Multifidelity Monte Carlo Methods for Uncertainty Quantification

Karen E. Willcox

Joint work with Tiangang Cui, Max Gunzburger, Boris Kramer,
Youssef Marzouk, Benjamin Peherstorfer

February 27, 2017

CSE17: SIAM Conference on Computational Science and Engineering

Minisymposium on Multifidelity Simulation and Approximation
in the Computational Sciences

MIT AEROASTRO

Multifidelity models and multifidelity methods

What are they and why use them?

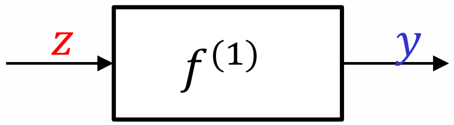
Multifidelity models

covering a range of different resolutions, scales, modeling assumptions, etc.

simplified physics, loosened tolerance, coarse grid, data-fit, projection-based ROM, etc.

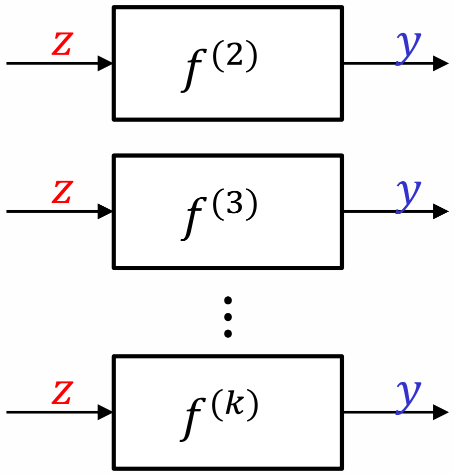
- high-fidelity model (“truth”) mapping input z to output y

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



- *k – 1* lower-fidelity models mapping input z to output y

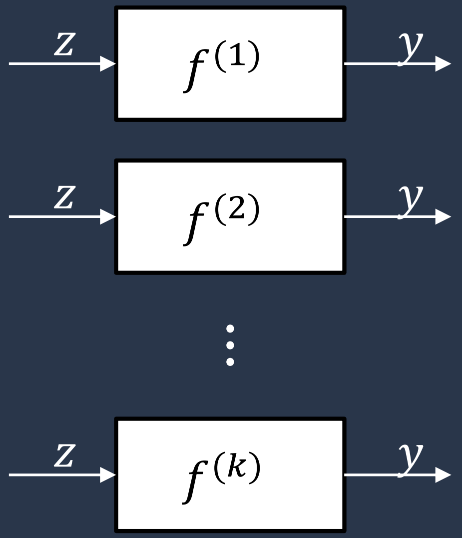
$$f^{(2)}, \ldots, f^{(k)}: \mathcal{Z} \to \mathcal{Y}$$



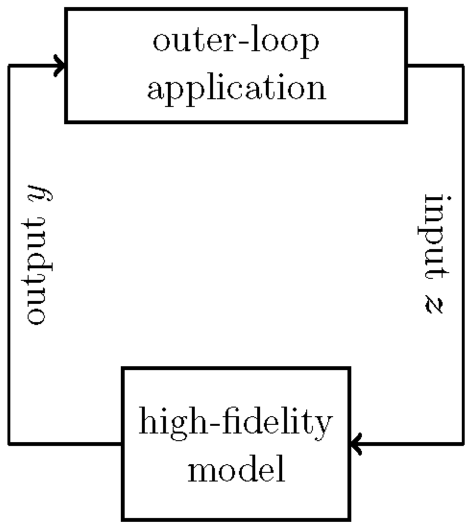
- model $f^{(i)}$ has cost w_i
- model $f^{(i)}$ has fidelity f_i
- models do not necessarily form a hierarchy

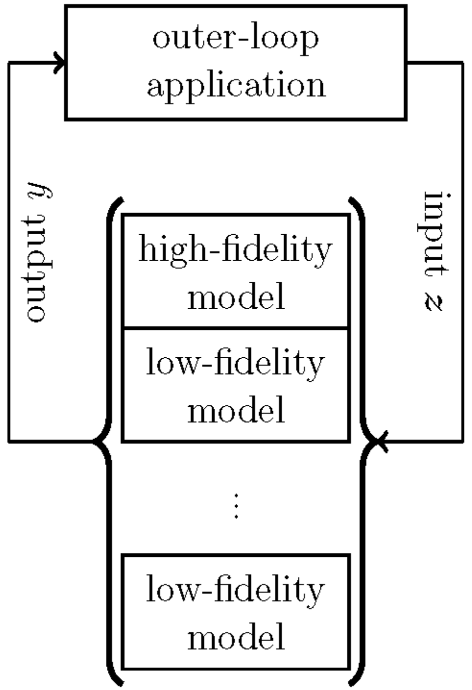
Multifidelity methods

for outer-loop problems



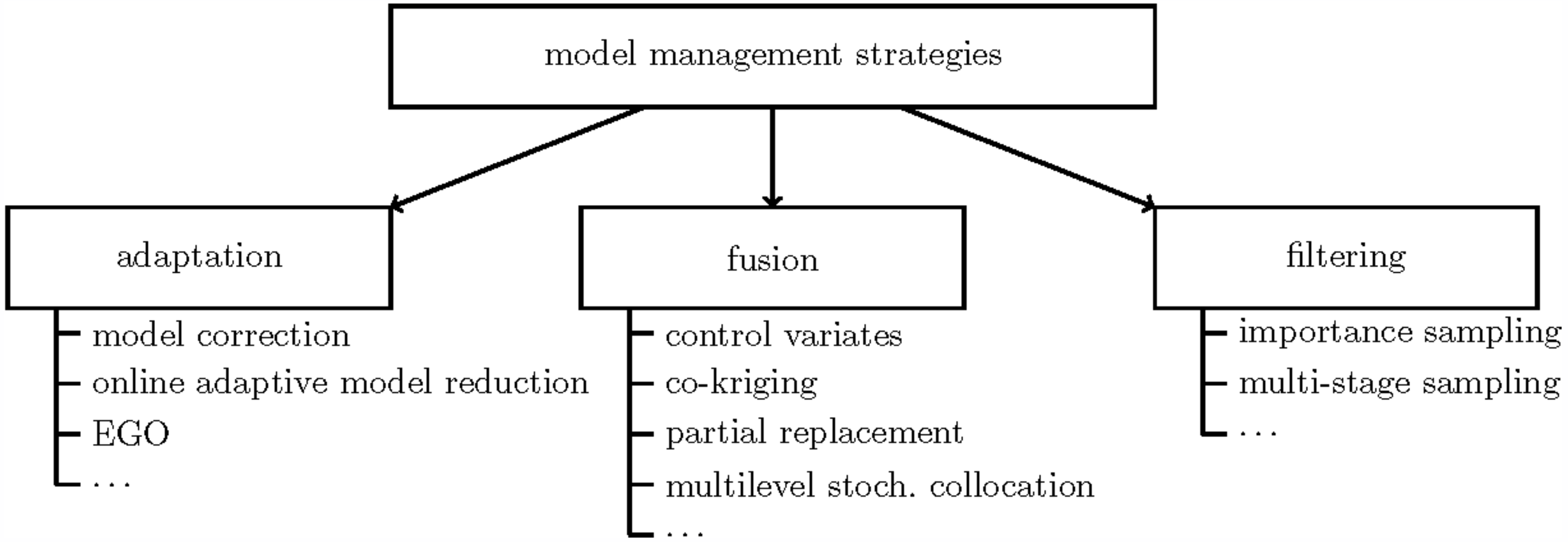
- **Outer-loop**: computational applications that form outer loops around a model
  - overall outer-loop result is obtained at the termination of the outer loop
  - examples: optimization, uncertainty propagation, inverse problems, data assimilation, control, sensitivity analysis
- **Multifidelity methods**: goal is to solve the outer-loop problem at high fidelity
  - invoke multiple models to reduce computational cost
  - maintains guarantees on outer-loop result
- Key questions
  - how to **combine model estimates**?
  - how to **balance evaluations** among models?
  - how to **guarantee outer-loop result**?



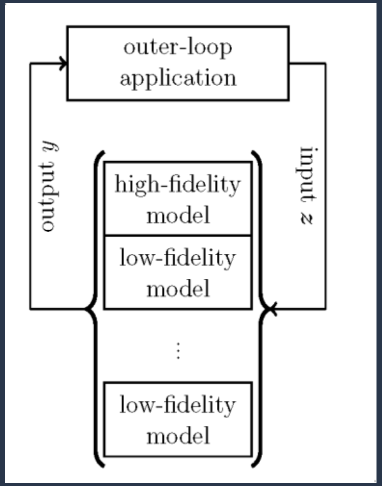


Multifidelity strategies

examples of multifidelity strategies for the outer loop



- optimization
  Alexandrov & Lewis, 1999; Eldred et al., 2004
- forward propagation of uncertainty
  Giles, 2008, Ng & Eldred, 2012, Ng & W., 2012, 2014; Peherstorfer et al., 2016
- failure probability estimation
  Bichon et al, 2008; Li & Xiu, 2010; Peherstorfer et al., 2016, Peherstorfer et al., 2017
- optimization under uncertainty
  Ng, Huynh, W., 2012; Ng & W., 2014, 2016
- statistical inverse problems
  Fox & Christensen, 2008; Efendiev & Hou, 2009; Cui et al., 2014



Why use multifidelity modeling?

Why use a multifidelity formulation?

Reduced model
(approximate)

Full model
(“truth”)

Why use a multifidelity formulation?

Reduced model
(approximate)

Computationally
cheap(er)

Full model
("truth")

Computationally
expensive

Why use a multifidelity formulation?

Reduced model
(approximate)

Full model
("truth")

Certified?

yes

- Replace full model with reduced model and solve {opt, UQ, inverse}
- Propagate error estimates on forward predictions to determine error in {opt, UQ, inverse} solutions (may be non-trivial)

Why use a multifidelity formulation?

Reduced model
(approximate)

Full model
("truth")

Certified?

↓ no

- Replace full model with reduced model and solve {opt, UQ, inverse}
- Hope for the best

Why use a multifidelity formulation?

Reduced model
(approximate)

Full model
(“truth”)

Certified?

no

- Use a multifidelity formulation that invokes both the reduced model and the full model
- Trade computational cost for the ability to place guarantees on the solution of {opt, UQ, inverse}

Why use a multifidelity formulation?

Reduced model
(approximate)

Full model
("truth")

Certified?

no

- Use a multifidelity formulation that invokes both the reduced model and the full model
- Trade computational cost for the ability to place guarantees on the solution of {opt, UQ, inverse}
- Certify the solution of {opt, UQ, inverse} even in the absence of guarantees on the reduced model itself

Multifidelity Monte Carlo (MFMC)

Efficient uncertainty propagation leveraging multiple models

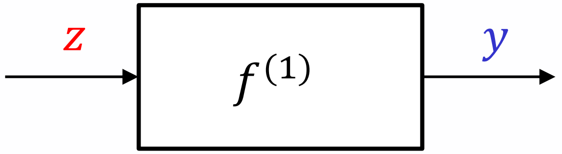
Ng & W., **Multifidelity approaches for optimization under uncertainty**, *IJNME*, 2014

Peherstorfer, W. & Gunzburger, **Optimal model management for multifidelity Monte Carlo estimation**, *SISC*, 2016

Estimating QoI statistics

via Monte Carlo sampling

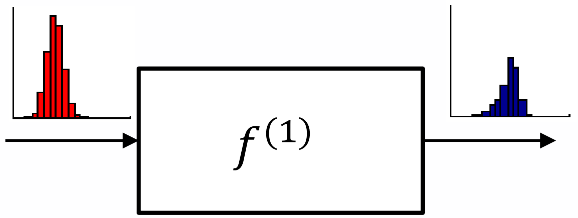
- uncertain **input** $z \in \mathcal{Z}$
- **output** quantity of interest $y \in \mathcal{Y}$
- high-fidelity **model**
  $f^{(1)}: \mathcal{Z} \to \mathcal{Y}$
  with cost $w_1 > 0$ (“truth”)



- Goal: given random input variable *Z*, estimate statistics *s* of $f^{(1)}(Z)$

- Example: expected value

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



- Monte Carlo estimator for s using n realizations $z_1, \dots, z_n$ of *Z* has costs nw_1:

$$\hat{s} = \bar{y}_n^{(1)} = \frac{1}{n}\sum_{i=1}^{n} f^{(1)}(z_i)$$

Multifidelity Monte Carlo

leveraging multiple approximate models to estimate statistics of the high-fidelity model

- high-fidelity model

  $f^{(1)}: \mathcal{Z} \to \mathcal{Y}$ (“truth”)

- *k – 1* surrogate models

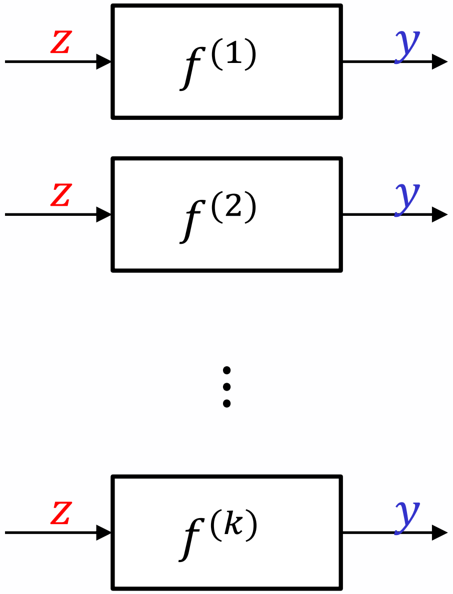
$$f^{(2)}, \ldots, f^{(k)}: \mathcal{Z} \to \mathcal{Y}$$

- model $f^{(i)}$ has cost w_i

- m_i evaluations for model i, with

$$m_1 \leq m_2 \leq \ldots \leq m_k$$

- Models do not necessarily form a hierarchy (cf. multi-level Monte Carlo)
  - How to combine models?
  - How to balance evaluations among them?

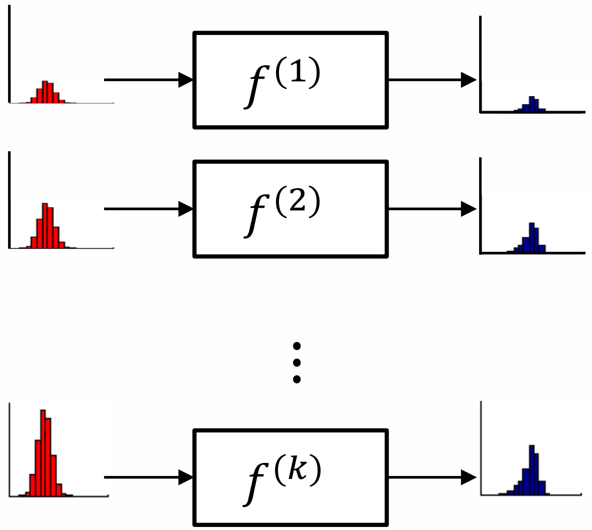


Multifidelity Monte Carlo

leveraging multiple approximate models to estimate statistics of the high-fidelity model

- Draw m_k realizations $z_1, \dots, z_{m_k}$ of Z and evaluate $f^{(i)}$:

  $f^{(i)}(z_1), \dots, f^{(i)}(z_{m_i})$



- Compute mean estimators

  $\bar{y}^{(1)}_{m_1}, \dots, \bar{y}^{(k)}_{m_k}$ and $\bar{y}^{(2)}_{m_1}, \dots, \bar{y}^{(k)}_{m_{k-1}}$

- MFMC estimator:

$$\hat{s} = \bar{y}^{(1)}_{m_1} + \sum_{i=2}^{k} \alpha_i \left(\bar{y}^{(i)}_{m_i} - \bar{y}^{(i)}_{m_{i-1}} \right)$$

MFMC estimate for the mean

mean estimate using m_1 evaluations of truth model

mean estimate using m_i evaluations of model i

mean estimate using m_{i-1} evaluations of model i

- MFMC estimator is unbiased, even with no error bounds for surrogates: $E[\hat{s}] = s$

Multifidelity Monte Carlo

General case with *k* models

Peherstorfer, W., Gunzburger, *SISC*, 2015

- MFMC estimator

$$\hat{s} = \bar{y}_{m_1}^{(1)} + \sum_{i=2}^{k} \alpha_i \left(\bar{y}_{m_i}^{(i)} - \bar{y}_{m_{i-1}}^{(i)} \right)$$

MFMC estimate for the mean

mean estimate using m_1 evaluations of truth model

mean estimate using m_i evaluations of model i

mean estimate using m_{i-1} evaluations of model i

- The costs of the MFMC estimator are

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

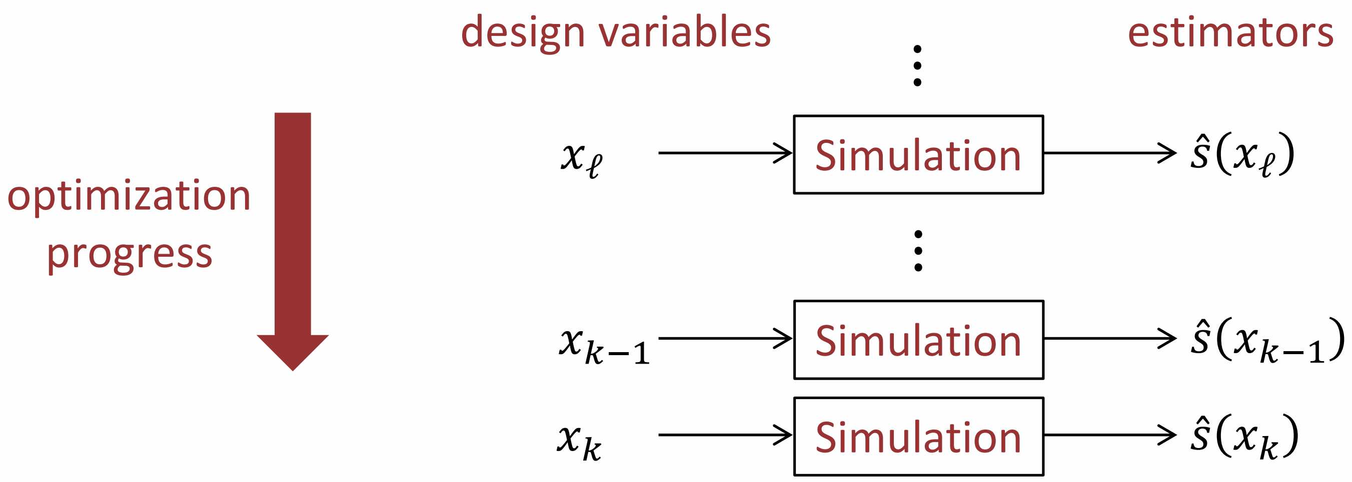
- Distinguishing features of MFMC method:
  - optimal selection of the number of model evaluations $m_1 \leq m_2 \leq \ldots \leq m_k$ and of coefficients $\alpha_2, \ldots, \alpha_k$
  - applicable to general information sources (e.g., any type of surrogate model, database curve fits, etc.)

A broad view of multifidelity models

in many outer-loop applications, can exploit past evaluations as a low-fidelity model

Ng & W., *J. Aircraft*, 2015

- in optimization under uncertainty, can exploit model correlation over design space
  - use $f^{(1)}(x+\Delta x, z)$ as **surrogate** for $f^{(1)}(x, z)$



- at current design point x_k
  - Define $A = f^{(1)}(x_k, z)$
  - Want to compute $\hat{s}$ as estimator of $s = \mathbb{E}[A]$
- previously visited design point x_ℓ where $\ell < k$
  - Define surrogate as $C = f^{(1)}(x_\ell, z)$
  - Reuse available data: $\hat{s}_C$ as estimator of $s_C = \mathbb{E}[C]$ with error $\mathrm{Var}[\hat{s}_C]$

Information Reuse Estimator

Multifidelity Importance Sampling (MFIS)

Efficient estimation of low probability events, leveraging multiple models

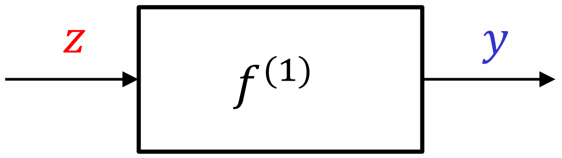
Peherstorfer, Cui, Marzouk, W., Multifidelity importance sampling, CMAME, 2016

Peherstorfer, Kramer, W., Combining multiple surrogate models to accelerate failure probability estimation with expensive high-fidelity models, in review

Estimating a failure probability

via Monte Carlo sampling

- uncertain **input** $z \in \mathcal{Z}$
- **output** quantity of interest $y \in \mathcal{Y}$
- high-fidelity model $f^{(1)}: \mathcal{Z} \to \mathcal{Y}$ with cost $w_1 > 0$ ("truth")



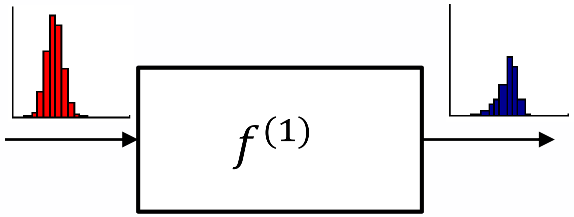
- define indicator function $I^{(1)}: \mathcal{Z} \to \mathcal{Y}$ as

$$I^{(1)}(z) = \begin{cases} 1, & \text{if } f^{(1)}(z) < 0 \\ 0, & \text{else} \end{cases} \quad \leftarrow \text{failure event}$$

- random variable Z with probability density p

- goal: estimate failure probability $P_f = \mathbb{E}_p[I^{(1)}(\mathrm{Z})]$
- Monte Carlo estimation of P_f using N realizations $z_1, \ldots, z_N$:

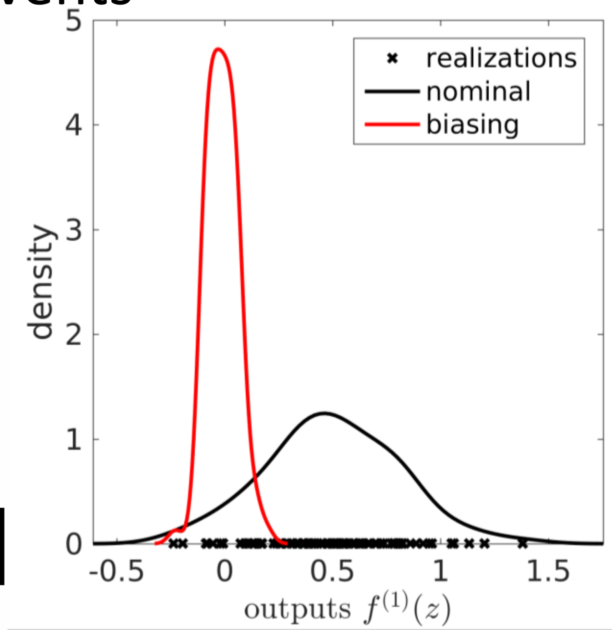
$$P_f^{\mathrm{MC}} = \frac{1}{N} \sum_{i=1}^{N} I^{(1)}(z_i)$$



Estimating a failure probability

via importance sampling

- Importance sampling: create biasing density q that puts more weight on failure events



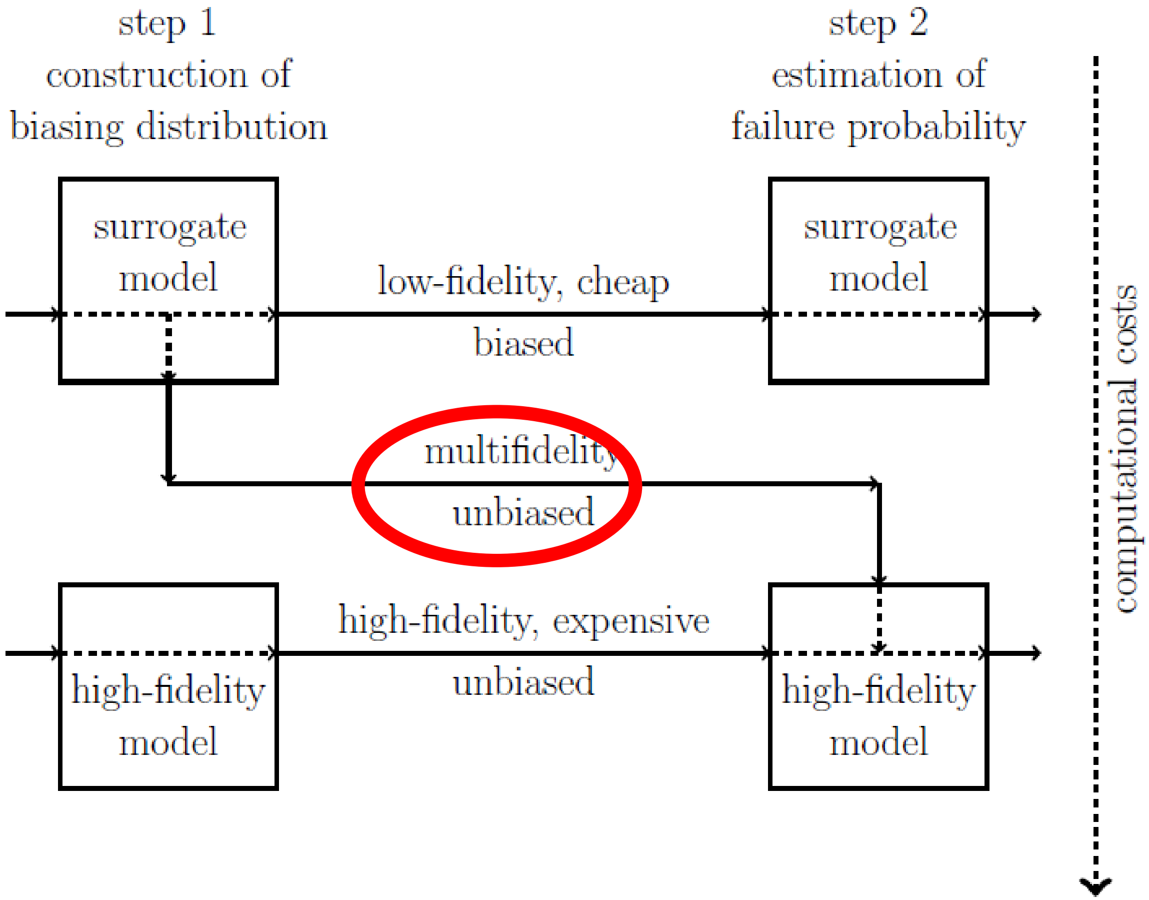
- Let $\hat{Z}$ be the corresponding RV
- Introduce the weight function
$$w(z) = \frac{p(z)}{q(z)}$$
- Reformulate failure probability as
$$P_f = \mathbb{E}_p\left[I^{(1)}(Z)\right] = \mathbb{E}_q\left[I^{(1)}(\hat{Z})w(\hat{Z})\right]$$
- Goal: construct a biasing density q such that
$$\mathrm{Var}_q\left[I^{(1)}(\hat{Z})w(\hat{Z})\right] < \mathrm{Var}_p\left[I^{(1)}(Z)\right]$$
- Lower variance means fewer realizations of $\hat{Z}$ than of Z are necessary to achieve the same MSE $\longrightarrow$ fewer model evaluations

Multifidelity importance sampling (MFIS)

with two models

Peherstorfer, Cui, Marzouk, W., *Computer Methods in Applied Mechanics and Engineering*, 2016

- We derive q with surrogate $f^{(2)}$, and use $f^{(1)}$ to estimate P_f
  - Step 1: Construction of biasing distribution (“speedup”)
  - Step 2: Estimation of P_f using q (“establish accuracy guarantees”)



Multifidelity importance sampling

Step 1: construction of biasing density

- Draw *many* realizations $z_1, \dots, z_N$ of Z (nominal)

- Evaluate **surrogate model** to obtain outputs

$$f^{(2)}(z_1), \dots, f^{(2)}(z_N)$$

- Fit normal dist. q to realizations that correspond to failure

$$\{z_i \mid I^{(2)}(z_i) = 1, i = 1, \dots, N\}$$

- Use Expectation-Maximization (EM) algorithm to fit density

- Derive random variable $\hat{Z}$ with distribution given by q

Multifidelity importance sampling

Step 2: estimation of failure probability

- Draw $M \in \mathbb{N}$ realizations $\hat{z}_1, \dots, \hat{z}_M$ of $\hat{Z}$ (biasing)
- Evaluate **high-fidelity model** to obtain outputs

$$f^{(1)}(\hat{z}_1), \dots, f^{(1)}(\hat{z}_M)$$

  - typically have $M \ll N$, and therefore fewer high-fidelity model evaluations
- Derive the multifidelity importance sampling (MFIS) estimate

$$P_f^{\mathrm{MFIS}} = \frac{1}{M} \sum_{i=1}^{M} I^{(1)}(\hat{z}_i)\, w(\hat{z}_i)$$

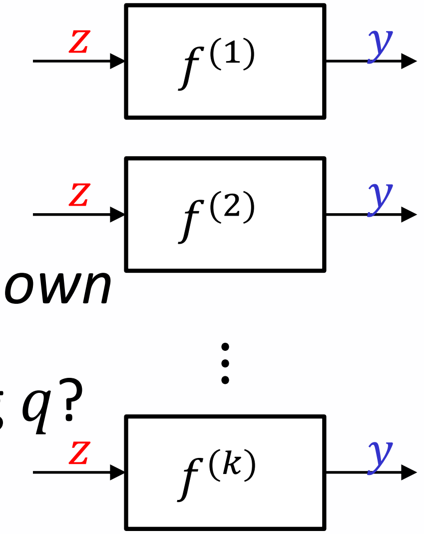
- We can show unbiasedness of the MFIS estimator

$$P_f = \mathbb{E}_q[P_f^{\mathrm{MFIS}}]$$

Mixed MFIS

extending MFIS to multiple models

- Given are $k-1$ models
  $$f^{(2)}, \dots, f^{(k)}: \mathcal{Z} \to \mathcal{Y}$$

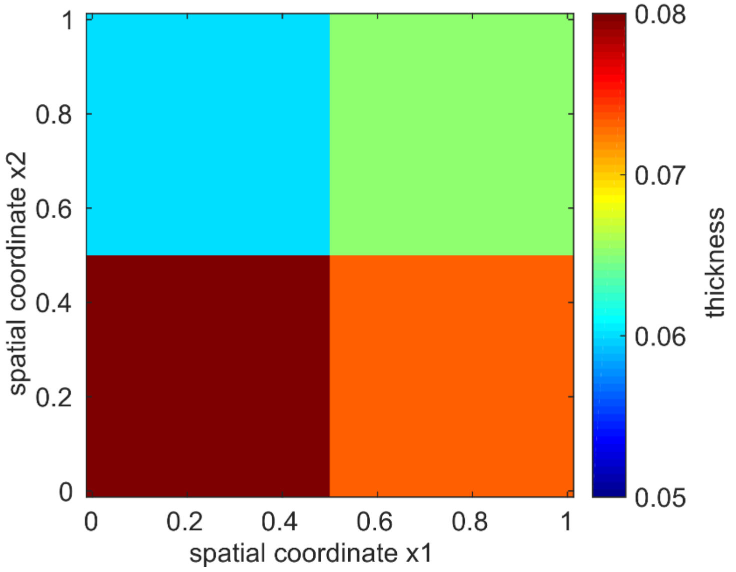


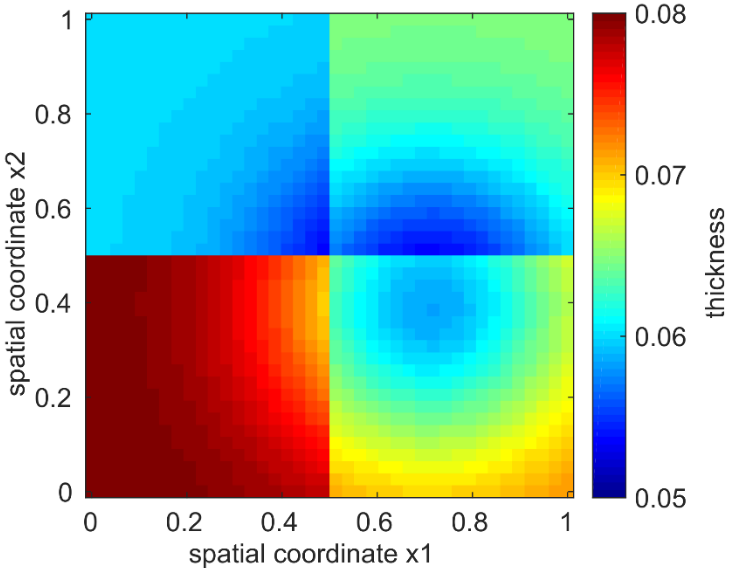
- Approximation qualities of these sources *unknown*
- Which of these should we use for constructing q?
- Our approach: Mixed MFIS
  - Use each surrogate $f^{(i)}$ to construct a density q_i, for $i = 2, \dots, k$
  - Sample from all these densities $q_2, \dots, q_k$ and combine samples
  - Mixed MFIS estimator P_f^{Mixed} derived as in [Owen et al, 2000]
- Known that
  $$\frac{\mathrm{Var}\left[P_f^{\mathrm{Mixed}}\right]}{k-1} \leq \min_{i=2,\dots,k} \mathrm{Var}\left[I^{(1)} \frac{p}{q_i}\right]$$
- Our P_f^{Mixed} is up to factor $k-1$ as good as using the surrogate that minimizes variance

Example: Locally damaged plate (multiple models)

Peherstorfer & W., 2015

- Locally damaged plate
- Inputs: nominal thickness, load, two damage parameters
- Inputs uniformly distributed in [0.05, 0.1] x [1, 100] x [0, 0.2] x (0, 0.05]
- QoI: maximum deflection of plate



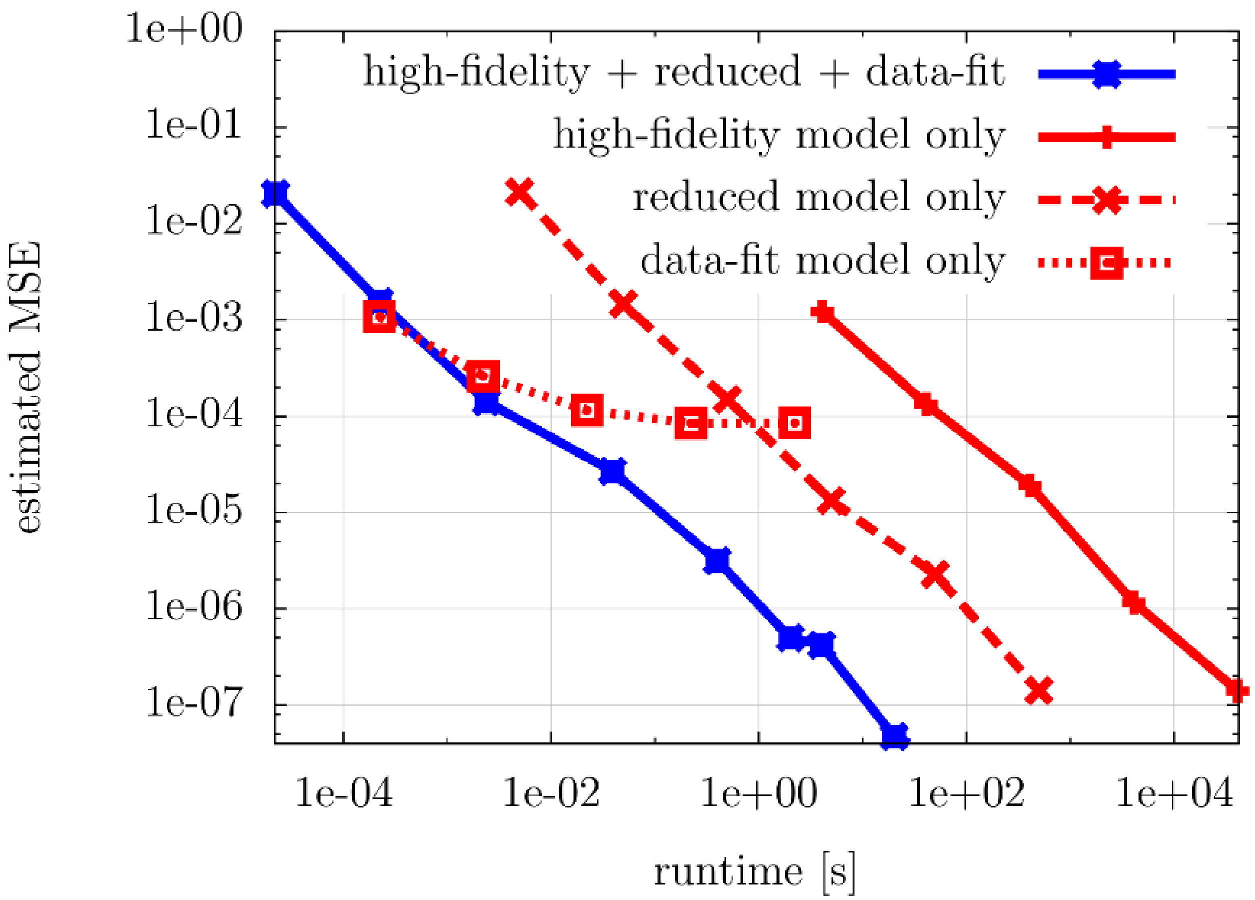


- Six models available
  - High-fidelity model: FEM, 300 dof
  - Reduced model: POD, 10 dof
  - Reduced model: POD, 2 dof
  - Reduced model: POD, 5 dof
  - Data-fit model: linear interpolation, 256 pts
  - Support vector machine: 256 pts

- Variance, correlation, runtime estimated from 100 samples

Locally damaged plate: MFMC

MFMC estimation of mean deflection achieves up to 4 orders of magnitude reduction in computational cost



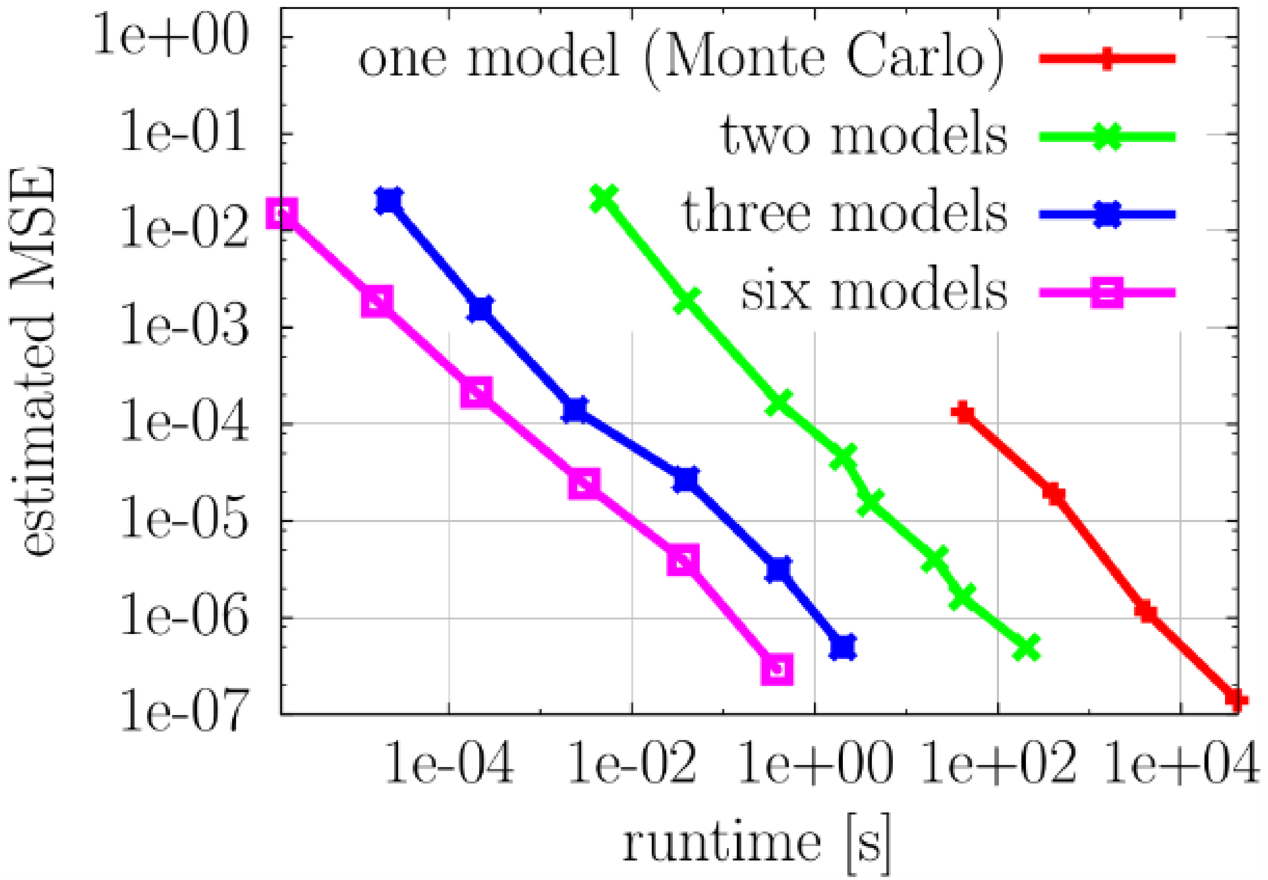
- Combine high-fidelity + reduced (POD, 10) + data-fit (linear interp, 256)
- Reduced and data-fit model lead to biased estimator, MFMC is unbiased

Locally damaged plate: MFMC mean estimate

Successively add reduced (POD, 10), data-fit (linear interp, 256), and then all others

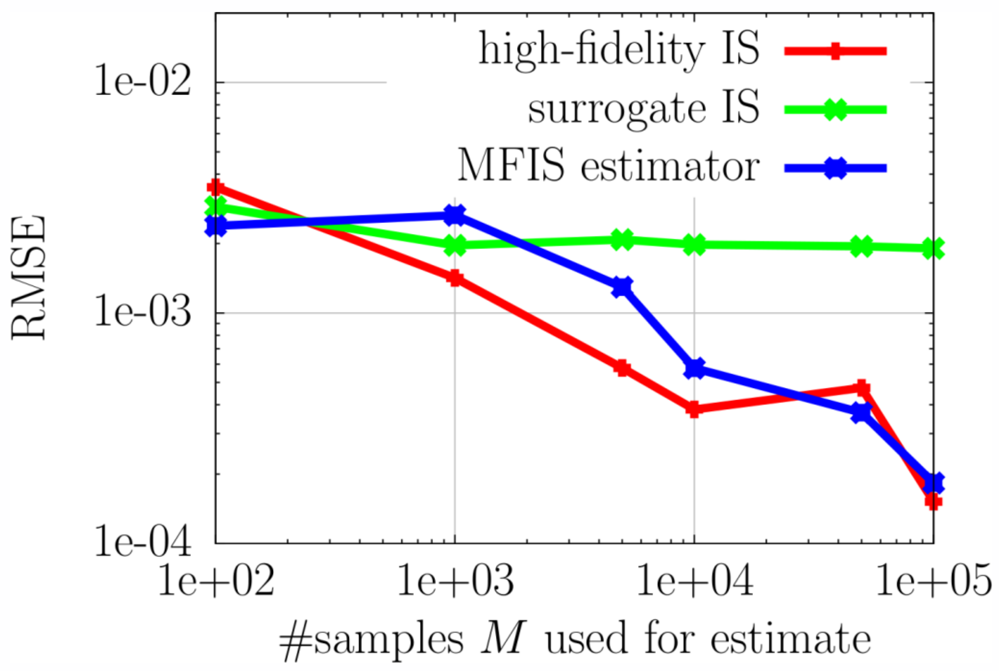
Adding data-fit model reduces variance, even though data-fit model is poor approximation of high-fidelity model

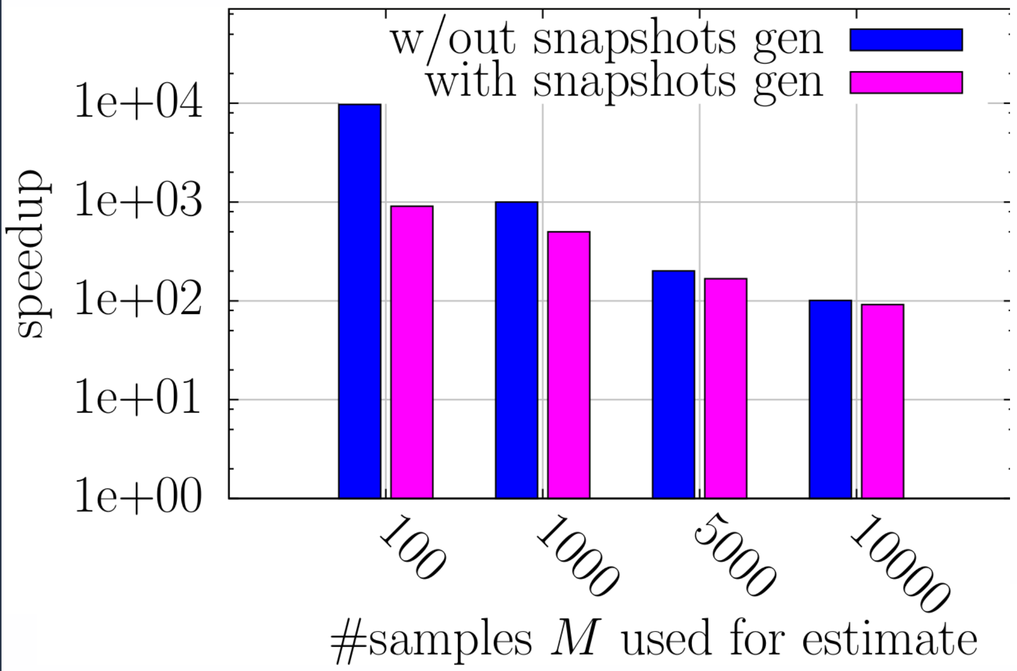
MFMC achieves almost **4 orders of magnitude improvement** over standard Monte Carlo simulation with high-fidelity model only.



Locally damaged plate: MFIS

Estimate the probability that the deflection exceeds a critical value





- Biasing density constructed from $N = 10^6$ realizations
- Using surrogate only leads to large bias
- MFIS leads to unbiased estimate of P_f
- If ROM available, speedup of up to 10^4, cf. high-fidelity

Conclusion

Multifidelity strategies for the outer loop:

leverage approximate models

but maintain guarantees on outer-loop result

- Multifidelity Monte Carlo (MFMC): a control variate formulation for estimating means
- Multifidelity Importance Sampling (MFIS): an importance sampling formulation for estimating probabilities
- MFMC extension to estimating variance and sensitivity indices (Qian MS89, MS145)

Acknowledgements

- Air Force Office of Sponsored Research MURI on Multi-Physics Multi-Source Information (J.-L. Cambier)
- Department of Energy Applied Mathematics Program: DiaMonD Multifaceted Mathematics Integrated Capability Center (A. Patra, S. Lee)
- DARPA EQUiPS program QUANTUM Project (F. Fahroo)