

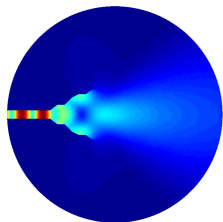
Data-Driven Multifidelity Methods for Monte Carlo Estimation

Benjamin Peherstorfer
Courant Institute of Mathematical Sciences
New York University

Collaborators: Karen Willcox, Max Gunzburger, Youssef Marzouk

September 2018

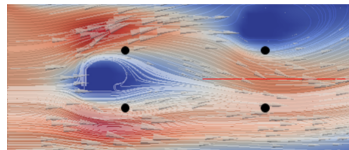
Outer loop applications



optimization



control



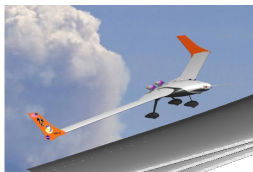
inference



visualization

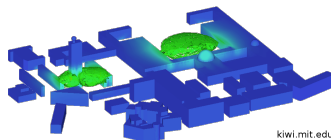


model calibration



U.S. Air Force/DLR

multi-discipline coupling



kiwi.mit.edu

uncertainty quantification

Surrogate models

Given is a high-fidelity model $f^{(1)} : \mathcal{D} \rightarrow \mathcal{Y}$

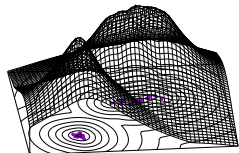
- ▶ Large-scale numerical simulation
- ▶ Achieves required accuracy
- ▶ Computationally expensive

Additionally, often have surrogate models

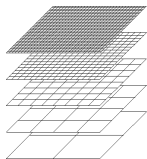
$$f^{(i)} : \mathcal{D} \rightarrow \mathcal{Y}, \quad i = 2, \dots, k$$

- ▶ Approximate high-fidelity $f^{(1)}$
- ▶ Often orders of magnitudes cheaper

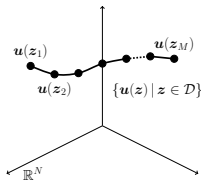
Examples of surrogate models



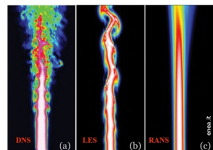
data-fit models,
response surfaces,
machine learning



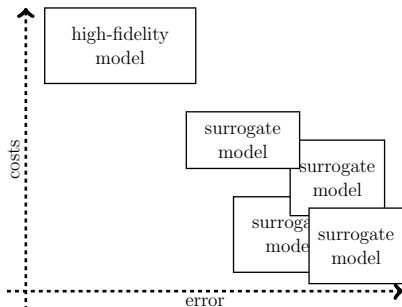
coarse-grid
approximations



reduced basis,
proper orthogonal
decomposition



simplified models,
linearized models



Replacing high-fidelity model with surrogate

Replace $f^{(1)}$ with a surrogate model

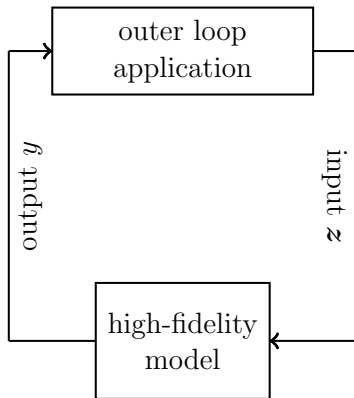
- ▶ Costs of outer loop reduced
- ▶ Often orders of magnitude speedups

Estimate depends on surrogate accuracy

- ▶ Control with error bounds/estimators
- ▶ Rebuild if accuracy too low
- ▶ No guarantees without bounds/estimators

Issues

- ▶ Propagation of surrogate error on estimate
- ▶ Surrogates without error control
- ▶ Costs of rebuilding a surrogate model



Replacing high-fidelity model with surrogate

Replace $f^{(1)}$ with a surrogate model

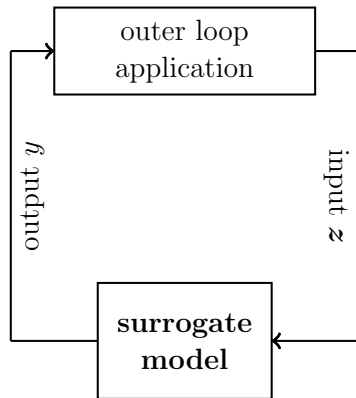
- ▶ Costs of outer loop reduced
- ▶ Often orders of magnitude speedups

Estimate depends on surrogate accuracy

- ▶ Control with error bounds/estimators
- ▶ Rebuild if accuracy too low
- ▶ No guarantees without bounds/estimators

Issues

- ▶ Propagation of surrogate error on estimate
- ▶ Surrogates without error control
- ▶ Costs of rebuilding a surrogate model



Our approach: Multifidelity methods

Combine high-fidelity and surrogate models

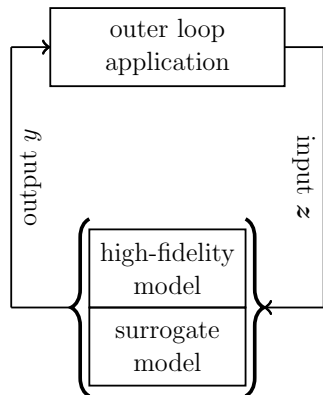
- ▶ Leverage surrogate models for speedup
- ▶ Recourse to high-fidelity for accuracy

Multifidelity guarantees high-fidelity accuracy

- ▶ Occasional recourse to high-fidelity model
- ▶ High-fidelity model is kept in the loop
- ▶ Independent of error control for surrogates

Multifidelity speeds up computations

- ▶ Adapt, fuse, filter with surrogate models
- ▶ Balance #solves among models



[Brandt, 1977], [Hackbusch, 1985], [Bramble et al, 1990], [Booker et al, 1999], [Jones et al, 1998], [Alexandrov et al, 1998], [Christen et al, 2005], [Cui et al, 2014]

[P., Willcox, Gunzburger, *Survey of multifidelity methods in uncertainty propagation, inference, and optimization*; SIAM Review, 60(3):550-591, 2018]

Our approach: Multifidelity methods

Combine high-fidelity and surrogate models

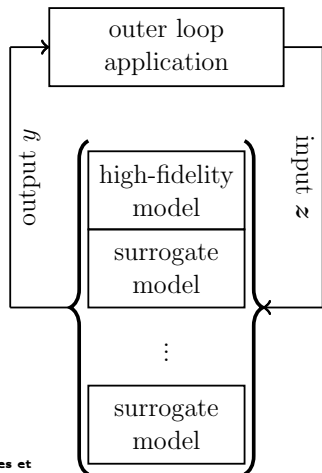
- ▶ Leverage surrogate models for speedup
- ▶ Recourse to high-fidelity for accuracy

Multifidelity guarantees high-fidelity accuracy

- ▶ Occasional recourse to high-fidelity model
- ▶ High-fidelity model is kept in the loop
- ▶ Independent of error control for surrogates

Multifidelity speeds up computations

- ▶ Adapt, fuse, filter with surrogate models
- ▶ Balance #solves among models



[Brandt, 1977], [Hackbusch, 1985], [Bramble et al, 1990], [Booker et al, 1999], [Jones et al, 1998], [Alexandrov et al, 1998], [Christen et al, 2005], [Cui et al, 2014]

[P., Willcox, Gunzburger, *Survey of multifidelity methods in uncertainty propagation, inference, and optimization*; SIAM Review, 60(3):550-591, 2018]

1. Motivation for multifidelity methods
2. Multifidelity Monte Carlo estimation (MFMC)
3. MFMC with context-aware surrogates
4. Multifidelity and Bayesian inverse problems
5. Outlook and conclusions

1. Motivation for multifidelity methods
2. **Multifidelity Monte Carlo estimation (MFMC)**
3. MFMC with context-aware surrogates
4. Multifidelity and Bayesian inverse problems
5. Outlook and conclusions

- 1 P., Willcox & Gunzburger *Optimal model management for multifidelity Monte Carlo estimation*. SISC, 2016.

Uncertainty propagation as an outer loop application

High-fidelity (“truth”) model with costs $w_1 > 0$

$$f^{(1)} : \mathcal{D} \rightarrow \mathcal{Y}$$

Given random variable Z , estimate

$$s = \mathbb{E}[f^{(1)}(Z)]$$

Monte Carlo estimator with realizations $\mathbf{z}_1, \dots, \mathbf{z}_n$ of Z

$$\bar{y}_n^{(1)} = \frac{1}{n} \sum_{i=1}^n f^{(1)}(\mathbf{z}_i)$$

Uncertainty propagation with Monte Carlo is outer-loop application

- ▶ Each high-fidelity model solve is computationally expensive
- ▶ Repeated model solves become prohibitive

MFMC: Control variates

Estimate $\mathbb{E}[A]$ of random variable A with Monte Carlo estimator

$$\bar{a}_n = \frac{1}{n} \sum_{i=1}^n a_i, \quad a_1, \dots, a_n \sim A$$

Unbiased estimator $\mathbb{E}[\bar{a}_n] = \mathbb{E}[A]$ with mean-squared error (MSE)

$$e(\bar{a}_n) = \frac{\text{Var}[A]}{n}$$

Combine \bar{a}_n with Monte Carlo estimator \bar{b}_n of $\mathbb{E}[B]$ of random variable B

$$\hat{s}_A = \bar{a}_n + \gamma (\mathbb{E}[B] - \bar{b}_n), \quad \gamma \in \mathbb{R}$$

Control variate estimator \hat{s}_A is unbiased estimator $\mathbb{E}[\hat{s}_A] = \mathbb{E}[A]$ with MSE

$$e(\hat{s}_A) = (1 - \rho^2)e(\bar{a}_n)$$

- ▶ Correlation coefficient $-1 \leq \rho \leq 1$ of A and B
- ▶ If $\rho = 0$, same MSE as regular Monte Carlo
- ▶ If $|\rho| > 0$, lower MSE
- ▶ The higher correlated, the lower MSE of \hat{s}_A

MFMC: Control variates and surrogate models

Models

- ▶ High-fidelity model $f^{(1)} : \mathcal{D} \rightarrow \mathcal{Y}$
- ▶ Surrogates $f^{(2)}, \dots, f^{(k)} : \mathcal{D} \rightarrow \mathcal{Y}$

P., Willcox, Gunzburger, *Optimal model management for multifidelity Monte Carlo estimation*. SISC, 2016

Exploit correlation of $f^{(1)}(Z)$ and $f^{(i)}(Z)$ for reducing MSE

$$\rho_i = \frac{\text{Cov}[f^{(1)}(Z), f^{(i)}(Z)]}{\sqrt{\text{Var}[f^{(1)}(Z)] \text{Var}[f^{(i)}(Z)]}}, \quad i = 2, \dots, k$$

Related work: Combine multiple models for Monte Carlo estimation

- ▶ Multilevel Monte Carlo [Giles 2008], [Heinrich 2001], [Speight, 2009]
- ▶ RBM and control variates [Boyaval et al, 2010, 2012], [Vidal et al 2015]
- ▶ Data-fit models and control variates [Tracey et al 2013]
- ▶ Monte Carlo with low-/high-fidelity model [Ng & Eldred 2012]
- ▶ Two models and control variates [Ng & Willcox 2012, 2014]

⇒ **Need for arbitrary number of surrogates, any type of surrogates**

MFMC: Multifidelity Monte Carlo estimator

Take realizations of input random variable Z

P., Willcox, Gunzburger, *Optimal model management for multifidelity Monte Carlo estimation*. SISC, 2016

$$z_1, z_2, z_3, \dots$$

Evaluate model $f^{(i)}$ at first m_i realizations z_1, \dots, z_{m_i} of Z

$$f^{(i)}(z_1), \dots, f^{(i)}(z_{m_i}), \quad i = 1, \dots, k$$

Multifidelity Monte Carlo (MFMC) estimator

$$\hat{s} = \underbrace{\bar{y}_{m_1}^{(1)}}_{\text{from HFM}} + \sum_{i=2}^k \gamma_i \underbrace{\left(\bar{y}_{m_i}^{(i)} - \bar{y}_{m_{i-1}}^{(i)} \right)}_{\text{from surrogates}}$$

- ▶ MFMC estimator \hat{s} is unbiased estimator of $s = \mathbb{E}[f^{(1)}(Z)]$
- ▶ Costs of each model evaluation $0 < w_1, \dots, w_k \in \mathbb{R}$ give costs of MFMC

$$c(\hat{s}) = \sum_{i=1}^k m_i w_i$$

- ▶ Selection of coefficients $\gamma_2, \dots, \gamma_k$ and model evaluations m_1, \dots, m_k ?
- ▶ Comparison in terms of costs/MSE to regular Monte Carlo estimation?

MFMC: Balancing work among models

Variance of MFMC estimator \hat{s} is

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

- ▶ Variance σ_i^2 of $f^{(i)}(Z)$
- ▶ Correlation coefficient ρ_i between $f^{(1)}(Z)$ and $f^{(i)}(Z)$

Find \mathbf{m} and γ that **minimize MSE for given computational budget q**

$$\begin{aligned} & \arg \min_{\mathbf{m} \in \mathbb{R}^k, \gamma_2, \dots, \gamma_k \in \mathbb{R}} \text{Var}[\hat{s}] \\ & \text{subject to} \quad m_{i-1} - m_i \leq 0, \quad i = 2, \dots, k, \\ & \quad \quad \quad -m_1 \leq 0, \\ & \quad \quad \quad \mathbf{w}^T \mathbf{m} = q. \end{aligned}$$

Theorem 1 (P., Willcox, Gunzburger, 2016).

Optimization problem has unique (analytic) solution if $\rho_1^2 > \dots > \rho_k^2 > 0$ and

$$\frac{w_{i-1}}{w_i} > \frac{\rho_{i-1}^2 - \rho_i^2}{\rho_i^2 - \rho_{i+1}^2}, \quad i = 2, \dots, k \quad (1)$$

Sketch of proof

- ▶ Establish necessary condition for local optima with Karush-Kuhn-Tucker
- ▶ Only one local optima with $m_1 < m_2 < \dots < m_k$
- ▶ This local optima has smaller objective value than any with “ \leq ”

Variance reduction of MFMC \hat{s} w.r.t. benchmark Monte Carlo $\bar{y}_q^{(1)}$

$$e(\hat{s}) = \left(\sum_{i=1}^k \sqrt{\frac{w_i}{w_1} (\rho_i^2 - \rho_{i+1}^2)} \right)^2 e(\bar{y}_q^{(1)})$$

MFMC: Numerical example

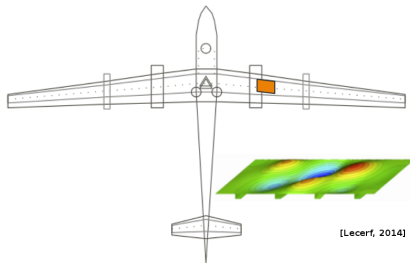
Locally damaged plate in bending

- ▶ Inputs: nominal thickness, load, damage
- ▶ Output: maximum deflection of plate
- ▶ **Only distribution of inputs known**
- ▶ Estimate **expected** deflection

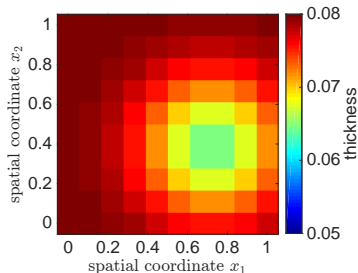
Six models

- ▶ High-fidelity model: FEM, 300 DoFs
- ▶ Reduced model: POD, 10 DoFs
- ▶ Reduced model: POD, 5 DoFs
- ▶ Reduced model: POD, 2 DoFs
- ▶ Data-fit model: linear interp., 256 pts
- ▶ Support vector machine: 256 pts

Var, corr, and costs est. from 100 samples

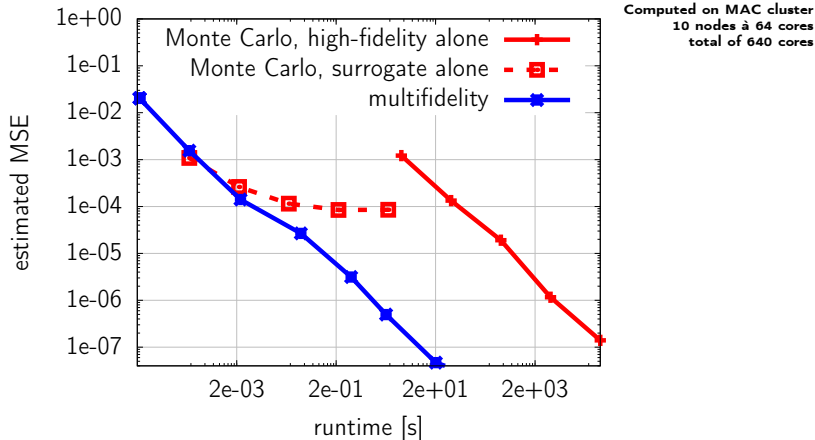


(a) wing panel



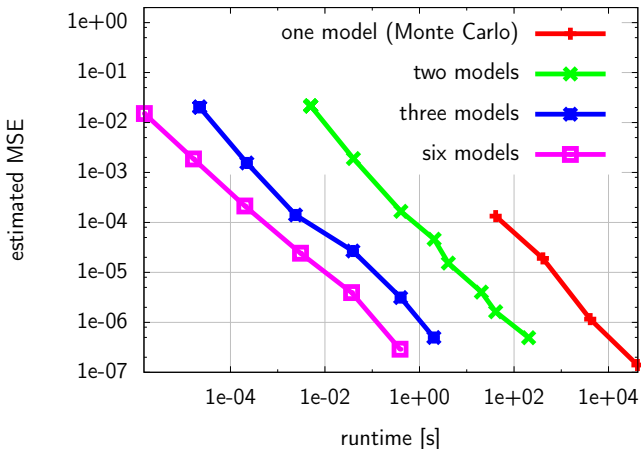
(b) damaged plate

MFMC: Speedups in uncertainty propagation



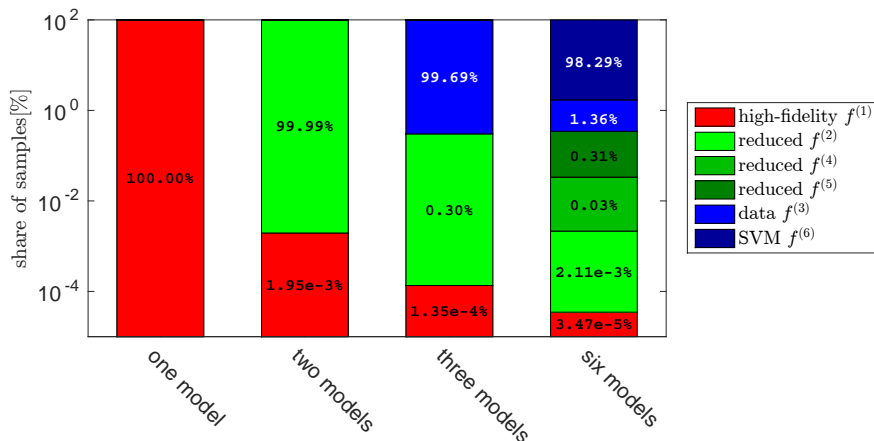
- ▶ Monte Carlo needs **12h runtime** for estimate with error below 10^{-7}
- ▶ Multifidelity provides estimator with error below 10^{-7} after **9 seconds**

MFMC: Combining many models



- ▶ Largest improvement from “single \rightarrow two” and “two \rightarrow three”
- ▶ Adding yet another reduced/SVM model reduces variance only slightly

MFMC: Distribution of #evals among models



- ▶ MFMC distributes #evals among models depending on corr/costs
- ▶ Number of evaluation changes exponentially between models
- ▶ Highest #evals in data-fit models (cost ratio $w_1/w_6 \approx 10^6$)

MFMC: Who else is using MFMC?

Multifidelity sensitivity analysis

- ▶ Identify the parameters of model with largest influence on quantity of interest
- ▶ Large-scale variance estimation problem
- ▶ Multifidelity makes tractable global sensitivity analysis with expensive models

→ Qian (MIT) with Earth Science at LANL

Uncertainty quantification in flutter problem

- ▶ Highly flexible, high-aspect-ratio wing
- ▶ Air density and root angle of attack uncertain
- ▶ Estimate expected flutter speed
- ▶ MFMC reduced runtime by more than 3 orders of magnitude

→ with Air Force Research Laboratory

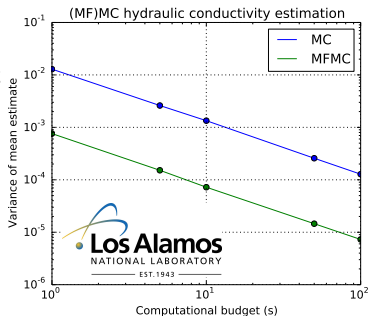


Figure: Elizabeth Qian

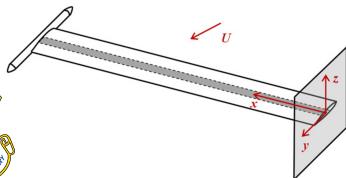


Figure: Philip S. Beran

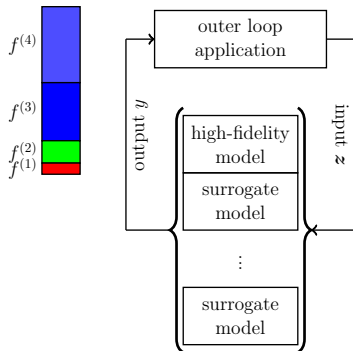
1. Motivation for multifidelity methods
2. Multifidelity Monte Carlo estimation (MFMC)
3. MFMC with context-aware surrogates
4. Multifidelity and Bayesian inverse problems
5. Outlook and conclusions

1. Motivation for multifidelity methods
2. Multifidelity Monte Carlo estimation (MFMC)
3. **MFMC with context-aware surrogates**
4. Multifidelity and Bayesian inverse problems
5. Outlook and conclusions

1 P., *Multifidelity Monte Carlo estimation with adaptive low-fidelity models*. submitted, 2017.

AMFMC: Integrate model reduction into MFMC

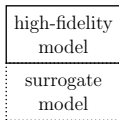
online budget



Adaptive MFMC (AMFMC)

- ▶ Trade off adaptation (“deterministic approximation”) and sampling
- ▶ Surrogate model is constructed with *outer-loop result* in mind
- ▶ Related to “exploration vs. exploitation” in Bayesian optimization
- ▶ Constructing goal-oriented surrogates [Oden et al, 2000], [Bui-Thanh et al, 2007], [Lieberman and Willcox, 2013], [Spantini et al, 2017], [Li et al, 2018]

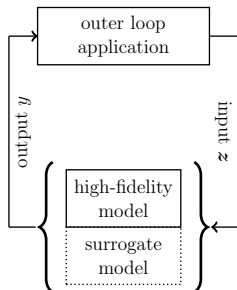
AMFMC: Integrate model reduction into MFMC



Adaptive MFMC (AMFMC)

- ▶ Trade off adaptation (“deterministic approximation”) and sampling
- ▶ Surrogate model is constructed with *outer-loop result* in mind
- ▶ Related to “exploration vs. exploitation” in Bayesian optimization
- ▶ Constructing goal-oriented surrogates [Oden et al, 2000], [Bui-Thanh et al, 2007], [Lieberman and Willcox, 2013], [Spantini et al, 2017], [Li et al, 2018]

AMFMC: Integrate model reduction into MFMC



Construct goal-oriented and context-aware surrogate for outer-loop application at hand

- Goal is outer-loop result
- Other models set the context in which surrogate model will be used

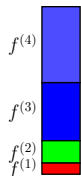
Cheap surrogate with poor approximation quality might be more useful than an expensive one that is more accurate

Adaptive MFMC (AMFMC)

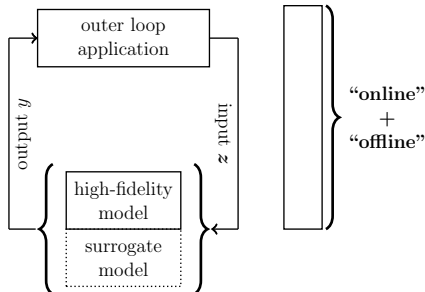
- ▶ Trade off adaptation (“deterministic approximation”) and sampling
- ▶ Surrogate model is constructed with *outer-loop result* in mind
- ▶ Related to “exploration vs. exploitation” in Bayesian optimization
- ▶ Constructing goal-oriented surrogates [Oden et al, 2000], [Bui-Thanh et al, 2007], [Lieberman and Willcox, 2013], [Spantini et al, 2017], [Li et al, 2018]

AMFMC: Integrate model reduction into MFMC

online budget



total budget

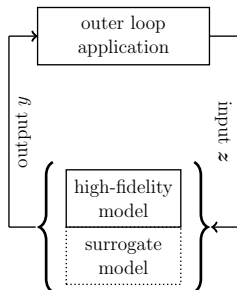
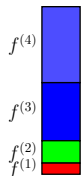


Adaptive MFMC (AMFMC)

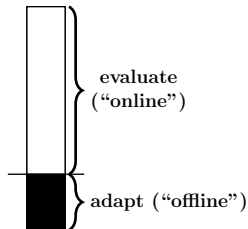
- ▶ Trade off adaptation (“deterministic approximation”) and sampling
- ▶ Surrogate model is constructed with *outer-loop result* in mind
- ▶ Related to “exploration vs. exploitation” in Bayesian optimization
- ▶ Constructing goal-oriented surrogates [Oden et al, 2000], [Bui-Thanh et al, 2007], [Lieberman and Willcox, 2013], [Spantini et al, 2017], [Li et al, 2018]

AMFMC: Integrate model reduction into MFMC

online budget



total budget

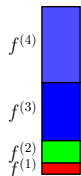


Adaptive MFMC (AMFMC)

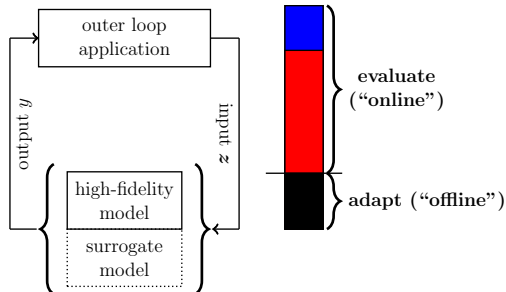
- ▶ Trade off adaptation (“deterministic approximation”) and sampling
- ▶ Surrogate model is constructed with *outer-loop result* in mind
- ▶ Related to “exploration vs. exploitation” in Bayesian optimization
- ▶ Constructing goal-oriented surrogates [Oden et al, 2000], [Bui-Thanh et al, 2007], [Lieberman and Willcox, 2013], [Spantini et al, 2017], [Li et al, 2018]

AMFMC: Integrate model reduction into MFMC

online budget



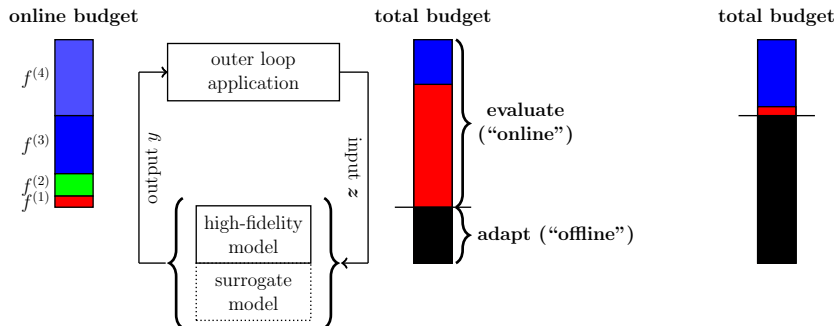
total budget



Adaptive MFMC (AMFMC)

- ▶ Trade off adaptation (“deterministic approximation”) and sampling
- ▶ Surrogate model is constructed with *outer-loop result* in mind
- ▶ Related to “exploration vs. exploitation” in Bayesian optimization
- ▶ Constructing goal-oriented surrogates [Oden et al, 2000], [Bui-Thanh et al, 2007], [Lieberman and Willcox, 2013], [Spantini et al, 2017], [Li et al, 2018]

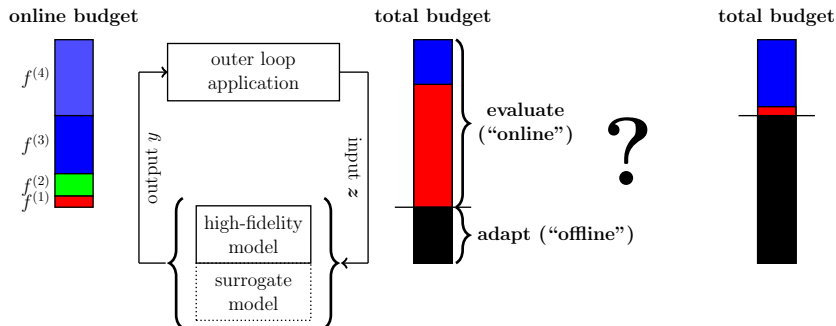
AMFMC: Integrate model reduction into MFMC



Adaptive MFMC (AMFMC)

- ▶ Trade off adaptation (“deterministic approximation”) and sampling
- ▶ Surrogate model is constructed with *outer-loop result* in mind
- ▶ Related to “exploration vs. exploitation” in Bayesian optimization
- ▶ Constructing goal-oriented surrogates [Oden et al, 2000], [Bui-Thanh et al, 2007], [Lieberman and Willcox, 2013], [Spantini et al, 2017], [Li et al, 2018]

AMFMC: Integrate model reduction into MFMC



Adaptive MFMC (AMFMC)

- ▶ Trade off adaptation (“deterministic approximation”) and sampling
- ▶ Surrogate model is constructed with *outer-loop result* in mind
- ▶ Related to “exploration vs. exploitation” in Bayesian optimization
- ▶ Constructing goal-oriented surrogates [Oden et al, 2000], [Bui-Thanh et al, 2007], [Lieberman and Willcox, 2013], [Spantini et al, 2017], [Li et al, 2018]

AMFMC: Problem setup

High-fidelity model with normalized evaluation costs $w_0 = 1$

$$f : \mathcal{D} \rightarrow \mathcal{Y}$$

Surrogate model with $n \in \mathbb{N}$

$$f^{(n)} : \mathcal{D} \rightarrow \mathcal{Y}$$

Surrogate model approximates high-fidelity model in the sense

$$1 - \rho_n^2 \leq c_1 n^{-\alpha}, \quad 0 < c_1, \alpha$$

Evaluation costs of surrogate model may grow with n as

$$w_n \leq c_2 n^\beta, \quad 0 < c_2, \beta$$

Costs of constructing surrogate $f^{(n)}$ are $w_0 n = n$

- ▶ Constructing $f^{(n)}$ requires n evaluations of f
- ▶ Evaluations of f dominate construction costs
- ▶ Construct costs are significant (e.g., model reduction)

AMFMC: Trading off construction costs and sampling costs

MFMC estimator \hat{s} with f and $f^{(n)}$ and (“online”) budget q has MSE

$$e(\hat{s}) = \frac{\sigma^2}{q} \left(\sqrt{1 - \rho_n^2} + \sqrt{w_n \rho_n^2} \right)^2$$

AMFMC splits total budget p between construction and sampling

- ▶ If spend n for constructing $f^{(n)}$, budget $q = p - n$ remains for sampling

$$e(\hat{s}_n) = \frac{\sigma^2}{p - n} \left(\sqrt{1 - \rho_n^2} + \sqrt{w_n \rho_n^2} \right)^2$$

- ▶ Measures error with respect to goal of estimating $\mathbb{E}[f(Z)]$
- ▶ Takes construction costs n into account
- ▶ Measures efficacy of surrogate model for variance reduction (context)

Upper bound on $e(\hat{s}_n)$

$$e(\hat{s}_n) \lesssim \frac{1}{p - n} (c_1 n^{-\alpha} + c_2 n^\beta)$$

AMFMC: Existence, uniqueness, and convexity

Consider the objective function

$$g(n) = \frac{1}{p-n} (c_1 n^{-\alpha} + c_2 n^\beta)$$

Find n such that $g(n)$ is minimized

$$\min_{n \in (0, p)} g(n)$$

Theorem 2 (P., 2017).

The objective g is convex in $(0, p)$ and therefore there exists a unique $\hat{n}^ \in (0, p)$ that minimizes $g(n) \Rightarrow$ there is an optimal trade-off*

Define the AMFMC estimator \hat{S}_n^*

- ▶ Computes \hat{n}^* evaluations of f to construct surrogate $f^{(\hat{n}^*)}$
- ▶ Use MFMC to combine f and surrogate $f^{(\hat{n}^*)}$ with budget $p - \hat{n}^*$

AMFMC: Adaptive multifidelity Monte Carlo estimator

Upper bound for \hat{n}^* that is useful for “small” budgets p

$$\hat{n}^* \leq \frac{\alpha}{\alpha + 1} p$$

There exists $\bar{n}^* \in \mathbb{N}$ independent of p such that $\hat{n}^* \leq \bar{n}^*$ for $p > 0$

- ▶ Number of adaptations \hat{n}^* is bounded with respect to p
- ▶ Stop adapting surrogate model even with unlimited budget $p \rightarrow \infty$
- ▶ Surrogate models can be “too accurate” for multifidelity methods

Corollary 3 (P., 2017).

Cost complexity of AMFMC with $w_n = 0$ is

$$e(\hat{S}_n^*) \in \mathcal{O}(p^{-1-\alpha})$$

- ▶ Can interpret $w_n = 0$ as $\mathbb{E}[f^{(\hat{n}^*)}(Z)]$ is known
 \Rightarrow control functionals [Oates, Girolami, Chopin, 2016]
- ▶ Helps to understand case $w_n \ll 1$ ($f^{(\hat{n}^*)}$ much cheaper than f)

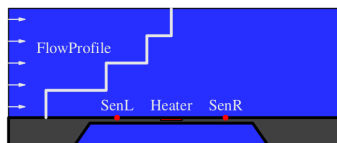
AMFMC: Anemometer

Anemometer problem

- ▶ Measure velocity of fluid
- ▶ Three inputs uniformly distributed in

$$[0, 10] \times [0.1, 10] \times [1, 10]$$

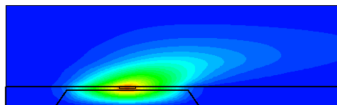
- ▶ Output is velocity
- ▶ Estimate expected velocity



Figures: MORWiki

High-fidelity model

- ▶ Based on convection-diffusion equation
- ▶ Discretized with finite elements
- ▶ High-fidelity model has 29008 DoFs



AMFMC: Surrogate model for anemometer problem

Surrogate model

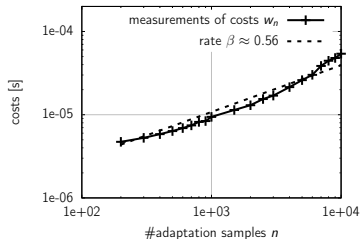
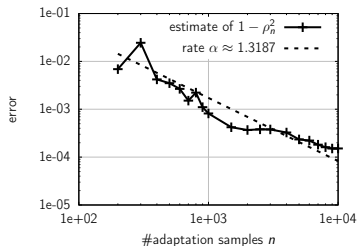
- ▶ Gaussian process regression
- ▶ Take n realizations of Z
- ▶ Train on corresponding n outputs of f

Optimizing for \hat{n}^*

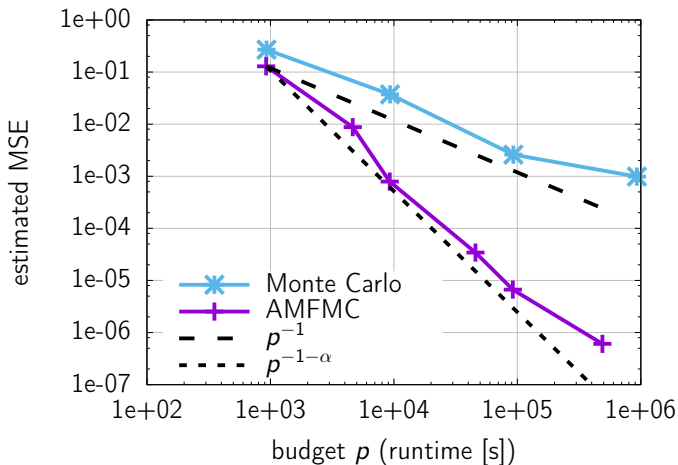
- ▶ One dimensional convex problem
- ▶ Numerically solve for \hat{n}^*

Adaptation of surrogate in AMFMC

- ▶ Numerically estimate rates from pilot runs
- ▶ Optimize for \hat{n}^* with Matlab's `fmincon`

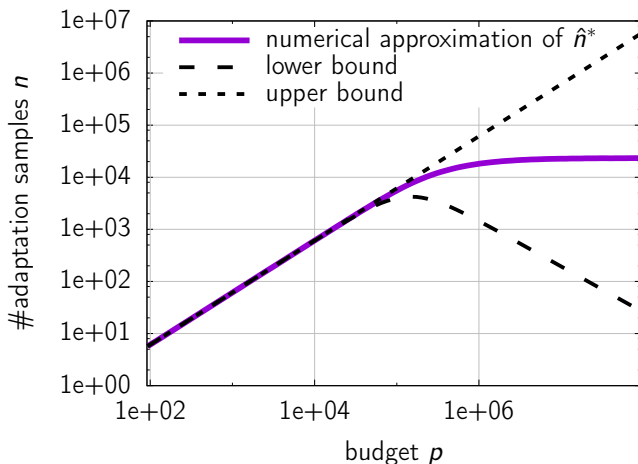


AMFMC: Anemometer results



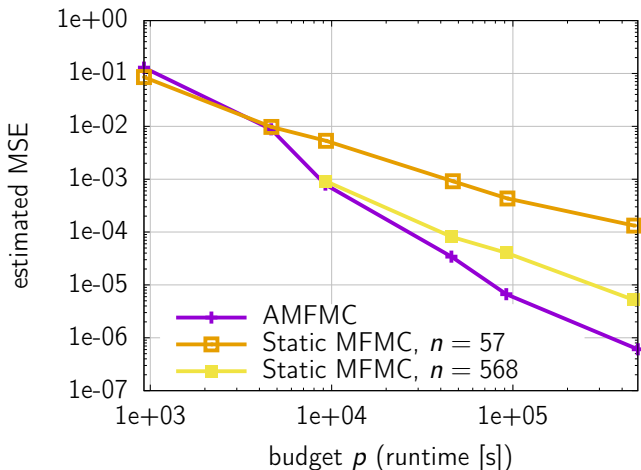
- ▶ Speedups of up to 3 orders of magnitude compared crude Monte Carlo
- ▶ MSE of AMFMC decays with $p^{-1-\alpha}$ in *pre-asymptotic* regime

AMFMC: Anemometer optimal trade-off



- ▶ Approximation of \hat{n}^* is bounded
- ▶ Lower and upper bounds seem tight in *pre-asymptotic* regime

AMFMC: Comparison to static models



- ▶ AMFMC optimally trades off adaptation and sampling costs
- ▶ Up to two orders of magnitude speedups compared to static models

AMFMC: Beam example

Beam problem

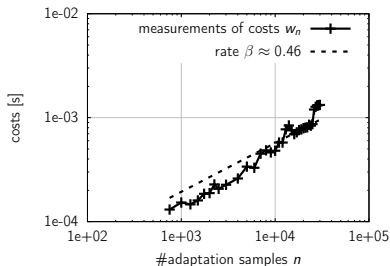
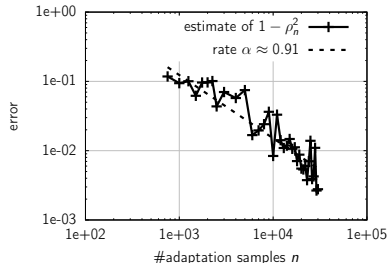
- ▶ Length and height uniformly distributed

$$[0.8, 1.2] \times [5 \times 10^{-4}, 5 \times 10^{-3}]$$

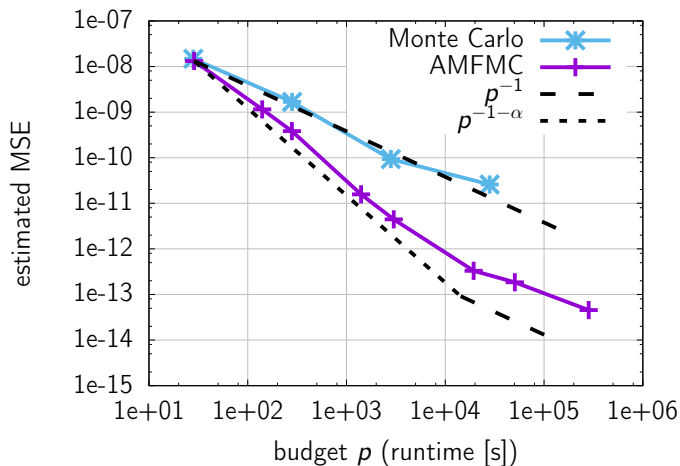
- ▶ Output is displacement of beam
- ▶ Estimate expected displacement

Models

- ▶ High-fidelity finite element model
- ▶ Surrogate is Gaussian process model
- ▶ Measure rates numerically



AMFMC: Beam results



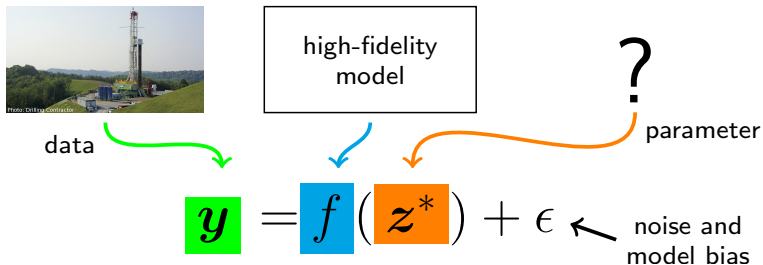
- ▶ AMFMC achieves about an order of magnitude speedup
- ▶ Decay of MSE slows down from $p^{-1-\alpha}$ to p^{-1}

1. Motivation for multifidelity methods
2. Multifidelity Monte Carlo estimation (MFMC)
3. MFMC with context-aware surrogates
4. Multifidelity and Bayesian inverse problems
5. Outlook and conclusions

1. Motivation for multifidelity methods
2. Multifidelity Monte Carlo estimation (MFMC)
3. MFMC with context-aware surrogates
- 4. Multifidelity and Bayesian inverse problems**
5. Outlook and conclusions

¹ P. and Marzouk, *A transport-based multifidelity preconditioner for Markov chain Monte Carlo*, arXiv:1808.09379, 2018 (*submitted*).

MFTMap: Bayesian inference



Bayesian inference of inputs \mathbf{z} from data \mathbf{y}

- ▶ Parameters represented as random variable \mathbf{z} with prior $\pi_0(\mathbf{z})$
- ▶ Define likelihood $L_{\mathbf{y}} : \mathcal{D} \rightarrow \mathbb{R}$ of data \mathbf{y}
- ▶ Update distribution of \mathbf{z} (“infer”) with Bayes’ rule

$$\underbrace{\pi(\mathbf{z})}_{\text{posterior}} \propto \underbrace{L_{\mathbf{y}}(\mathbf{z})}_{\text{likelihood}} \underbrace{\pi_0(\mathbf{z})}_{\text{prior}}$$

Sampling posterior π

- ▶ Evaluate posterior expectation for function g

$$\mathbb{E}_{\pi}[g] = \int g(\mathbf{z})\pi(\mathbf{z})d\mathbf{z}$$

- ▶ Samples required as inputs in upstream simulations
- ▶ Explore posterior to decide where to take new data points
- ▶ Estimate quantiles

Making sampling tractable

- ▶ Low-cost approximation of forward model (model reduction)
- ▶ Sampling scheme that exploits structure in problem

Metropolis-Hastings

- ▶ Propose a sample $\mathbf{z}' \sim q(\cdot | \mathbf{z}_{i-1})$
- ▶ Compute acceptance probability

$$\alpha(\mathbf{z}_{i-1}, \mathbf{z}') = \min \left\{ 1, \frac{\pi(\mathbf{z}')q(\mathbf{z}_{i-1} | \mathbf{z}')}{\pi(\mathbf{z}_{i-1})q(\mathbf{z}' | \mathbf{z}_{i-1})} \right\}$$

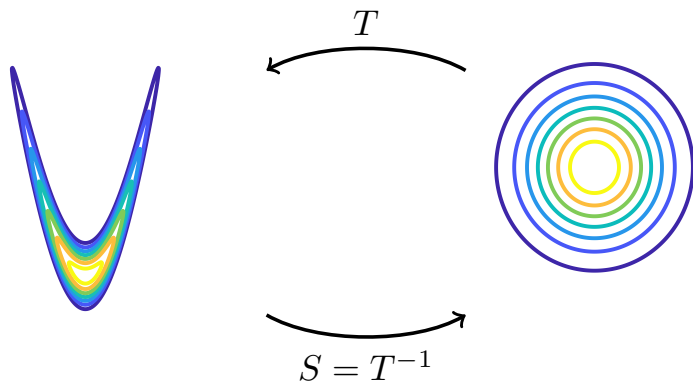
- ▶ Update

$$\mathbf{z}_i = \begin{cases} \mathbf{z}' & \text{with probability } \alpha \\ \mathbf{z}_{i-1} & \text{with probability } 1 - \alpha \end{cases}$$



Effective proposal distribution q required

Transform posterior for better sampling

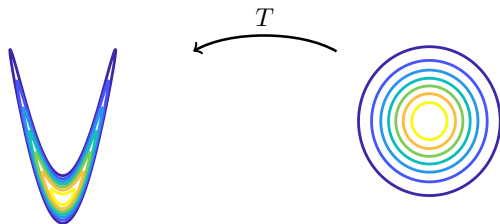


[Marzouk, Moselhy, Parno, Spantini: *Sampling via measure transport: An introduction*, Handbook of Uncertainty Quantification, 2016],

[Spantini, Bigoni, Marzouk: Inference via low-dimensional couplings. arXiv preprint arXiv:1703.06131, 2017]

MFTMap: Two ways to use transport maps

Offline/online approach [Moselhy & Marzouk, 2012]



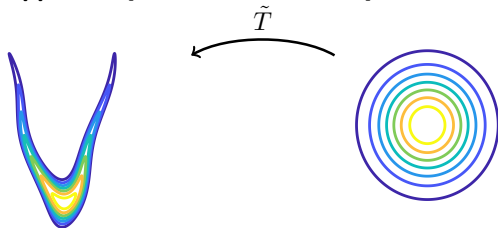
- ▶ Find map from prior to posterior
- ▶ One time high costs to construct map (nonconvex optimization)
- ▶ No samples from posterior necessary to construct map
- ▶ BUT: Only *approximate* posterior samples

Transport maps to accelerate MCMC [Parno & Marzouk, 2014, 2018]

- ▶ Adapt map during MCMC sampling
- ▶ Optimization problem becomes convex and cheaper to solve
- ▶ Asymptotically, samples from posterior
- ▶ BUT: Posterior samples required for initialization (adaptive MCMC)

MFTMap: Two ways to use transport maps

Offline/online approach [Moselhy & Marzouk, 2012]



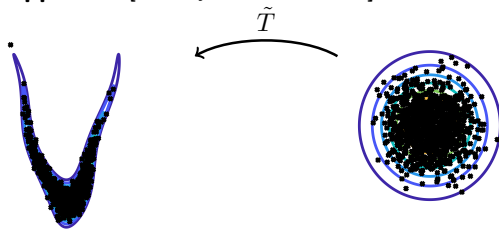
- ▶ Find map from prior to posterior
- ▶ One time high costs to construct map (nonconvex optimization)
- ▶ No samples from posterior necessary to construct map
- ▶ BUT: Only *approximate* posterior samples

Transport maps to accelerate MCMC [Parno & Marzouk, 2014, 2018]

- ▶ Adapt map during MCMC sampling
- ▶ Optimization problem becomes convex and cheaper to solve
- ▶ Asymptotically, samples from posterior
- ▶ BUT: Posterior samples required for initialization (adaptive MCMC)

MFTMap: Two ways to use transport maps

Offline/online approach [Moselhy & Marzouk, 2012]



- ▶ Find map from prior to posterior
- ▶ One time high costs to construct map (nonconvex optimization)
- ▶ No samples from posterior necessary to construct map
- ▶ BUT: Only *approximate* posterior samples

Transport maps to accelerate MCMC [Parno & Marzouk, 2014, 2018]

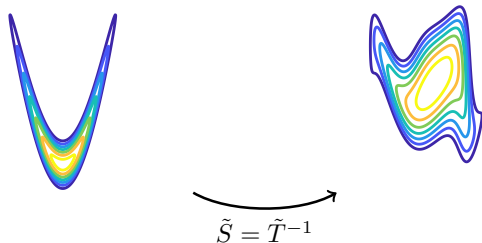
- ▶ Adapt map during MCMC sampling
- ▶ Optimization problem becomes convex and cheaper to solve
- ▶ Asymptotically, samples from posterior
- ▶ BUT: Posterior samples required for initialization (adaptive MCMC)

MFTMap: Two ways to use transport maps

Offline/online approach [Moselhy & Marzouk, 2012]

- ▶ Find map from prior to posterior
- ▶ One time high costs to construct map (nonconvex optimization)
- ▶ No samples from posterior necessary to construct map
- ▶ BUT: Only *approximate* posterior samples

Transport maps to accelerate MCMC [Parno & Marzouk, 2014, 2018]



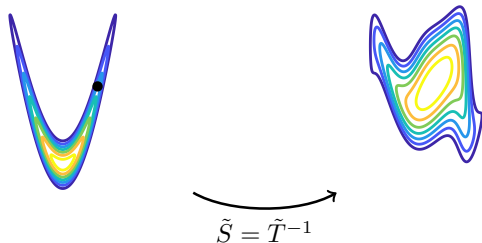
- ▶ Adapt map during MCMC sampling
- ▶ Optimization problem becomes convex and cheaper to solve
- ▶ Asymptotically, samples from posterior
- ▶ BUT: Posterior samples required for initialization (adaptive MCMC)

MFTMap: Two ways to use transport maps

Offline/online approach [Moselhy & Marzouk, 2012]

- ▶ Find map from prior to posterior
- ▶ One time high costs to construct map (nonconvex optimization)
- ▶ No samples from posterior necessary to construct map
- ▶ BUT: Only *approximate* posterior samples

Transport maps to accelerate MCMC [Parno & Marzouk, 2014, 2018]



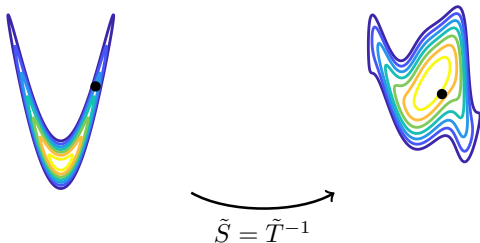
- ▶ Adapt map during MCMC sampling
- ▶ Optimization problem becomes convex and cheaper to solve
- ▶ Asymptotically, samples from posterior
- ▶ BUT: Posterior samples required for initialization (adaptive MCMC)

MFTMap: Two ways to use transport maps

Offline/online approach [Moselhy & Marzouk, 2012]

- ▶ Find map from prior to posterior
- ▶ One time high costs to construct map (nonconvex optimization)
- ▶ No samples from posterior necessary to construct map
- ▶ BUT: Only *approximate* posterior samples

Transport maps to accelerate MCMC [Parno & Marzouk, 2014, 2018]



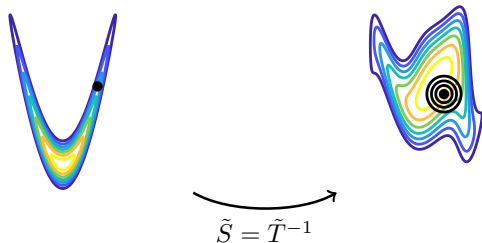
- ▶ Adapt map during MCMC sampling
- ▶ Optimization problem becomes convex and cheaper to solve
- ▶ Asymptotically, samples from posterior
- ▶ BUT: Posterior samples required for initialization (adaptive MCMC)

MFTMap: Two ways to use transport maps

Offline/online approach [Moselhy & Marzouk, 2012]

- ▶ Find map from prior to posterior
- ▶ One time high costs to construct map (nonconvex optimization)
- ▶ No samples from posterior necessary to construct map
- ▶ BUT: Only *approximate* posterior samples

Transport maps to accelerate MCMC [Parno & Marzouk, 2014, 2018]



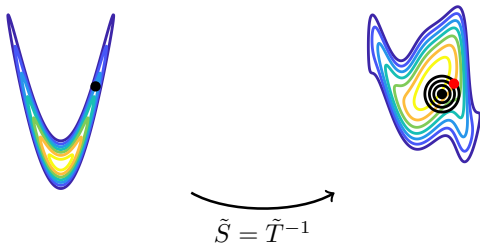
- ▶ Adapt map during MCMC sampling
- ▶ Optimization problem becomes convex and cheaper to solve
- ▶ Asymptotically, samples from posterior
- ▶ BUT: Posterior samples required for initialization (adaptive MCMC)

MFTMap: Two ways to use transport maps

Offline/online approach [Moselhy & Marzouk, 2012]

- ▶ Find map from prior to posterior
- ▶ One time high costs to construct map (nonconvex optimization)
- ▶ No samples from posterior necessary to construct map
- ▶ BUT: Only *approximate* posterior samples

Transport maps to accelerate MCMC [Parno & Marzouk, 2014, 2018]



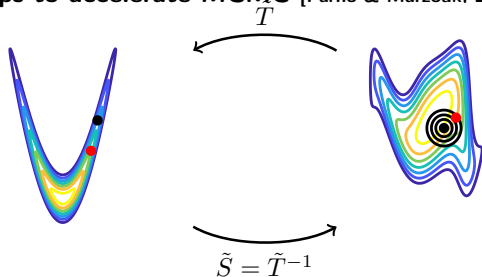
- ▶ Adapt map during MCMC sampling
- ▶ Optimization problem becomes convex and cheaper to solve
- ▶ Asymptotically, samples from posterior
- ▶ BUT: Posterior samples required for initialization (adaptive MCMC)

MFTMap: Two ways to use transport maps

Offline/online approach [Moselhy & Marzouk, 2012]

- ▶ Find map from prior to posterior
- ▶ One time high costs to construct map (nonconvex optimization)
- ▶ No samples from posterior necessary to construct map
- ▶ BUT: Only *approximate* posterior samples

Transport maps to accelerate MCMC [Parno & Marzouk, 2014, 2018]

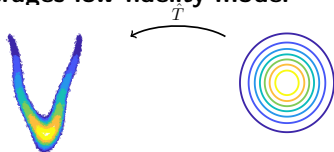


- ▶ Adapt map during MCMC sampling
- ▶ Optimization problem becomes convex and cheaper to solve
- ▶ Asymptotically, samples from posterior
- ▶ BUT: Posterior samples required for initialization (adaptive MCMC)

MFTMap: A tmap-based multifidelity preconditioner

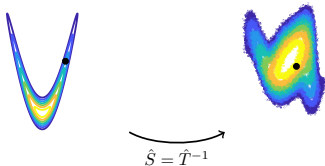
Offline - MFMH leverages low-fidelity model

[P. & Marzouk, *A transport-based multifidelity preconditioner for Markov chain Monte Carlo*. arXiv:1808.09379, 2018.]



- ▶ Build on offline/online decomposition of [Moselhy & Marzouk, 2012]
- ▶ Mitigate high offline costs by leveraging low-fidelity model
- ▶ No samples of high-fidelity posterior required to construct map

Online - MFMH evaluates high-fidelity model to correct

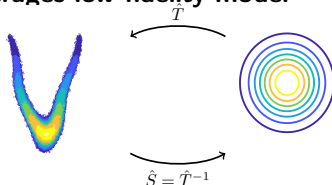


- ▶ Use (inverse) tmap as preconditioner [Parno & Marzouk, 2014, 2018]
- ▶ MCMC corrects error introduced by low-fidelity model, etc
- ▶ Scheme is ergodic for high-fidelity posterior

MFTMap: A tmap-based multifidelity preconditioner

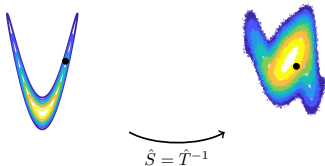
Offline - MFMH leverages low-fidelity model

[P. & Marzouk, *A transport-based multifidelity preconditioner for Markov chain Monte Carlo*. arXiv:1808.09379, 2018.]



- ▶ Build on offline/online decomposition of [Moselhy & Marzouk, 2012]
- ▶ Mitigate high offline costs by leveraging low-fidelity model
- ▶ No samples of high-fidelity posterior required to construct map

Online - MFMH evaluates high-fidelity model to correct

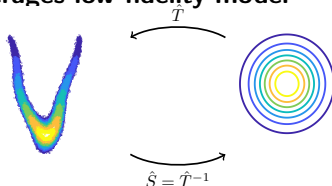


- ▶ Use (inverse) tmap as preconditioner [Parno & Marzouk, 2014, 2018]
- ▶ MCMC corrects error introduced by low-fidelity model, etc
- ▶ Scheme is ergodic for high-fidelity posterior

MFTMap: A tmap-based multifidelity preconditioner

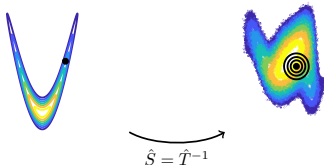
Offline - MFMH leverages low-fidelity model

[P. & Marzouk, *A transport-based multifidelity preconditioner for Markov chain Monte Carlo*. arXiv:1808.09379, 2018.]



- ▶ Build on offline/online decomposition of [Moselhy & Marzouk, 2012]
- ▶ Mitigate high offline costs by leveraging low-fidelity model
- ▶ No samples of high-fidelity posterior required to construct map

Online - MFMH evaluates high-fidelity model to correct

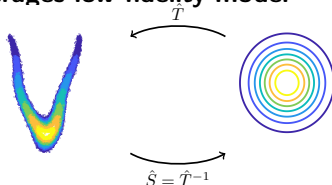


- ▶ Use (inverse) tmap as preconditioner [Parno & Marzouk, 2014, 2018]
- ▶ MCMC corrects error introduced by low-fidelity model, etc
- ▶ Scheme is ergodic for high-fidelity posterior

MFTMap: A tmap-based multifidelity preconditioner

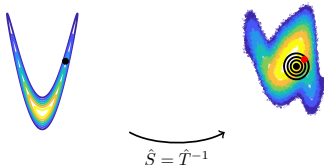
Offline - MFMH leverages low-fidelity model

[P. & Marzouk, *A transport-based multifidelity preconditioner for Markov chain Monte Carlo*. arXiv:1808.09379, 2018.]



- ▶ Build on offline/online decomposition of [Moselhy & Marzouk, 2012]
- ▶ Mitigate high offline costs by leveraging low-fidelity model
- ▶ No samples of high-fidelity posterior required to construct map

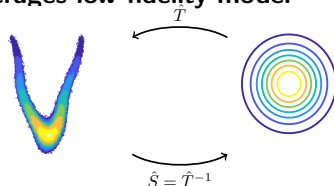
Online - MFMH evaluates high-fidelity model to correct



- ▶ Use (inverse) tmap as preconditioner [Parno & Marzouk, 2014, 2018]
- ▶ MCMC corrects error introduced by low-fidelity model, etc
- ▶ Scheme is ergodic for high-fidelity posterior

MFTMap: A tmap-based multifidelity preconditioner

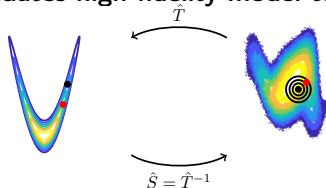
Offline - MFMH leverages low-fidelity model



[P. & Marzouk, *A transport-based multifidelity preconditioner for Markov chain Monte Carlo*. arXiv:1808.09379, 2018.]

- ▶ Build on offline/online decomposition of [Moselhy & Marzouk, 2012]
- ▶ Mitigate high offline costs by leveraging low-fidelity model
- ▶ No samples of high-fidelity posterior required to construct map

Online - MFMH evaluates high-fidelity model to correct



- ▶ Use (inverse) tmap as preconditioner [Parno & Marzouk, 2014, 2018]
- ▶ MCMC corrects error introduced by low-fidelity model, etc
- ▶ Scheme is ergodic for high-fidelity posterior

MFTMap: Euler-Bernoulli beam¹

Euler-Bernoulli beam

$$\frac{\partial^2}{\partial x^2} \left(E(x) \frac{\partial^2}{\partial x^2} u(x) \right) = g(x), \quad x \in (0, 1)$$

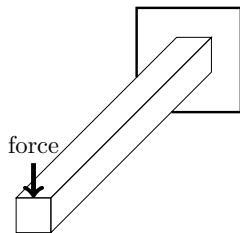
- ▶ Vertical deflection of beam $u : [0, 1] \rightarrow \mathbb{R}$
- ▶ Effective stiffness $E : [0, 1] \rightarrow \mathbb{R}$
- ▶ Load $g : [0, 1] \rightarrow \mathbb{R}$

Models

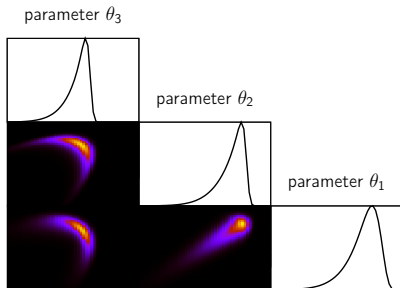
- ▶ FD model $f^{(1)}$ with $N = 601$ DOFs
- ▶ Spline interpolant $f^{(2)}$ of model $f^{(1)}$

Inverse problem

- ▶ Data \mathbf{y} are u evaluated at 41 points
- ▶ Infer piecewise constant E



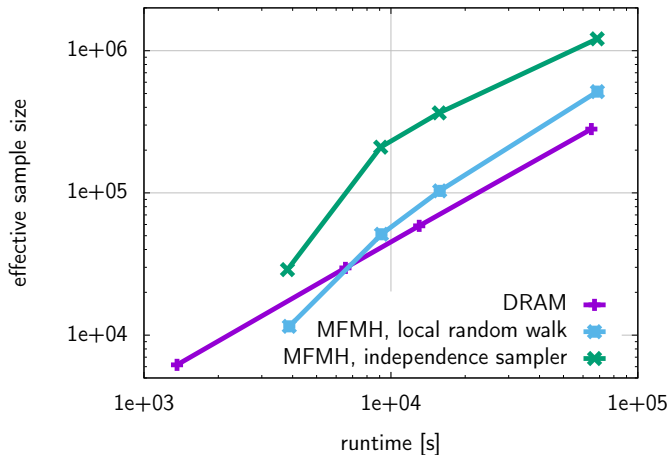
(a) Euler-Bernoulli beam



(b) dimension-wise plot of posterior

¹Adapted from <https://github.com/g2s3-2018/labs>

MFTMap: Euler-Bernoulli beam - numerical results



- ▶ Multifidelity (MFMH) leads to higher ESS than DRAM
- ▶ Offline costs need to be compensated before speedups are obtained
- ▶ Independence sampler (reference) can be used

1. Motivation for multifidelity methods
2. Multifidelity Monte Carlo estimation (MFMC)
3. MFMC with context-aware surrogates
4. Multifidelity and Bayesian inverse problems
5. **Outlook and conclusions**

Outlook

Optimization under uncertainty

- ▶ Estimate statistics in optimization iteration
- ▶ Robust optimization

Rare event simulation

- ▶ Estimate probability of rare event
- ▶ Crucial for risk-averse optimization

Sensitivity analysis

- ▶ Identify parameters of model that lead to largest variance in quantity of interest
- ▶ Large-scale variance estimation problem

Bayesian inverse problems

- ▶ Markov chain Monte Carlo sampling
- ▶ Increase acceptance probability of moves

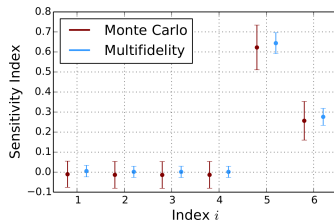
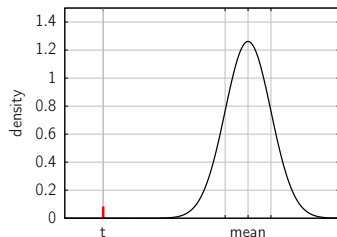
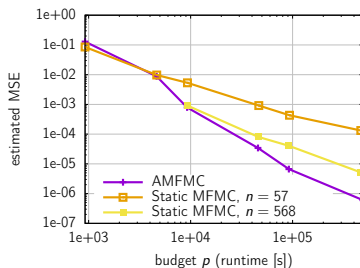
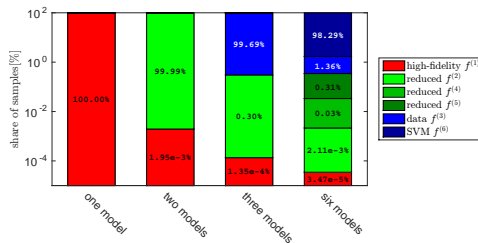


Figure: Elizabeth Qian

[P., Willcox, Gunzburger, *Survey of multifidelity methods in uncertainty propagation, inference, and optimization*; SIAM Review, 60(3):550-591, 2018]

Conclusions



Multifidelity methods

- ▶ Leverage surrogate models for runtime speedup
- ▶ Recourse to high-fidelity model for accuracy guarantees
- ▶ Optimally trade off approximation, sampling, and construction
- ▶ Context aware construction of surrogate models

Our references

- 1 P., Willcox & Gunzburger *Optimal model management for multifidelity Monte Carlo estimation*. SISC, 2016.
- 2 P., Gunzburger & Willcox: *Convergence analysis of multifidelity Monte Carlo estimation*. Numerische Mathematik, 2018
- 3 P. *Multifidelity Monte Carlo estimation with adaptive low-fidelity models*. submitted, 2017.
- 4 P. & Marzouk: *A transport-based multifidelity preconditioner for Markov chain Monte Carlo*. arXiv:1808.09379, 2018.