Data-Driven Multifidelity Methods for Monte Carlo Estimation

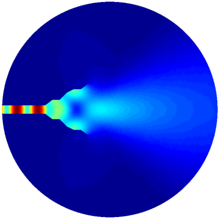
Benjamin Peherstorfer
University of Wisconsin-Madison

Karen Willcox
Massachusetts Institute of Technology

Max Gunzburger
Florida State University

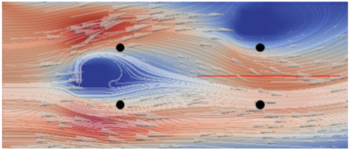
April 2018

Outer loop applications



optimization

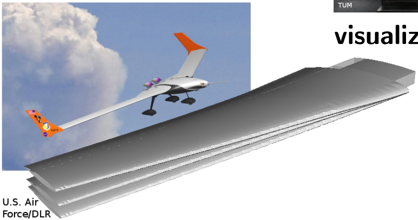
control



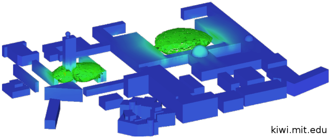
inference

visualization

model calibration



multi-discipline coupling



uncertainty quantification

Surrogate models

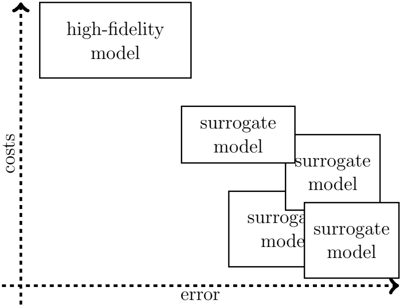
Given is a high-fidelity model $f^{(1)} : \mathcal{D} \to \mathcal{Y}$

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

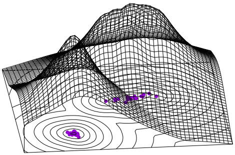
Additionally, often have surrogate models

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

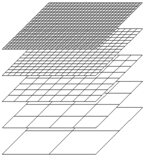
- Approximate high-fidelity $f^{(1)}$
- Often orders of magnitudes cheaper



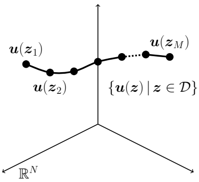
Examples of surrogate models



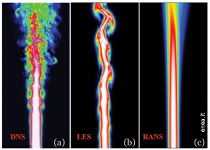
data-fit models, response surfaces, machine learning



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

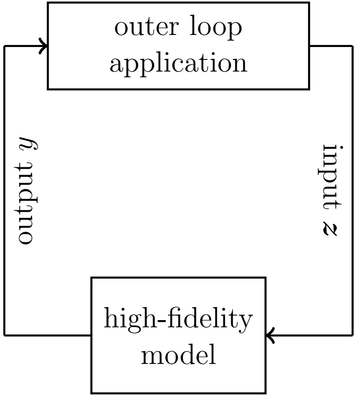
- Costs of outer loop reduced
- Often orders of magnitude speedups

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

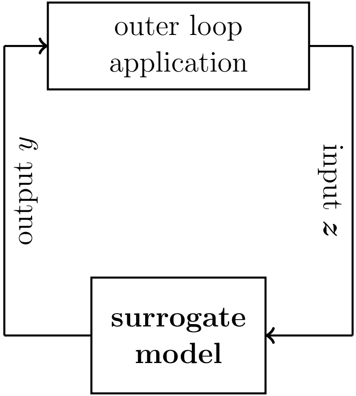
- Costs of outer loop reduced
- Often orders of magnitude speedups

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

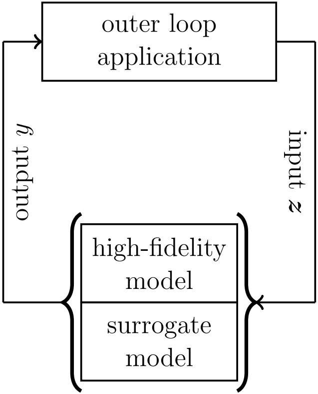
- Leverage surrogate models for speedup
- Recourse to high-fidelity for accuracy

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, 2018 (to appear)]

Our approach: Multifidelity methods

Combine high-fidelity and surrogate models

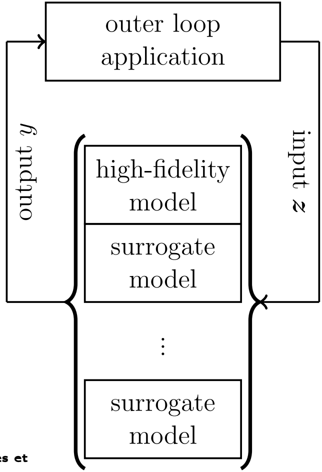
- Leverage surrogate models for speedup
- Recourse to high-fidelity for accuracy

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, 2018 (to appear)]

Outline

1. Motivation for multifidelity methods
2. Multifidelity Monte Carlo estimation (MFMC)
3. Asymptotic analysis of MFMC
4. Adaptive surrogates and MFMC
5. Outlook and conclusions

Outline

1. Motivation for multifidelity methods

2. **Multifidelity Monte Carlo estimation (MFMC)**

3. Asymptotic analysis of MFMC

4. Adaptive surrogates and MFMC

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 $w_1 > 0$

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

Given random variable Z, estimate

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

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

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

Uncertainty propagation with Monte Carlo is outer-loop application

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

[Rozza, Carlberg, Manzoni, Ohlberger, Veroy-Grepl, Willcox, Kramer, Benner, Ullmann, Nouy, Zahm, etc]

MFMC: Control variates

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

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

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

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

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

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

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 $\hat{s}_A$

[Nelson, 87]

MFMC: Control variates and surrogate models

Models

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

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

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

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]

$\Rightarrow$ **Need for arbitrary number of surrogates, any type of surrogates**

MFMC: Multifidelity Monte Carlo estimator

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

Take realizations of input random variable Z

$$\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3, \dots$$

Evaluate model $f^{(i)}$ at first m_i realizations $\boldsymbol{z}_1, \dots, \boldsymbol{z}_{m_i}$ of Z

$$f^{(i)}(\boldsymbol{z}_1), \dots, f^{(i)}(\boldsymbol{z}_{m_i}), \qquad i = 1, \dots, k$$

Multifidelity Monte Carlo (MFMC) estimator

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

- 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

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

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

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

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

MFMC: Optimal sampling

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

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

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

Sketch of proof

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

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

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

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

MFMC: Numerical example

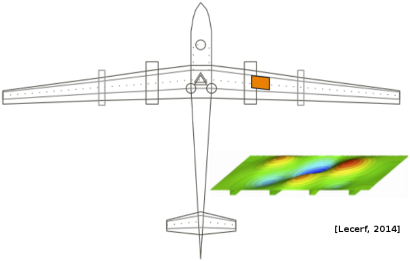
Locally damaged plate in bending

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

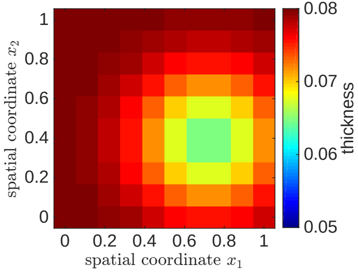
Six models

- 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



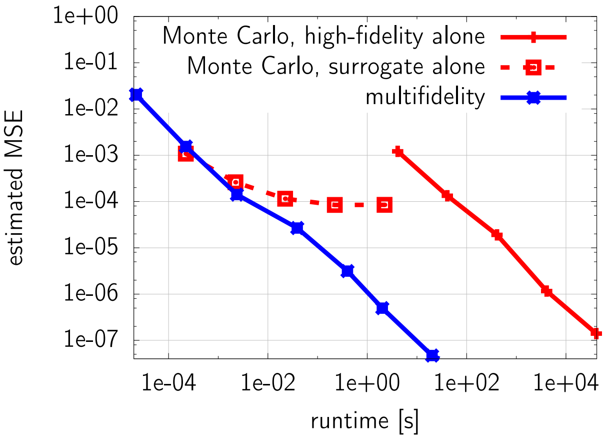
(a) wing panel



(b) damaged plate

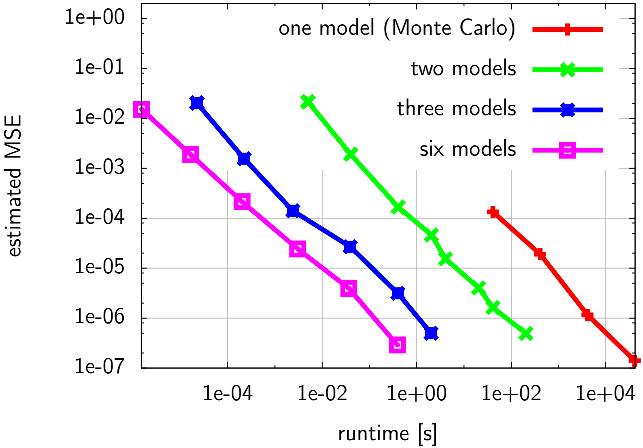
MFMC: Speedups in uncertainty propagation

Computed on MAC cluster
10 nodes à 64 cores
total of 640 cores



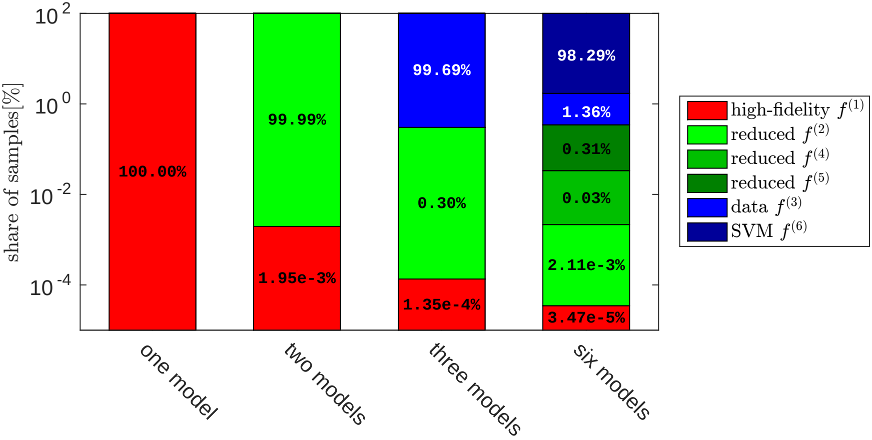
- Monte Carlo needs 12h runtime for estimate with error below 10^{-7}
- Multifidelity provides estimator with error below 10^{-7} after 9 seconds

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

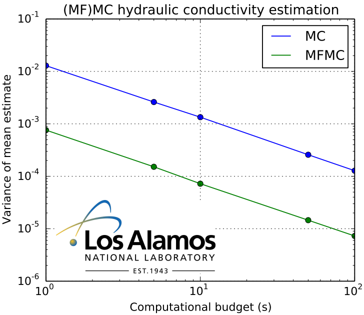


Figure: Elizabeth Qian

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

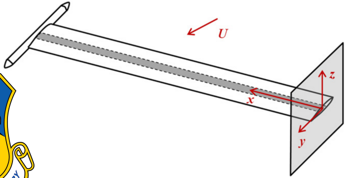


Figure: Philip S. Beran

Outline

1. Motivation for multifidelity methods

2. Multifidelity Monte Carlo estimation (MFMC)

3. Asymptotic analysis of MFMC

4. Adaptive surrogates and MFMC

5. Outlook and conclusions

Outline

1. Motivation for multifidelity methods
2. Multifidelity Monte Carlo estimation (MFMC)
3. **Asymptotic analysis of MFMC**
4. Adaptive surrogates and MFMC
5. Outlook and conclusions

1 P., Gunzburger & Willcox *Convergence analysis of multifidelity Monte Carlo estimation*. Numerische Mathematik, 2018.

MFMC: Asymptotic analysis

Properties of MFMC in setting with $f^{(1)}, f^{(2)}, \dots, f^{(k)}$

- Existence and uniqueness
- Unbiased estimator of statistics of high-fidelity model $f^{(1)}$
- MSE in terms of costs and correlation coefficients

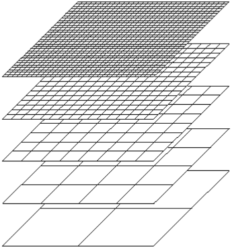
Now ("exact") f **and sequence** $f^{(1)}, f^{(2)}, \dots$

- Estimate

$$\mathbb{E}[f(Z)]$$

- MSE of MFMC estimator $\hat{s}$ that uses $f^{(1)}, \dots, f^{(L)}$ is

$$e(\hat{s}) = \underbrace{\mathrm{Var}[\hat{s}]}_{\text{variance}} + \underbrace{\mathbb{E}[f(Z) - f^{(L)}(Z)]^2}_{\text{bias w.r.t. } f}$$



Example: $f^{(1)}, f^{(2)}, \dots$ correspond to multilevel discretization of f

Goal: Given tolerance $\epsilon > 0$

- Find $L \in \mathbb{N}$, #model evaluations $\boldsymbol{m}$, coefficients γ such that $e(\hat{s}) \lesssim \epsilon$
- Bound costs $c(\hat{s})$

[Brandt, 1977], [Goodman et al, 1989], [Heinrich, 2001], [Giles, 2008], [Cliffe, 2011]

MFMC: Asymptotic results

Assumption: There exists $1 < h \in \mathbb{R}$ and rates $0 < \alpha, \beta, \tau \in \mathbb{R}$ such that

- $|\mathbb{E}[f - f^{(\ell)}]| \lesssim h^{-\alpha\ell}, \quad \ell \in \mathbb{N}$
- $w_\ell \lesssim h^{\beta\ell}, \quad \ell \in \mathbb{N}$
- $\text{Var}\left[f^{(\ell)} - f^{(\ell-1)}\right] \lesssim h^{-\tau\ell}, \quad \ell \in \mathbb{N}$

(Regular) Monte Carlo estimator $\bar{y}_q^{(L)}$ achieves $e(\bar{y}_q^{(L)}) \lesssim \epsilon$ with

$$c(\bar{y}_q^{(L)}) \lesssim \epsilon^{-1}\epsilon^{-\beta/(2\alpha)}$$

Theorem 2 (P., Gunzburger, Willcox, 2018).

If MFMC estimator exists and $\tau > \beta$, then MFMC achieves $e(\hat{s}) \lesssim \epsilon$ with

$$c(\hat{s}) \lesssim \epsilon^{-1}$$

- Costs bound independent of rates α and β
- Agrees with results in multilevel Monte Carlo estimation [Giles, 2008]

[P., Gunzburger & Willcox: ***Convergence analysis of multifidelity Monte Carlo estimation***. Numerische Mathematik, 2018]

Outline

1. Motivation for multifidelity methods
2. Multifidelity Monte Carlo estimation (MFMC)
3. Asymptotic analysis of MFMC
4. Adaptive surrogates and MFMC
5. Outlook and conclusions

Outline

1. Motivation for multifidelity methods

2. Multifidelity Monte Carlo estimation (MFMC)

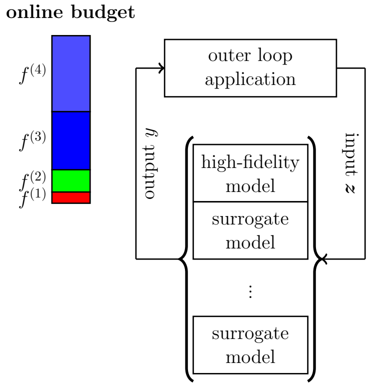
3. Asymptotic analysis of MFMC

4. **Adaptive surrogates and MFMC**

5. Outlook and conclusions

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

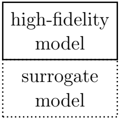
AMFMC: Integrate model reduction into MFMC



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]**

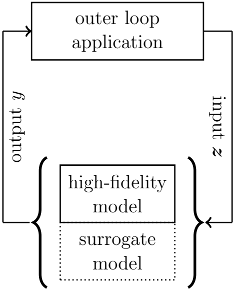
AMFMC: Integrate model reduction into MFMC



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: Integrate model reduction into MFMC



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

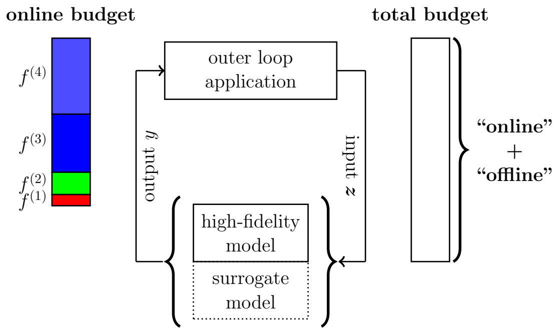
- **Goal is outer-loop result**
- **Other models set the context in which surrogate model will be used**

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]**

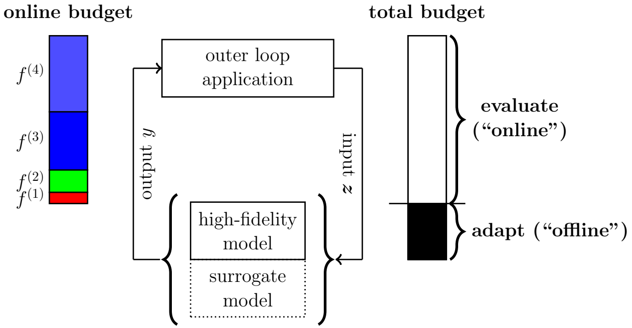
AMFMC: Integrate model reduction into MFMC



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]**

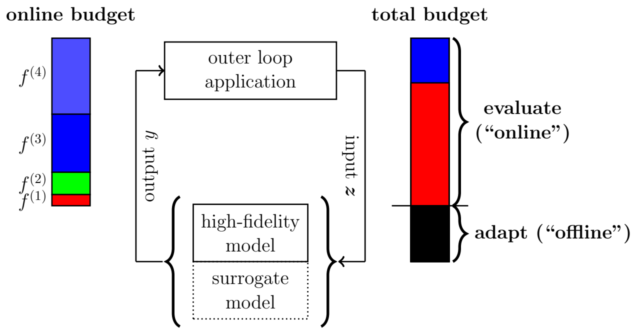
AMFMC: Integrate model reduction into MFMC



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]**

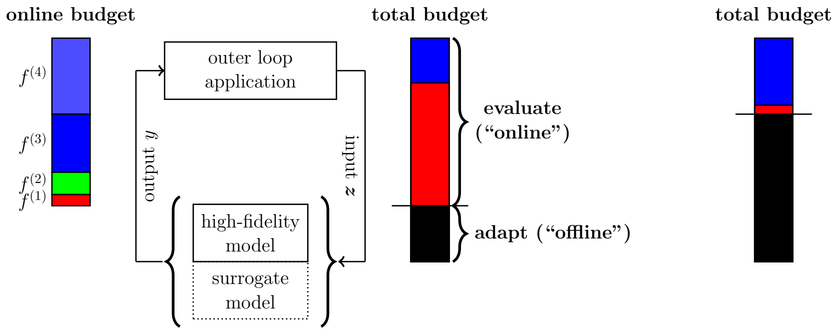
AMFMC: Integrate model reduction into MFMC



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]**

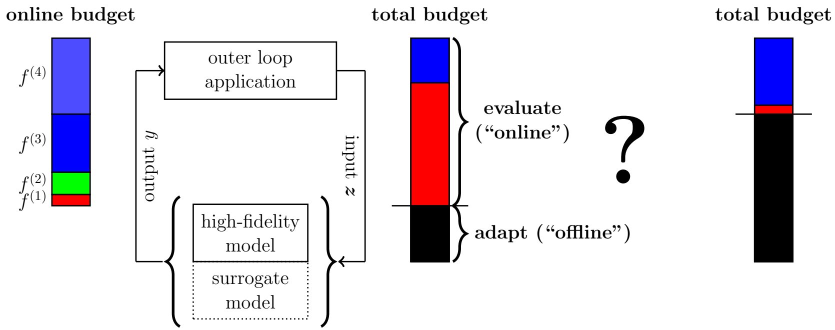
AMFMC: Integrate model reduction into MFMC



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 $w_0 = 1$

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

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

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

Surrogate model approximates high-fidelity model in the sense

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

Evaluation costs of surrogate model may grow with n as

$$w_n \leq c_2 n^{\beta}\,, \qquad 0 < c_2, \beta$$

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 $e(\hat{s}_n)$

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

AMFMC: Existence, uniqueness, and convexity

Consider the objective function

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

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

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

Theorem 3 (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 $p - \hat{n}^*$

[P., Multifidelity Monte Carlo estimation with adaptive low-fidelity models, 2017 (submitted)]

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 $p \to \infty$
- Surrogate models can be "too accurate" for multifidelity methods

Corollary 4 (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

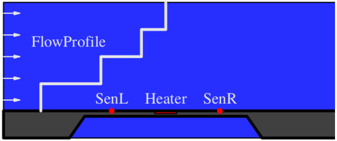
  $\Rightarrow$ control functionals **[Oates, Girolami, Chopin, 2016]**
- Helps to understand case $w_n \ll 1$ ($f^{(\hat{n}^*)}$ much cheaper than f)

Anemometer problem

- Measure velocity of fluid
- Three inputs uniformly distributed in

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

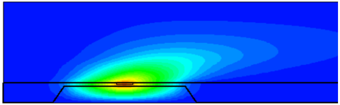
- Output is velocity
- Estimate expected velocity



Figures: MORWiki

High-fidelity model

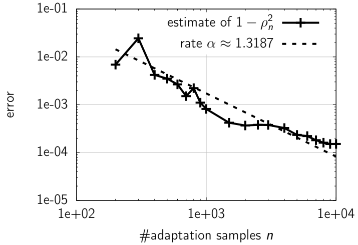
- Based on convection-diffusion equation
- Discretized with finite elements
- High-fidelity model has 29008 DoFs



https://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Anemometer

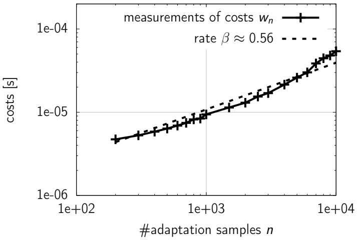
Surrogate model

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



Optimizing for $\hat{n}^*$

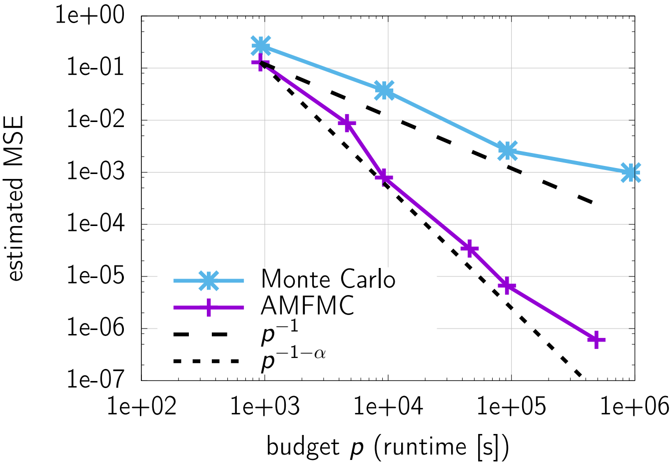
- One dimensional convex problem
- Numerically solve for $\hat{n}^*$



Adaptation of surrogate in AMFMC

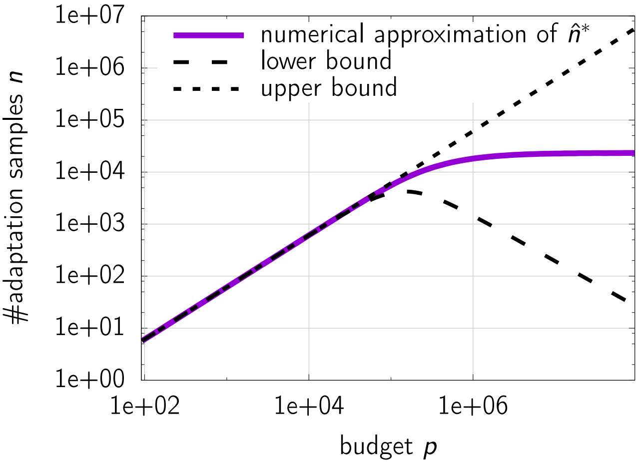
- Numerically estimate rates from pilot runs
- Optimize for $\hat{n}^*$ with Matlab's `fmincon`

AMFMC: Anemometer results



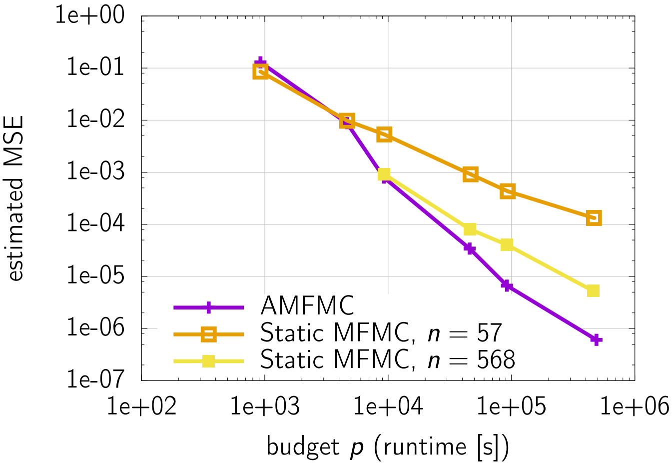
- Speedups of up to 3 orders of magnitude compared crude Monte Carlo
- MSE of AMFMC decays with $p^{-1-\alpha}$ in *pre-asymptotic* regime

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

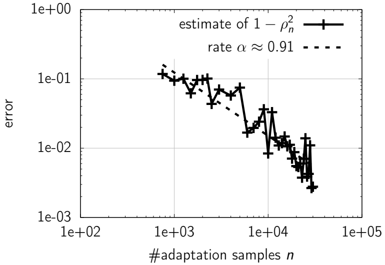
- Length and height uniformly distributed

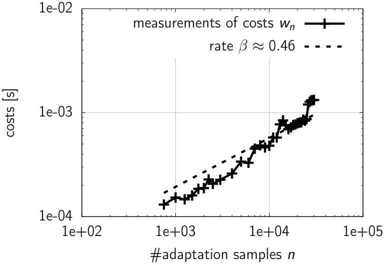
$$[0.8, 1.2] \times [5 \times 10^{-4}, 5 \times 10^{-3}]$$

- Output is displacement of beam
- Estimate expected displacement

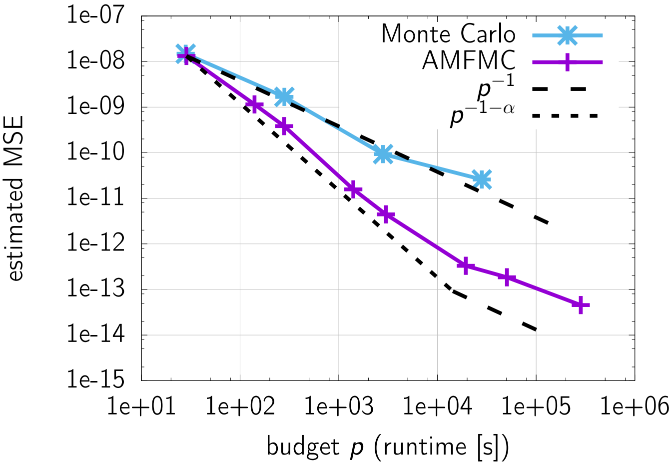
Models

- 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}

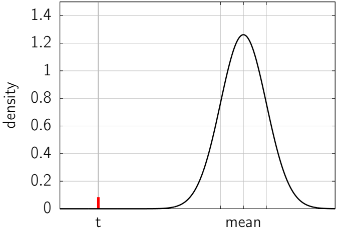
Outline

1. Motivation for multifidelity methods
2. Multifidelity Monte Carlo estimation (MFMC)
3. Asymptotic analysis of MFMC
4. Adaptive surrogates and MFMC
5. **Outlook and conclusions**

Outlook

Optimization under uncertainty

- Estimate statistics in optimization iteration
- Robust optimization

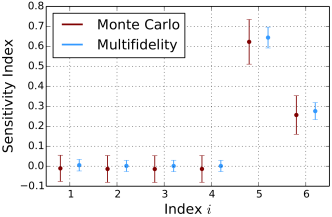


Rare event simulation

- 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



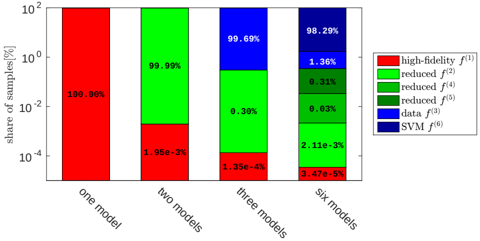
Bayesian inverse problems

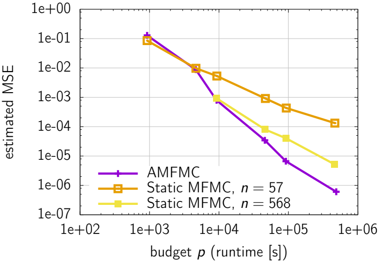
- Markov chain Monte Carlo sampling
- Increase acceptance probability of moves

Figure: Elizabeth Qian

[P., Willcox, Gunzburger, *Survey of multifidelity methods in uncertainty propagation, inference, and optimization*; SIAM Review, 2018 (to appear)]

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

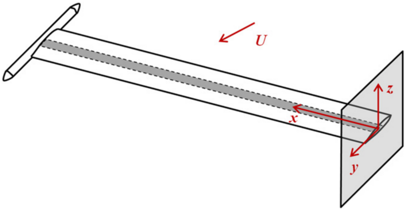
Our references

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.

MFMC: Wing flutter problem setup

Flutter problem

- Uncertain inputs
  - Angle of attack from 0.5° to 2.5°
  - Air density, mass of tip vary by 5%
- Estimate expected flutter speed



High-fidelity model

- Based on Hodges-Dowell equations
- Nonlinear terms of $\geq$3rd order ignored
- FEM discretization with 10 elements

Low-fidelity models

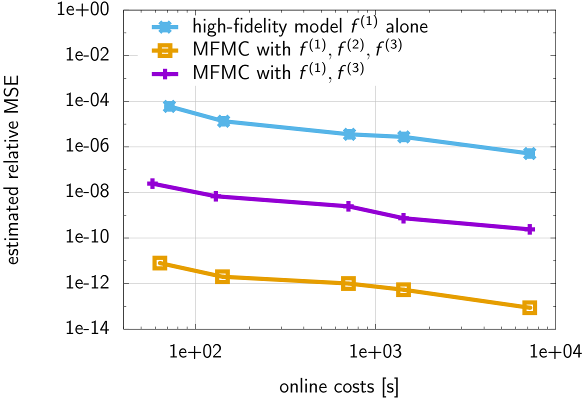
- Spline interpolants on equidistant grid
- Low-fidelity model $f^{(2)}$ from 343 points
- Low-fidelity model $f^{(3)}$ from 125 points

collaboration with Philip Beran (Air Force Research Laboratory)

[Stanford and Beran, 2013], [Beran, Stanford, and Wang, 2017]

[P., Beran, Willcox, 2018]

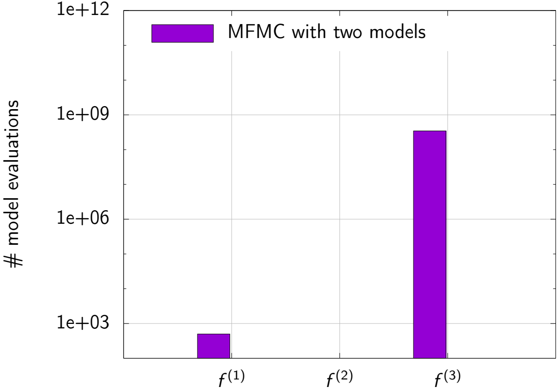
MFMC: Wing flutter speedup results



MFMC achieves significant speedup

- Low-fidelity models are 8 orders of magnitude cheaper than $f^{(1)}$
- MFMC achieves about 7 orders of magnitude speedup

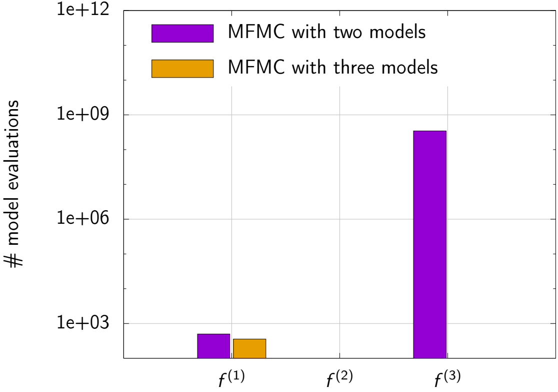
MFMC: Distribution of work in wing flutter problem



Adding model $f^{(2)}$ changes #evals of model $f^{(3)}$

- With $f^{(2)}$, model $f^{(3)}$ is evaluated more often
- Demonstrates that interactions between models drives efficiency of MFMC

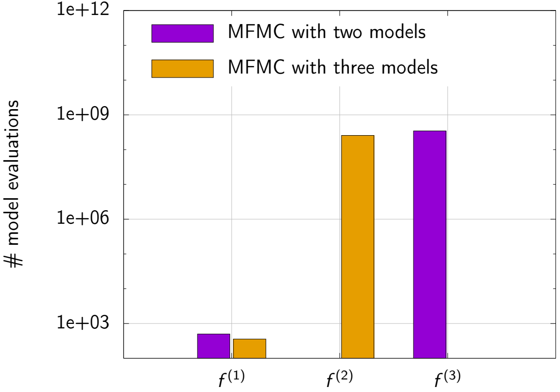
MFMC: Distribution of work in wing flutter problem



Adding model $f^{(2)}$ changes #evals of model $f^{(3)}$

- With $f^{(2)}$, model $f^{(3)}$ is evaluated more often
- Demonstrates that interactions between models drives efficiency of MFMC

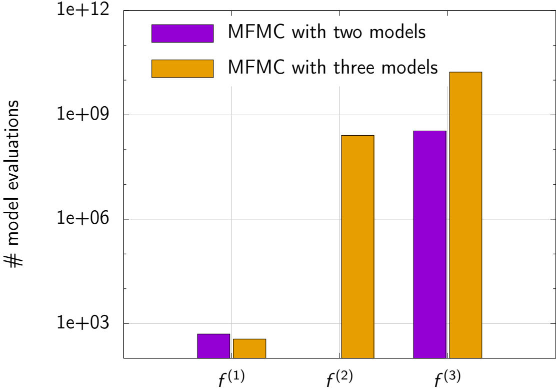
MFMC: Distribution of work in wing flutter problem



Adding model $f^{(2)}$ changes #evals of model $f^{(3)}$

- With $f^{(2)}$, model $f^{(3)}$ is evaluated more often
- Demonstrates that interactions between models drives efficiency of MFMC

MFMC: Distribution of work in wing flutter problem



Adding model $f^{(2)}$ changes #evals of model $f^{(3)}$

- With $f^{(2)}$, model $f^{(3)}$ is evaluated more often
- Demonstrates that interactions between models drives efficiency of MFMC