling



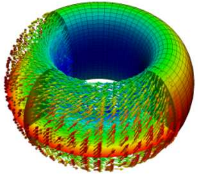
Multi-scale Materials Modeling

# Motivation

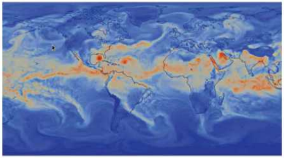
Flow in Nuclear Reactor (Turbulent CFD)



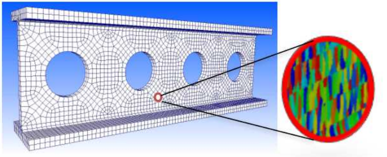
Tokamak Equilibrium (MHD)



We are working to develop **data-informed** models ...

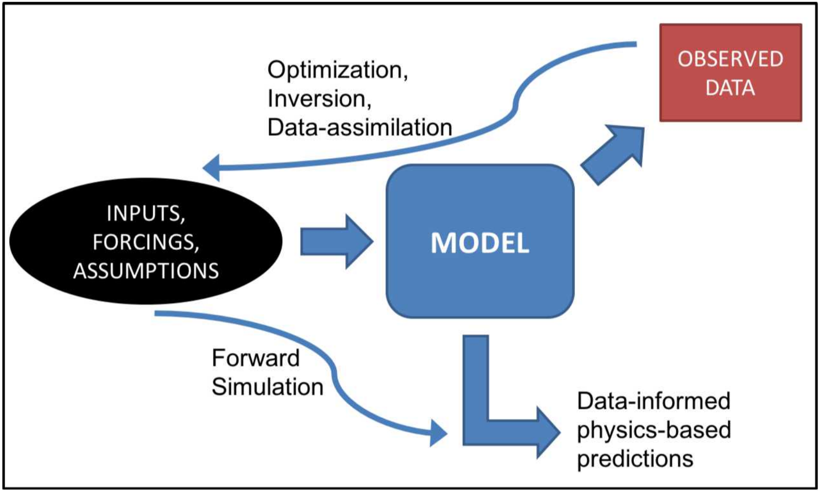


Climate Modeling

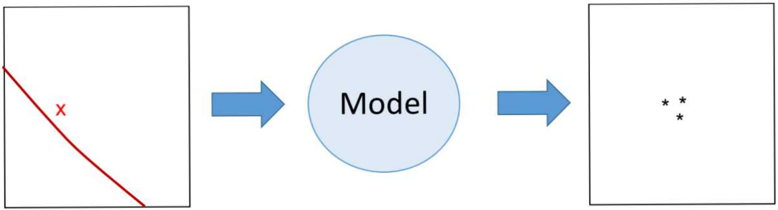


Multi-scale Materials Modeling

# Data-informed Physics-Based Predictions



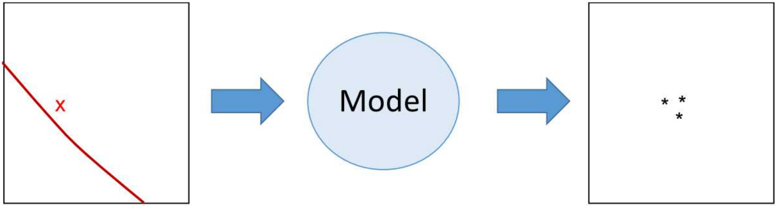
# A Deterministic Inverse Problem




## Problem

Given some observed data, find $x \in \mathbf{X}$ that best predicts the data.

# A Deterministic Inverse Problem

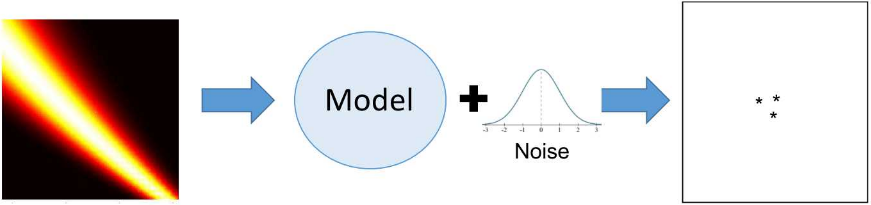



**Problem**

Given some observed data, find $x \in \mathbf{X}$ that best predicts the data.

- Solutions may not be unique without additional assumptions.
- Requires solving several deterministic forward problems.

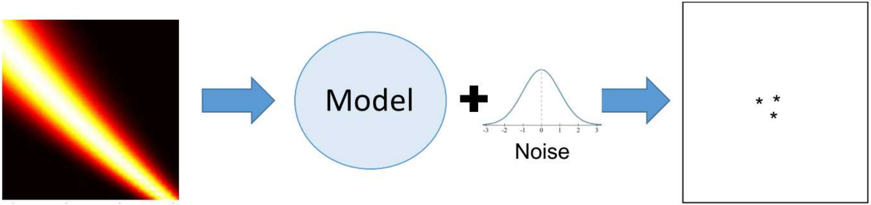
# A Stochastic Inverse Problem




## Problem

Given some observed data and an assumed noise model, find the parameters that are most likely to have produced the data.

# A Stochastic Inverse Problem

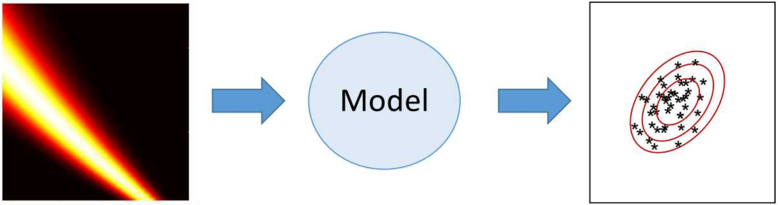



## Problem

Given some observed data and an assumed noise model, find the parameters that are most likely to have produced the data.

- Solutions may not be unique without additional assumptions.
- Requires solving several deterministic forward problems.

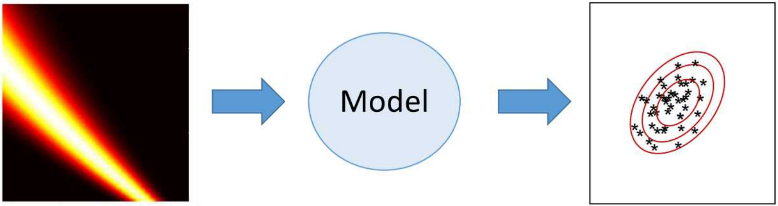
# A Different Stochastic Inverse Problem




## Problem

Given a probability density on observations, find a probability density on $\mathbf{X}$ such that the push-forward matches the given density on the observed data.

# A Different Stochastic Inverse Problem




## Problem

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

- Solutions may not be unique without additional assumptions.
- **We only need to solve a single stochastic forward problem.**

# Notation

## We assume we are given:

1. A finite-dimensional **parameter space**, $\mathbf{X}$.
2. A **parameter-to-observation/data map**, $f : \mathbf{X} \to \mathcal{D} = f(\mathbf{X})$
3. An **observed probability measure** on $(\mathcal{D}, \mathcal{B}_{\mathcal{D}})$, denoted $\mathbb{P}_{\mathcal{D}}^{\text{obs}}$, that has a density, $\pi_{\mathcal{D}}^{\text{obs}}$.
4. An **initial probability measure** on $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$, denoted $\mathbb{P}_{\mathbf{X}}^{\text{init}}$, that has a density, $\pi_{\mathbf{X}}^{\text{init}}$.

# Notation

## We assume we are given:

1. A finite-dimensional **parameter space**, $\mathbf{X}$.
2. A **parameter-to-observation/data map**, $f : \mathbf{X} \to \mathcal{D} = f(\mathbf{X})$
3. An **observed probability measure** on $(\mathcal{D}, \mathcal{B}_{\mathcal{D}})$, denoted $\mathbb{P}_{\mathcal{D}}^{\text{obs}}$, that has a density, $\pi_{\mathcal{D}}^{\text{obs}}$.
4. An **initial probability measure** on $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$, denoted $\mathbb{P}_{\mathbf{X}}^{\text{init}}$, that has a density, $\pi_{\mathbf{X}}^{\text{init}}$.

## We need to compute:

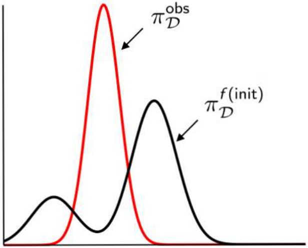
1. The **push-forward of the initial density** through the model.

- In other words, **we need to solve a forward UQ problem using the initial.**
- We use $\pi_{\mathcal{D}}^{f(\text{init})}$ to denote this push-forward density.

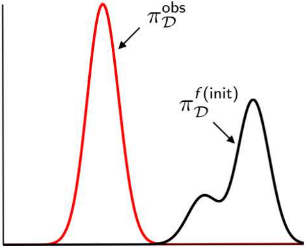
# A Key Assumption

## Predictability Assumption

We assume that the observed probability measure, $\mathbb{P}_{\mathcal{D}}^{\text{obs}}$, is absolutely continuous with respect to the push-forward of the initial, $\mathbb{P}_{\mathcal{D}}^{f(\text{init})}$.




Good Initial




Bad Initial
(Cannot predict all observations)

# A Solution to the Stochastic Inverse Problem

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

*Given an initial probability measure, $\mathbb{P}_{\mathbf{X}}^{init}$ on $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$ and an observed probability measure, $\mathbb{P}_{\mathcal{D}}^{obs}$, on $(\mathcal{D}, \mathcal{B}_{\mathcal{D}})$, the probability measure $P_{\mathbf{X}}^{up}$ on $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$ defined by*

$$\mathbb{P}_{\mathbf{X}}^{up}(A) = \int_{\mathcal{D}} \left( \int_{A \cap f^{-1}(q)} \pi_{\mathbf{X}}^{init}(x) \frac{\pi_{\mathcal{D}}^{obs}(f(x))}{\pi_{\mathcal{D}}^{f(init)}(f(x))} d\mu_{\mathbf{X},q}(x) \right) d\mu_{\mathcal{D}}(q), \; \forall A \in \mathcal{B}_{\mathbf{X}}$$

*solves the stochastic inverse problem.*

# A Solution to the Stochastic Inverse Problem

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

*Given an initial probability measure, $\mathbb{P}_{\mathbf{X}}^{init}$ on $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$ and an observed probability measure, $\mathbb{P}_{\mathcal{D}}^{obs}$, on $(\mathcal{D}, \mathcal{B}_{\mathcal{D}})$, the probability measure $P_{\mathbf{X}}^{up}$ on $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$ defined by*

$$\mathbb{P}_{\mathbf{X}}^{up}(A) = \int_{\mathcal{D}} \left( \int_{A \cap f^{-1}(q)} \pi_{\mathbf{X}}^{init}(x) \frac{\pi_{\mathcal{D}}^{obs}(f(x))}{\pi_{\mathcal{D}}^{f(init)}(f(x))} d\mu_{\mathbf{X},q}(x) \right) d\mu_{\mathcal{D}}(q), \ \forall A \in \mathcal{B}_{\mathbf{X}}$$

*solves the stochastic inverse problem.*

## Corollary (BJW., SISC 2018a)

*The updated measure of* $\mathbf{X}$ *is 1.*

# A Solution to the Stochastic Inverse Problem

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

*Given an initial probability measure, $\mathbb{P}_{\mathbf{X}}^{init}$ on $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$ and an observed probability measure, $\mathbb{P}_{\mathcal{D}}^{obs}$, on $(\mathcal{D}, \mathcal{B}_{\mathcal{D}})$, the probability measure $P_{\mathbf{X}}^{up}$ on $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$ defined by*

$$\mathbb{P}_{\mathbf{X}}^{up}(A) = \int_{\mathcal{D}} \left( \int_{A \cap f^{-1}(q)} \pi_{\mathbf{X}}^{init}(x) \frac{\pi_{\mathcal{D}}^{obs}(f(x))}{\pi_{\mathcal{D}}^{f(init)}(f(x))} d\mu_{\mathbf{X},q}(x) \right) d\mu_{\mathcal{D}}(q), \ \forall A \in \mathcal{B}_{\mathbf{X}}$$

*solves the stochastic inverse problem.*

## Corollary (BJW., SISC 2018a)

*The updated measure of $\mathbf{X}$ is 1.*

## Theorem (BJW., SISC 2018a)

*$\mathbb{P}_{\mathbf{X}}^{up}$ is stable with respect to perturbations in $\mathbb{P}_{\mathcal{D}}^{obs}$ and in $\mathbb{P}_{\mathbf{X}}^{init}$.*

# A Solution to the Stochastic Inverse Problem

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

*Given an initial probability measure,* $\mathbb{P}_{\mathbf{X}}^{init}$ *on* $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$ *and an observed probability measure,* $\mathbb{P}_{\mathcal{D}}^{obs}$*, on* $(\mathcal{D}, \mathcal{B}_{\mathcal{D}})$*, the probability measure* $P_{\mathbf{X}}^{up}$ *on* $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$ *defined by*

$$\mathbb{P}_{\mathbf{X}}^{up}(A) = \int_{\mathcal{D}} \left( \int_{A \cap f^{-1}(q)} \pi_{\mathbf{X}}^{init}(x) \frac{\pi_{\mathcal{D}}^{obs}(f(x))}{\pi_{\mathcal{D}}^{f(init)}(f(x))} d\mu_{\mathbf{X},q}(x) \right) d\mu_{\mathcal{D}}(q), \ \forall A \in \mathcal{B}_{\mathbf{X}}$$

*solves the stochastic inverse problem.*

## Corollary (BJW., SISC 2018a)

*The updated measure of* $\mathbf{X}$ *is 1.*

## Theorem (BJW., SISC 2018a)

$\mathbb{P}_{\mathbf{X}}^{up}$ *is stable with respect to perturbations in* $\mathbb{P}_{\mathcal{D}}^{obs}$ *and in* $\mathbb{P}_{\mathbf{X}}^{init}$.

For details: [Combining Push-forward Measures and Bayes' Rule to Construct Consistent Solutions to Stochastic Inverse Problems, BJW. SISC 40 (2), 2018.]

# A Solution to the Stochastic Inverse Problem

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

*Given an initial probability measure, $\mathbb{P}_{\mathbf{X}}^{init}$ on $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$ and an observed probability measure, $\mathbb{P}_{\mathcal{D}}^{obs}$, on $(\mathcal{D}, \mathcal{B}_{\mathcal{D}})$, the probability measure $P_{\mathbf{X}}^{up}$ on $(\mathbf{X}, \mathcal{B}_{\mathbf{X}})$ defined by*

$$\mathbb{P}_{\mathbf{X}}^{up}(A) = \int_{\mathcal{D}} \left( \int_{A \cap f^{-1}(q)} \pi_{\mathbf{X}}^{init}(x) \frac{\pi_{\mathcal{D}}^{obs}(f(x))}{\pi_{\mathcal{D}}^{f(init)}(f(x))} \, d\mu_{\mathbf{X},q}(x) \right) d\mu_{\mathcal{D}}(q), \ \forall A \in \mathcal{B}_{\mathbf{X}}$$

*solves the stochastic inverse problem.*

The updated density is:

$$\pi_{\mathbf{X}}^{\text{up}}(x) = \pi_{\mathbf{X}}^{\text{init}}(x) \frac{\pi_{\mathcal{D}}^{\text{obs}}(f(x))}{\pi_{\mathcal{D}}^{f(\text{init})}(f(x))}.$$

- Both $\pi_{\mathbf{X}}^{\text{init}}$ and $\pi_{\mathcal{D}}^{\text{obs}}$ are given.
- Computing $\pi_{\mathcal{D}}^{f(\text{init})}$ requires a forward propagation of the initial density.

# A Parameterized Nonlinear System

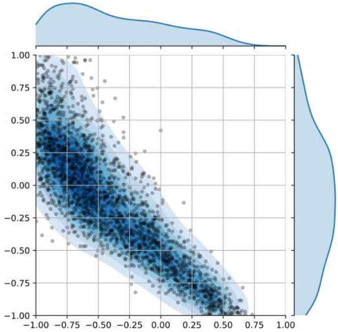
## Example

Consider a parameterized nonlinear system of equations:

$$\begin{aligned} x_1 u_1^2 + u_2^2 &= 1, \\ u_1^2 - x_2 u_2^2 &= 1 \end{aligned}$$

- The quantity of interest is the second component: $f(x) = u_2$.
- Assume that we observe $\pi_{\mathcal{D}}^{\text{obs}} \sim N(0.3, 0.025^2)$.
- We consider a uniform initial density.
- We use 10,000 samples from the initial and a standard KDE to approximate the push-forward.

# A Parameterized Nonlinear System



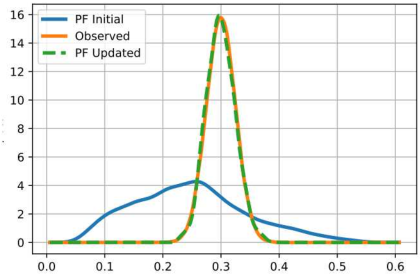




Figure: Samples from the updated density (left) and a comparison of $\pi_{\mathcal{D}}^{\text{obs}}$, $\pi_{\mathcal{D}}^{f(\text{init})}$ and $\pi_{\mathcal{D}}^{f(\text{up})}$ (right).

Additional demonstrations and interactive lecture materials can be found at https://github.com/eecsu/SIAM-AN18-Tutorial.

# Nice in theory, but not very practical!

- We cannot expect to be able to generate a large number of samples from a high-fidelity computational model!

# Nice in theory, but not very practical!

- We cannot expect to be able to generate a large number of samples from a high-fidelity computational model!
- Can we use approximate models, e.g., discretizations or surrogate models?
  - Yes, see [Butler, Jakeman, W. SISC 2018b].

# Nice in theory, but not very practical!

- We cannot expect to be able to generate a large number of samples from a high-fidelity computational model!
- Can we use approximate models, e.g., discretizations or surrogate models?
  - Yes, see [Butler, Jakeman, W. SISC 2018b].
- Can we use dimension reduction techniques, e.g., active subspaces?
  - Yes, initial results presented at SIAM CSE 2019. Journal article in preparation.

# Nice in theory, but not very practical!

- We cannot expect to be able to generate a large number of samples from a high-fidelity computational model!
- Can we use approximate models, e.g., discretizations or surrogate models?
  - Yes, see [Butler, Jakeman, W. SISC 2018b].
- Can we use dimension reduction techniques, e.g., active subspaces?
  - Yes, initial results presented at SIAM CSE 2019. Journal article in preparation.
- Can we leverage connections with deterministic optimization with regularization and randomized methods to develop scalable approaches?
  - Yes, see [Marvin, Bui-Thanh, W. CCR Proceedings 2018].

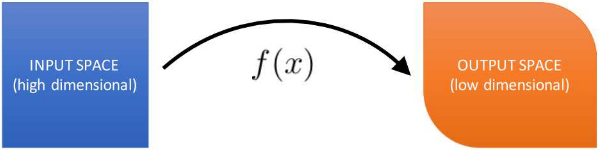
# Nice in theory, but not very practical!

- We cannot expect to be able to generate a large number of samples from a high-fidelity computational model!
- Can we use approximate models, e.g., discretizations or surrogate models?
    - Yes, see [Butler, Jakeman, W. SISC 2018b].
- Can we use dimension reduction techniques, e.g., active subspaces?
    - Yes, initial results presented at SIAM CSE 2019. Journal article in preparation.
- Can we leverage connections with deterministic optimization with regularization and randomized methods to develop scalable approaches?
    - Yes, see [Marvin, Bui-Thanh, W. CCR Proceedings 2018].
- **Can we leverage lower-fidelity models in a multi-fidelity context?**

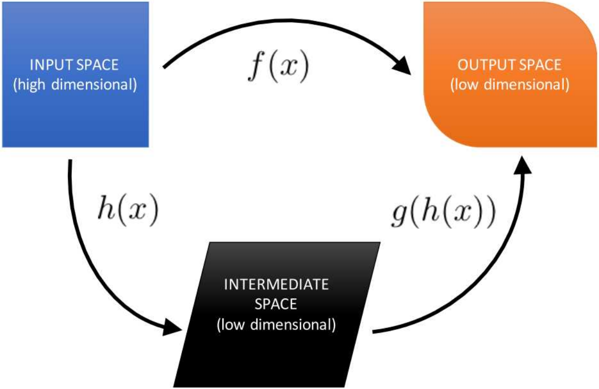
# Multilevel/Multifidelity Methods

- Objective: leverage cheaper (typically lower-fidelity) models to reduce the computational burden in performing UQ with high-fidelity (typically expensive) models.
- Multi-level Monte Carlo are theoretically robust approach for expectations (moments, probabilities of events, etc.) [Giles 2015; Cliffe et al 2011; Nobile et al 2015; Beskos et al 2017]
- Multi-index [Haji-Ali et al 2016] and multi-fidelity [Zhu et al 2017; Ng et al 2014; Perherstorfer et al 2016; Geraci et al 2015] approaches are generalization with the same objective.
- In order to solve our stochastic inverse problem, we require the push-forward of the initial density.
- Extensions of the aforementioned methodologies to probability densities have been developed [Elferson et al 2016; Giles et al 2015; Biereg et al 2016]
- Typically limited to parametric density estimates.
- We utilize an approach that is a bit more broadly applicable ...

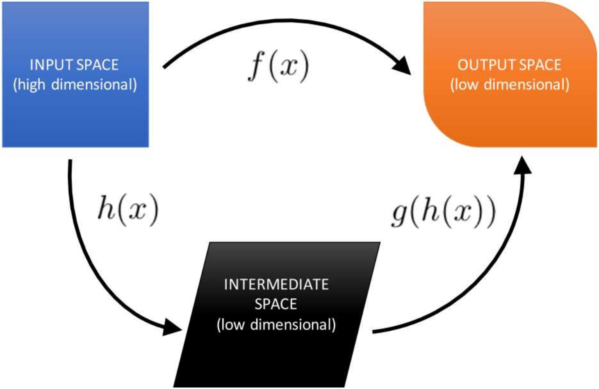
# A General Framework



# A General Framework



# A General Framework




- If $h(x)$ is lower-fidelity model, then we recover a particular multi-fidelity formulation [Koutsourelakis 2009; Biehler, Gee, Wall 2015; Bruder, Gee, W. 2019]

# A MFMC Framework

- For simplicity, consider one low-fidelity map, $f_{\text{LO}} : \mathbf{X}_{\text{LO}} \to \mathcal{D}_{\text{LO}}$, and one high-fidelity map $f_{\text{HI}} : \mathbf{X}_{\text{HI}} \to \mathcal{D}_{\text{HI}}$.
- At this point, we do not make any assumptions about a possible structure or hierarchy among the low-fidelity models.
- We further note that approximations can be very poor as long as there is a statistical dependence, even highly nonlinear, between the models.
- By basic rules of probability theory, we can rewrite this density as the marginal of a joint density:

$$\pi_{\mathcal{D}_{\text{HI}}}^{f_{\text{HI}}}(q) = \int_{\mathcal{D}_{\text{LO}}} \pi(q_{\text{HI}} | q_{\text{LO}}) \, \pi_{\mathcal{D}_{\text{LO}}}^{f_{\text{LO}}}(q_{\text{LO}}) \, \mathrm{d}q,$$

- The parameterization of the random variables $x$ is not required to be the same over all levels of fidelity,
- We do require that the parameters for a given lower fidelity model be a subset of the parameters for all of the higher fidelity models, or that there exists a relationship between the parameters of the different models.
- The key ingredient is constructing the conditional density $\pi(q_{\text{HI}} | q_{\text{LO}})$.

# Approximating the Conditional Densities

- In [Koutsourelakis 2009], a custom Bayesian regression model was employed.
- An appealing alternative is to use a Gaussian process (GP) model.
- Easy to generate samples from conditional densities.
- A GP comes with certain assumptions on the noise and joint distribution of the process variables.
- Hence, we cannot expect it to yield an exact conditional distribution, even if the number of training points becomes large.
- We can compare the GP results with those obtained using a kernel density estimator.

# Summary of the MFMC Framework with Two Models

1. Define a high-fidelity model $Q(x)$ and a low-fidelity model $q(x)$.
2. Generate $N_{\text{LO}}$ samples from the initial density and compute the low-fidelity push-forward of the initial density using a standard KDE model.
3. Generate a training data set $\{q_{\text{LO},i}, q_{\text{HI},i}\}_{i=1}^{N_{\text{HI}}}$.
4. Train a regression model on the training set, generate high-fidelity samples $q_{\text{HI},i}$ corresponding to the low-fidelity evaluations $q_{\text{LO},i}$ from the noise model and approximate the high-fidelity push-forward.

# A Very Simple Motivational Example

## Example

Consider a simple polynomial model,

$$f(x) = x^p, \quad p = 1, 3, 5, \ldots, \qquad \mathbf{\Lambda} = [-1, 1].$$

- The initial density is assumed uniform over the input space.
- The observed density a Gaussian distribution $\pi_{\mathcal{D}}^{\text{obs}}(q) = \mathcal{N}(0.25, 10^{-2})$.

- Consider a three-level hierarchy:

$$\begin{aligned} f_{\text{LO}}(x) &= x, \\ f_{\text{MID}}(x) &= x^3, \\ f_{\text{HI}}(x) &= x^5. \end{aligned}$$

- We choose training sets with $N_{\text{LO}} = 20,000$, $N_{\text{MID}} = 50$ and $N_{\text{HI}} = 25$.
- Utilize Gaussian process regression models with homoscedastic noise for the mapping between fidelities.

# Comparison of push-forward densities

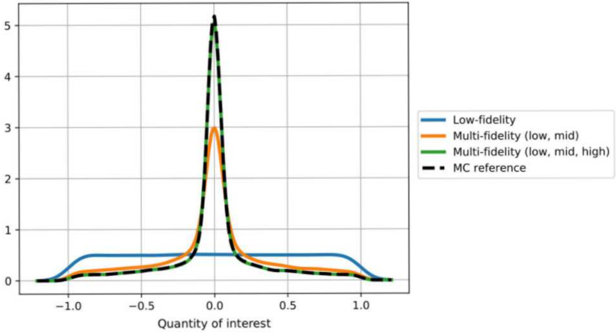



Figure: Estimated push-forward of the initial density using the low-, mid- and high-fidelity models along with the Monte Carlo reference for the one-to-one example.

# Regression models between fidelities

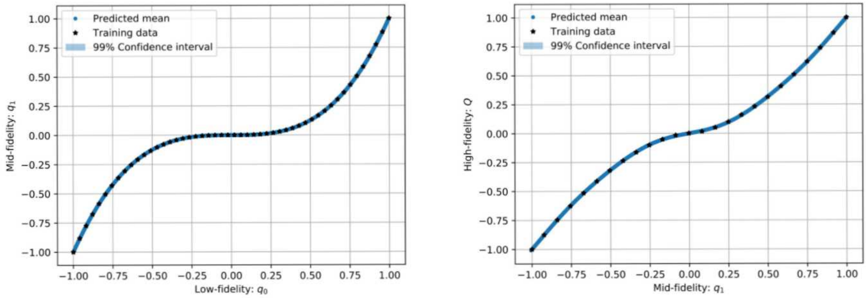
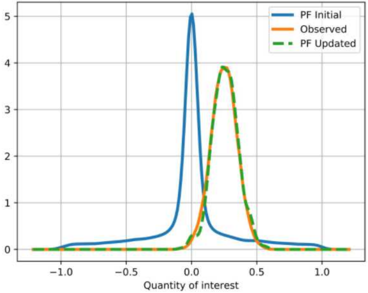



Figure: Training data, Gaussian process predicted means and 99% confidence intervals for the low- / mid-fidelity (left) and the mid- / high-fidelity (right) regression models for the one-to-one example.

# Updated densities and push-forwards



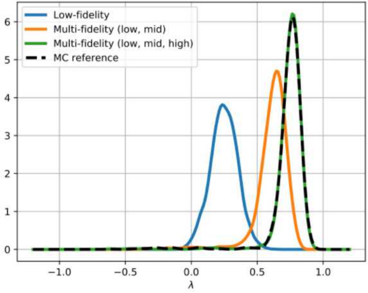




Figure: The push forward of the initial and updated densities using the multi-fidelity model (left). Updated densities using $f_{\text{LO}}$, a multi-fidelity model involving $f_{\text{LO}}$ and $f_{\text{MID}}$, and a multi-fidelity model involving $f_{\text{LO}}$, $f_{\text{MID}}$ and $f_{\text{HI}}$ (right).

# Diagnostic Data

We consider the following diagnostics:

1. The integral of the updated density (should be close to 1).
2. The Kullback-Leibler (KL) divergence between the push-forward of the updated density and the observed density (should be close to 0).
3. The information gained between the initial and updated densities, measured by the KL divergence
4. The acceptance rate from rejection sampling.

| Diagnostic | HF Monte Carlo | Multi-fidelity |
|---|---|---|
| Updated integral | 0.945883 | 0.942736 |
| $\mathrm{KL}(\pi_{\mathcal{D}}^{f(\mathrm{up})} \vert \pi_{\mathcal{D}}^{\mathrm{obs}})$ | 0.026284 | 0.013593 |
| Information gain | 1.854633 | 1.855338 |
| Acceptance rate | 7.19% | 7.10% |

Table: Diagnostic values for the one-to-one example.

# Results

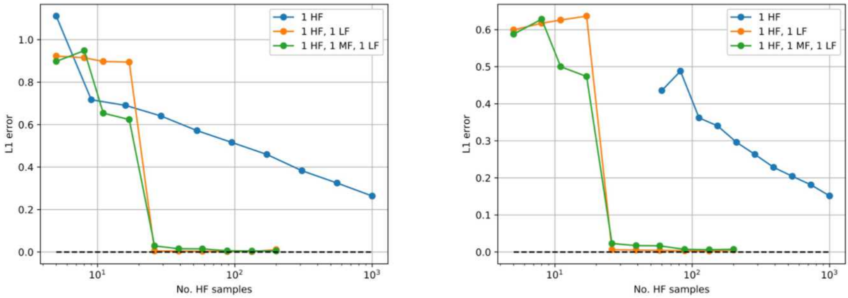



Figure: Convergence behavior of the $L_1$ error between the approximate high-fidelity push-forward of the initial density and the Monte Carlo reference with an increasing number of training samples (left) and the corresponding $L_1$ error for the updated densities for the one-to-one example (right).

# Single-phase Incompressible Flow

## Example

Consider the following elliptic partial differential equation (PDE)

$$\begin{cases} -\nabla \cdot (K(x)\nabla p) = 0, & \text{on } (0,1)^2, \\ p = 1, & \text{right side}, \\ p = 0, & \text{left side}, \\ K(x)\nabla p \cdot \mathbf{n} = 0, & \text{top and bottom}. \end{cases}$$

- $K(x)$ is a truncated Karhunen-Loéve expansion of $Y = \log K$ with 100 terms.
- The initial density is given by $\pi_{\boldsymbol{\Lambda}}^{\text{init}}(x) = \mathcal{N}(\mathbf{0}, \mathbf{I})$.
- The quantities of interest are pressure values at different locations.
- Solution approximated using a finite element discretization with three uniform grids with element sizes $h = 1/40$, $h = 1/80$ and $h = 1/160$.
- We choose training sets with $N_{\text{LO}} = 10,000$, $N_{\text{MID}} = 200$ and $N_{\text{HI}} = 20$.
- Utilize Gaussian process regression models with heteroscedastic noise for the mapping between fidelities.

# Push-forward Densities

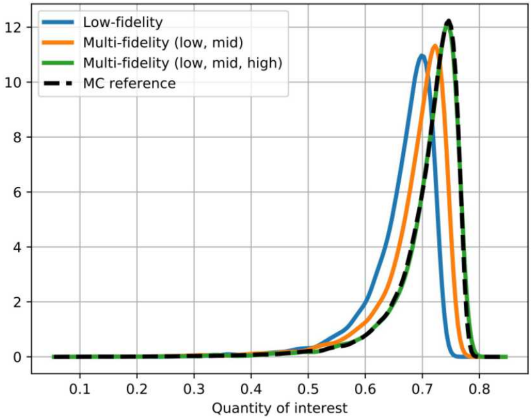



Figure: Estimated push-forward of the initial density using the low-, mid- and high-fidelity models along with the Monte Carlo reference for the elliptic PDE example.

# Diagnostic Data

| Diagnostic | HF Monte Carlo | Multi-fidelity |
|---|---|---|
| Updated integral | 0.997210 | 0.995296 |
| $\mathrm{KL}(\pi_{\mathcal{D}}^{f(\mathrm{up})} \vert \pi_{\mathcal{D}}^{\mathrm{obs}})$ | 0.001138 | 0.003231 |
| Information gain | 0.520369 | 0.519468 |
| Acceptance rate | 33.63% | 34.21% |

Table: Diagnostic values for the linear elliptic example.

# Regression Models using Heteroscedastic Noise

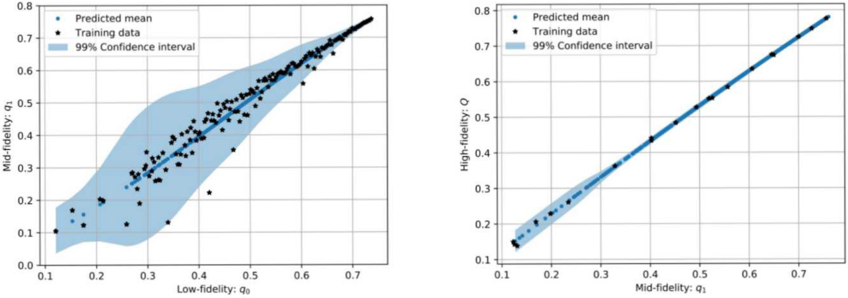



Figure: Training data, Gaussian process predicted means and 99% confidence intervals for the low- / mid-fidelity (left) and the mid- / high-fidelity (right) regression models for the elliptic PDE example.

# Samples From Regression Models

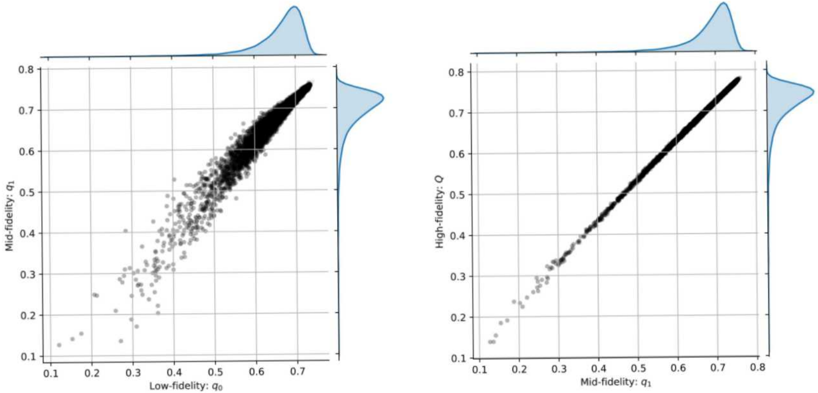



Figure: Obtained joint and marginal densities for the low- / mid-fidelity (left) and the mid- / high-fidelity (right) regression models for the elliptic PDE example.

# Convergence of the densities

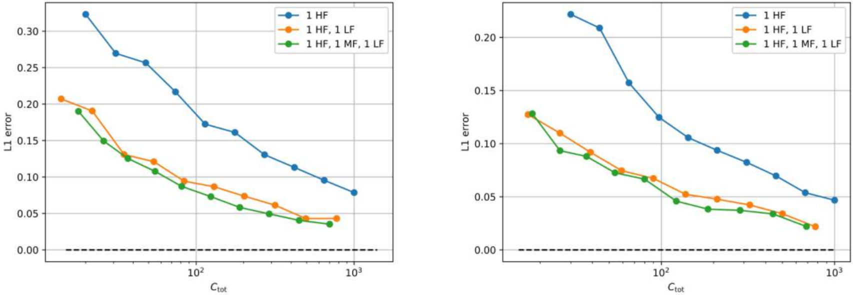



Figure: Convergence behavior of the $L_1$ error between the approximate high-fidelity push-forward of the initial density and the Monte Carlo reference (left) and the $L_1$ error in the corresponding updated densities (right), using one high-fidelity model (1 HF / Monte Carlo sampling), a high- and low-fidelity model (1 HF, 1 LF) and a high-, mid- and low-fidelity model (1 HF, 1 MF, 1 LF) for the elliptic PDE example.

# Results using GP regression and a KDE noise model

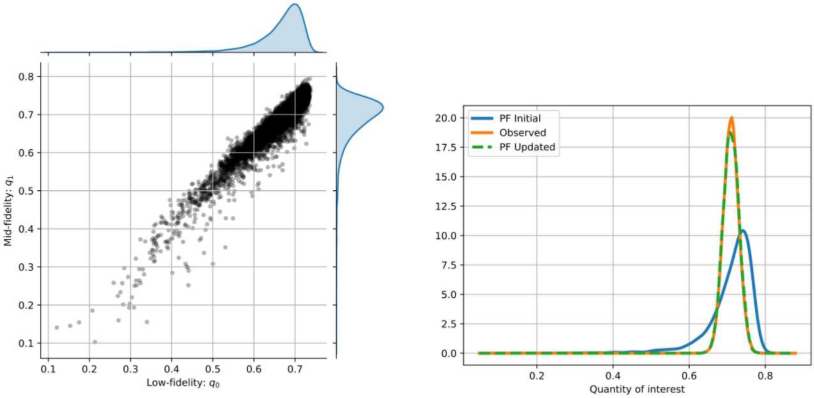



Figure: Obtained joint and marginal densities for the low- / mid-fidelity (left) and the comparison of the push-forwards of the initial and updated densities with the observed density for the elliptic PDE example with GP regression and a kernel density noise model.

# Diagnostics using GP regression and a KDE noise model

| Diagnostic | HF Monte Carlo | Multi-fidelity | Multi-fidelity (KDE) |
|---|---|---|---|
| Updated integral | 0.997210 | 0.995296 | 1.003031 |
| $\mathrm{KL}(\pi_{\mathcal{D}}^{f(\mathrm{up})} \| \pi_{\mathcal{D}}^{\mathrm{obs}})$ | 0.001138 | 0.003231 | 0.001173 |
| Information gain | 0.520369 | 0.519468 | 0.518980 |
| Acceptance rate | 33.63% | 34.21% | 36.32% |

Table: Diagnostic values for the linear elliptic example.

# Multiple quantities of interest

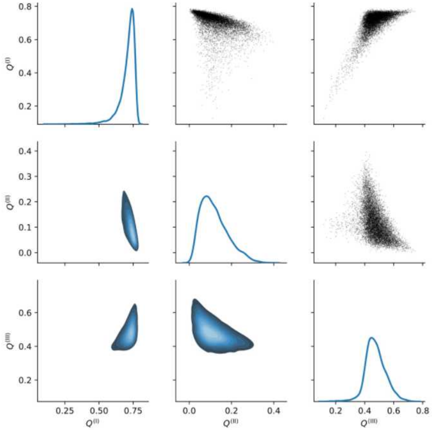



Figure: Visualization of the approximate high-fidelity push-forward of the initial density for the elliptic PDE example with multiple QoI. Marginal distributions of the three quantities of interest are shown on the diagonal and pairwise joint distributions on the off-diagonals.

# Convergence with multiple quantities of interest

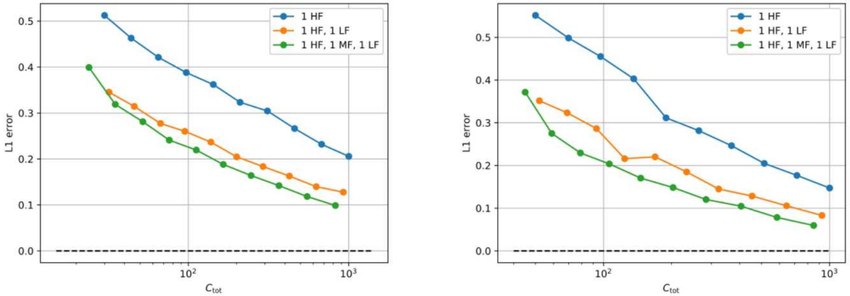


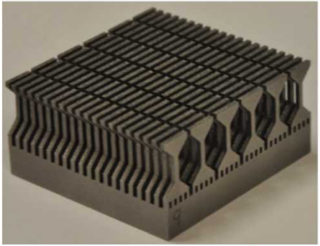

Figure: Convergence behavior of the $L_1$ error between the approximate high-fidelity push-forward of the initial density and the Monte Carlo reference (left) and the $L_1$ error in the corresponding updated densities (right), using one high-fidelity model (1 HF / Monte Carlo sampling), a high- and low-fidelity model (1 HF, 1 LF) and a high-, mid- and low-fidelity model (1 HF, 1 MF, 1 LF) for the elliptic PDE example with multiple QoI.

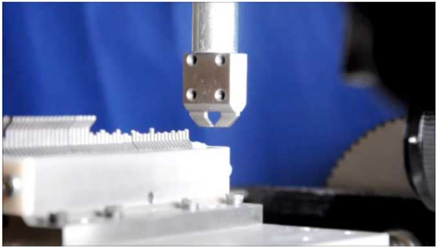
# Diagnostics with multiple quantities of interest

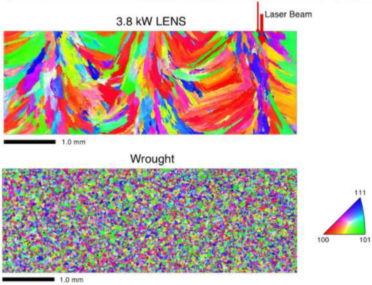
| Diagnostic | HF Monte Carlo | Multi-fidelity |
|---|---|---|
| Updated integral | 0.994670 | 1.010853 |
| $\mathrm{KL}(\pi_{\mathcal{D}}^{f(\mathrm{up})} \vert \pi_{\mathcal{D}}^{\mathrm{obs}})$ | 0.093677 | 0.187659 |
| Information gain | 1.589674 | 1.647271 |
| Acceptance rate | 4.58% | 4.47% |

Table: Diagnostic values for the linear elliptic example with multiple QoI.

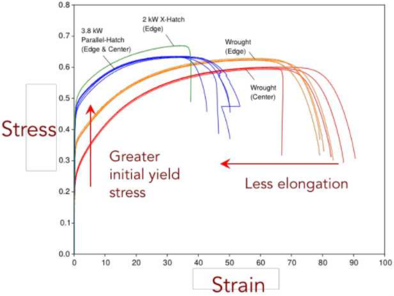
# An Example Inspired by Additive Manufacturing








304 L

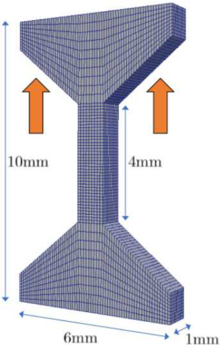


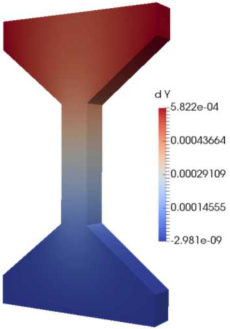
# A Low-fidelity Model

- For the low-fidelity model, we solve the linear isotropic elasticity equations,

$$-\nabla \cdot \boldsymbol{\sigma}(\mathbf{u}) = \mathbf{f}, \quad \sigma_{ij} = C_{ijkl}\epsilon_{kl}, \quad \boldsymbol{\epsilon} = \frac{1}{2}\nabla\mathbf{u} + \frac{1}{2}\nabla\mathbf{u}^T,$$

  using bilinear finite elements on a mesh with 16,600 hexahedral elements.
- Assume $E \sim \mathcal{U}(1.8 \cdot 10^4, 2.2 \cdot 10^4)$ and $\nu \sim \mathcal{U}(0.27, 0.33)$.
- Quantity of interest is the average vertical displacement within the connector.
- Assume the observed density is given by $\pi_{\mathcal{D}}^{\mathrm{obs}}(Q) = \mathcal{N}(0.95, 1 \cdot 10^{-4})$.



# A High-fidelity Model

- The high-fidelity model utilizes a crystal elasticity formulation where

$$\sigma_{ij} = \hat{C}_{ijkl}\epsilon_{kl}, \quad \hat{C}_{ijkl} = C_{mnop}R_{lp}R_{ko}R_{jn}R_{im}.$$

- Assume a grain diameter of 50$\mu$m and use well-spaced Voronoi cells to produce approximately 224,000 grains within the computational domain.
- High-fidelity model uses a finer discretization with 17,040,000 elements.
- This resolution provides approximately 76 elements per grain whereas the coarse discretization would have approximately 14 grains per element.
- Each grain is endowed with a random orientation defined by 4 independent normal random variables.
- Both models use these distributions for $E$ and $\nu$, but the high-fidelity has 896,000 additional random parameters associated with the random orientations.

# A High-fidelity Model

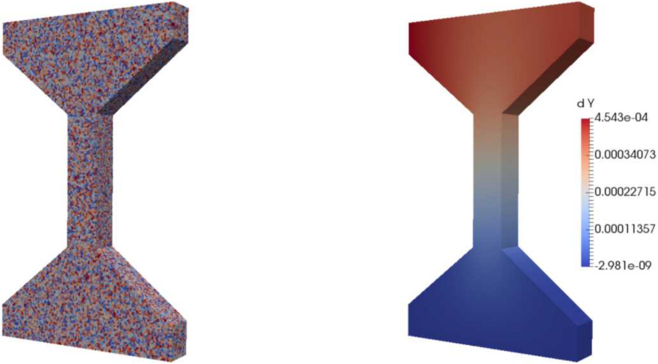
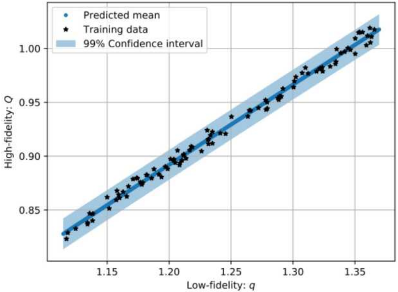



Figure: On the left, the grains for the high-fidelity model. On the right, the vertical displacement using the high-fidelity model and the nominal parameter values ($E = 2.0 \cdot 10^4$ and $\nu = 0.3$).

# Regression Model and Push-forward Densities



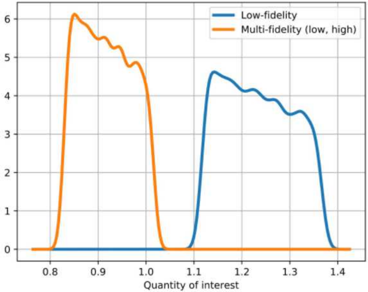
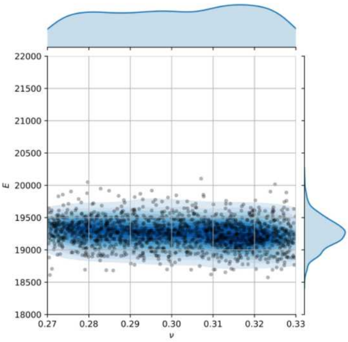




Figure: The training data, the Gaussian process mean and the 99% confidence intervals for the mapping between the low- and high-fidelity models for the linear elastic example (left). The push-forward of the initial density through the low-fidelity model and through the multi-fidelity model (right).

# Stochastic Inversion Results Using Multi-fidelity Framework



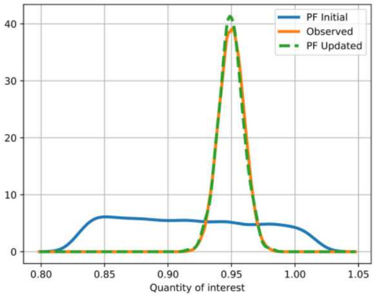




Figure: Samples from the updated density using the MFMC framework (left) and the comparison of the push-forwards of the initial and updated densities with the observed density (right) for the linear elastic example.

# Conclusions

- Our goal is to develop **data-informed physics-based** models.
- Many approaches exist for incorporating data into a model.
  - Deterministic optimization, Bayesian methods, OUU, data assimilation, etc.
- Our approach provides a solution to a specific stochastic inverse problem.
- Main computational expense is the forward UQ problem to obtain the push-forward of the initial density.
- We demonstrated that a **multi-fidelity Monte Carlo** approach can be utilized within this framework.
- Numerical results indicate that this combination works well for problems with well-correlated hierarchy of models.

# Thanks! Questions?

- SIP interactive lecture materials:

  https://github.com/eecsu/SIAM-AN18-Tutorial

- SIP combined with multifidelity framework:

  https://github.com/TimWildey/CBayes-MLMF


## Acknowledgments

L. Bruder's research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), project GE2254/1-1. This material is based upon work supported by the U.S. Department of Energy, Office of Science, ASCR, Early Career Research Program.


# Thank you for your attention!

# Questions?