The University of Texas at Austin
Institute for Computational Engineering and Sciences

MULTI-FIDELITY MODELS & METHODS

Fusing models and data to achieve efficient design, optimization, and uncertainty quantification

KAREN E. WILLCOX

Science at Extreme Scales:
Where Big Data Meets Large-Scale Computing Tutorials

Institute for Pure and Applied Mathematics

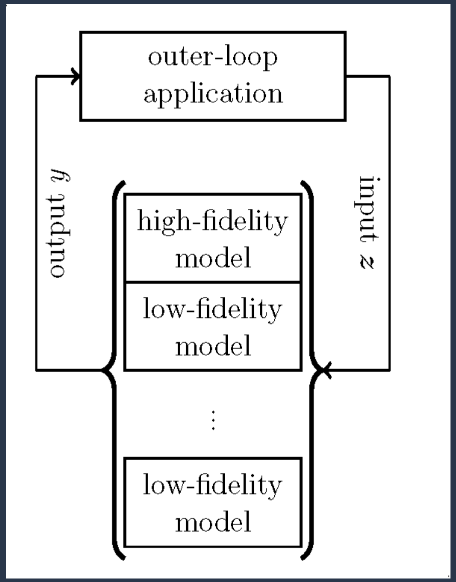
September 18, 2018

ICES

Tutorial Outline

Content largely follows Peherstorfer, W. & Gunzburger. “Survey of multifidelity methods in uncertainty propagation, inference, and optimization,” *SIAM Review*, Vol. 60, No. 3, 2018.

1. Why use multifidelity modeling?
2. What is multifidelity modeling?
3. Multifidelity optimization (MFO)
4. Multifidelity Monte Carlo (MFMC)
5. Multifidelity Importance Sampling (MFIS)
6. Multi Information Source Optimization (MISO)
7. Conclusions



1. Motivation

Why use multifidelity modeling?

Why use a multifidelity formulation?

Reduced model
(approximate)

Full model
(“truth”)

Why use a multifidelity formulation?

Reduced model
(approximate)

Full model
("truth")

Computationally
cheap(er)

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

2. Multifidelity models and multifidelity methods

"All models are wrong, but some are useful."

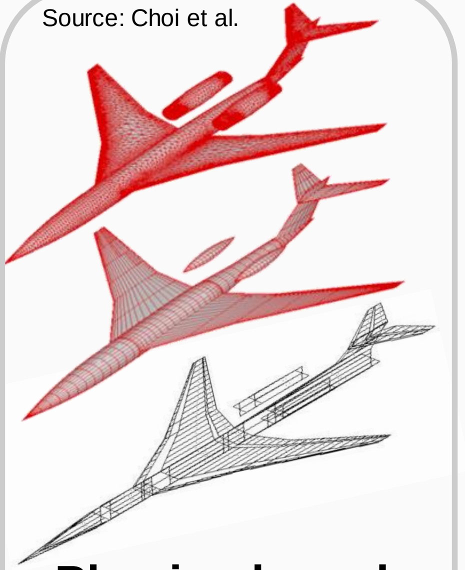
George Box, 1979

Decisions are informed by multiple sources of information

Analysis and design typically **begin with low-fidelity** models and progressively incorporate higher fidelity tools

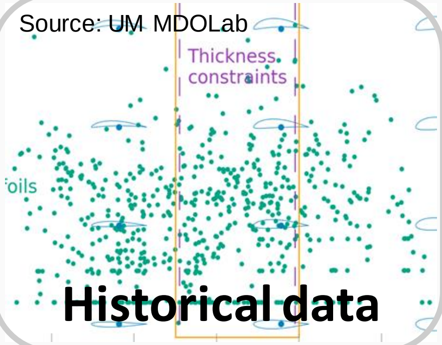
Many information sources available: multifidelity models, historical data, operational data, experimental data, expert opinions

Source: Choi et al.

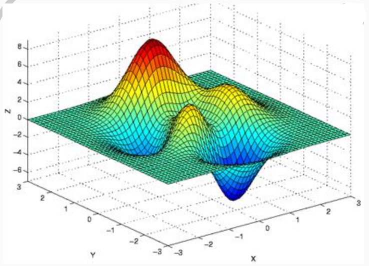


Physics-based models

Source: UM MDOLab



Historical data



Surrogate & reduced models

Source: wikipedia.org

Experts

Experimental data

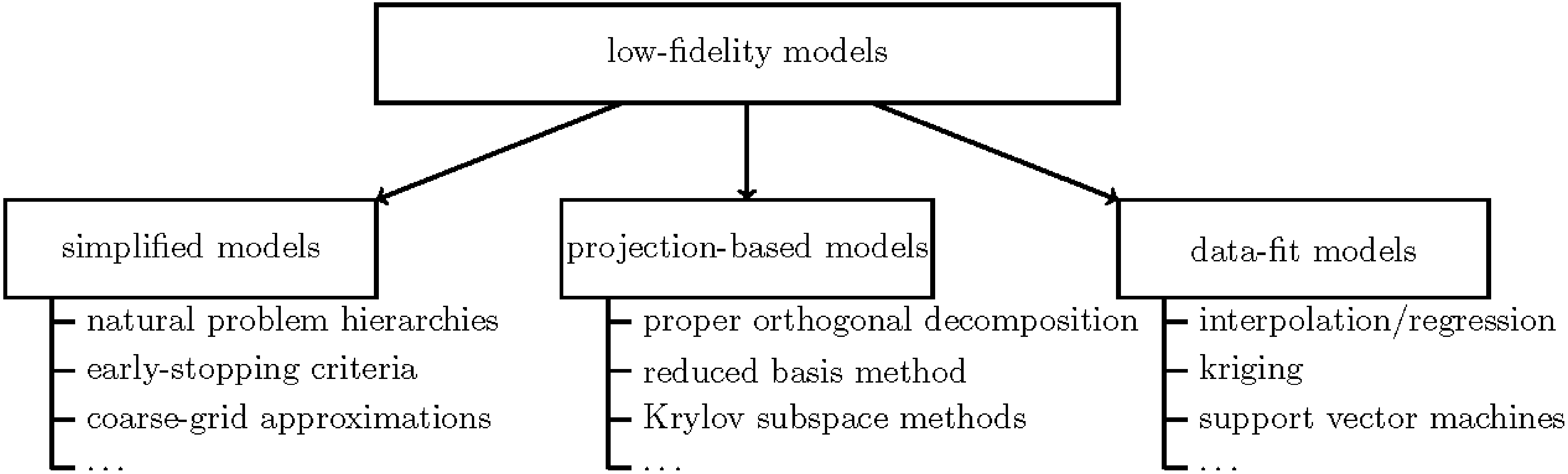
Telling us different things about the system:
the collective information is greater than the individual parts

Critical to get the right information early in the decision process

Multifidelity models can come in different forms

- Covering a range of different resolutions, scales, modeling assumptions, etc.

- Simplified physics, loosened tolerance, coarse grid, data-fit, projection-based reduced models, etc.

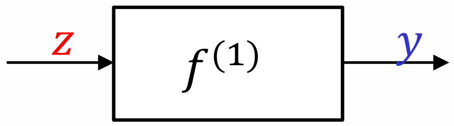


Multifidelity models can come in different forms

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

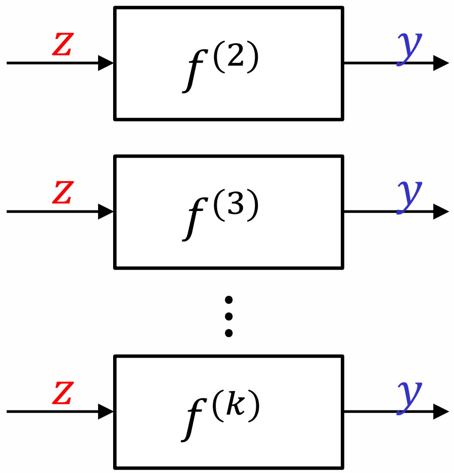
Simplified physics, loosened tolerance, coarse grid, data-fit, projection-based ROM, etc.

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



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

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

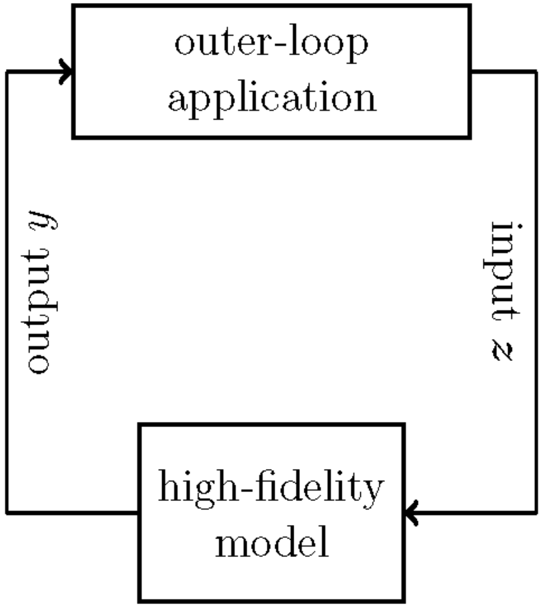


$$f^{(2)}, \dots, f^{(k)}: \mathcal{Z} \rightarrow \mathcal{Y}$$

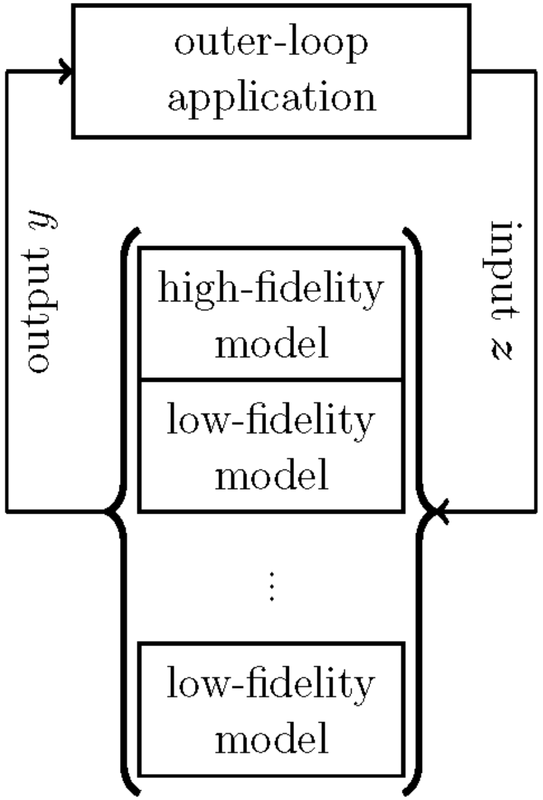
- model $f^{(i)}$ has cost w_i
- model $f^{(i)}$ has fidelity f_i
- models do not necessarily form a hierarchy

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

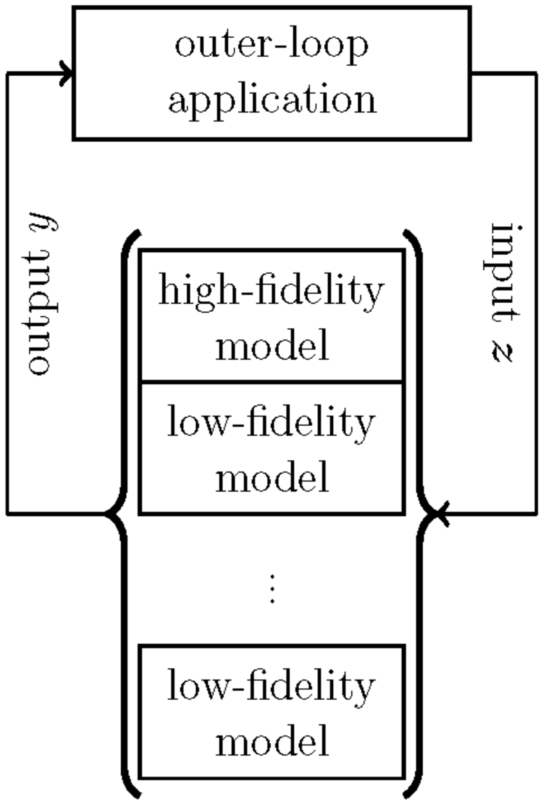


- **Multifidelity methods**: goal is to solve the outer-loop problem at high fidelity
  - invoke multiple models to reduce computational cost
  - maintains guarantees on outer-loop result



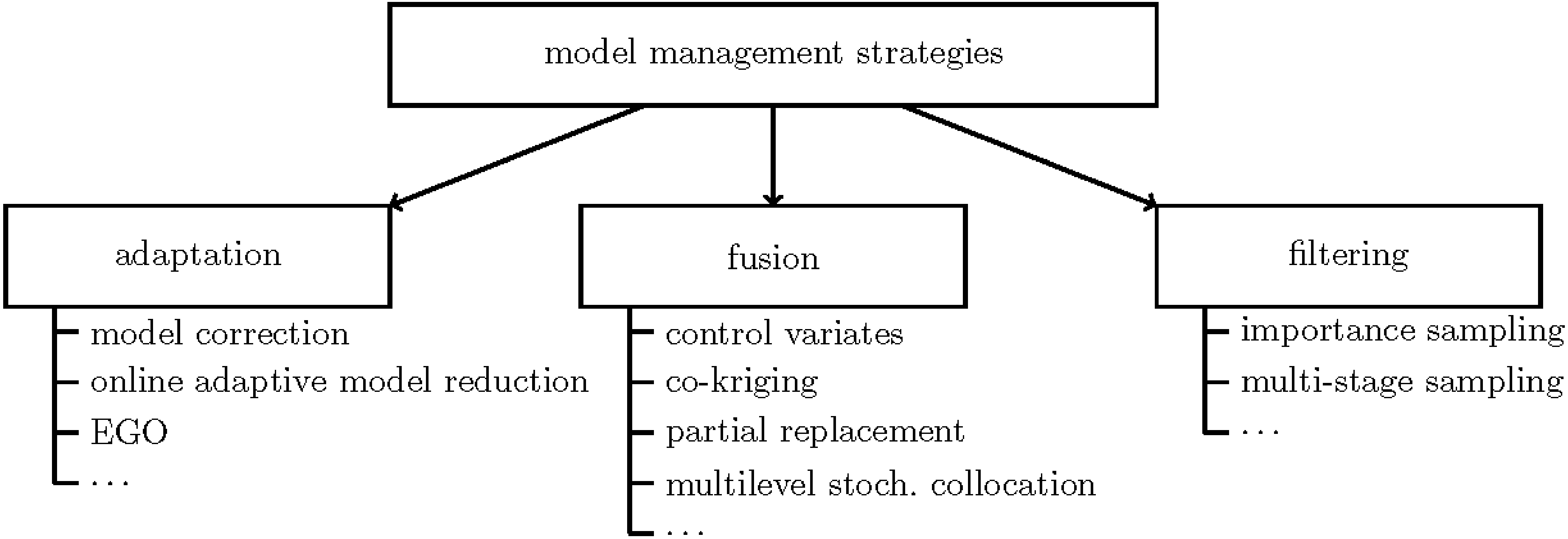
Multifidelity methods for outer-loop problems

- **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



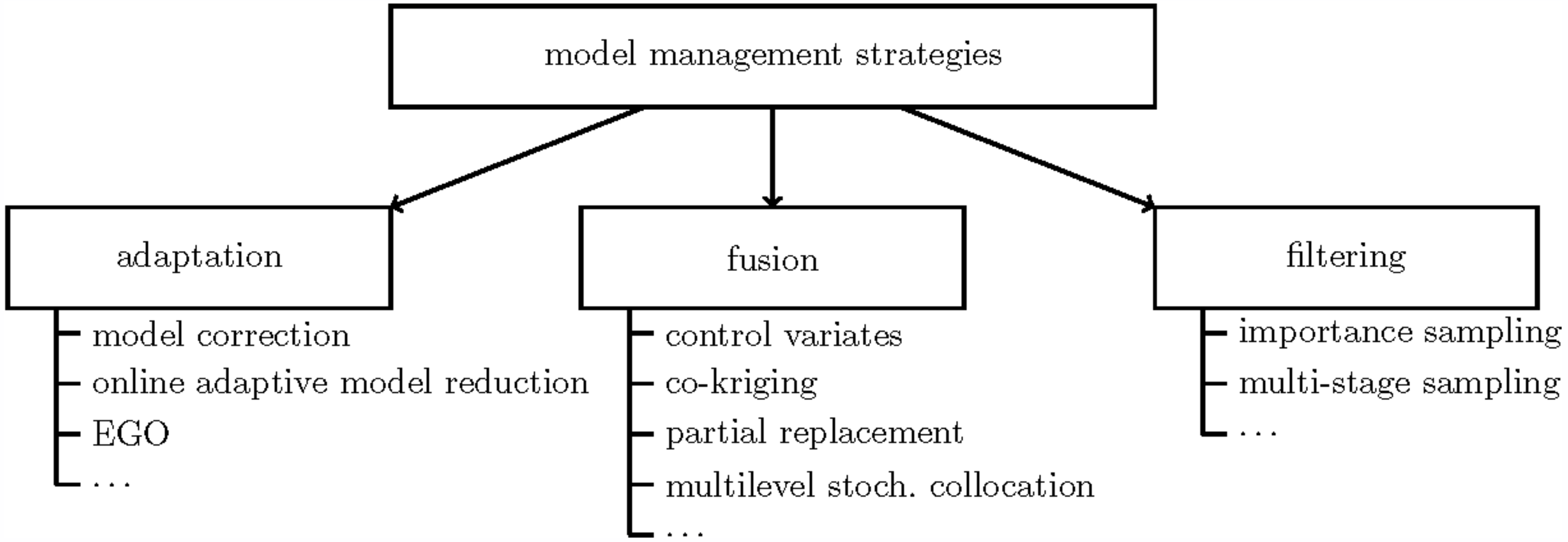
- Key questions
  - how to **combine model estimates**?
  - how to **balance evaluations** among models?
  - how to **guarantee outer-loop result**?

Multifidelity strategies for the outer loop



Multifidelity strategies

Examples of multifidelity strategies for the outer loop



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

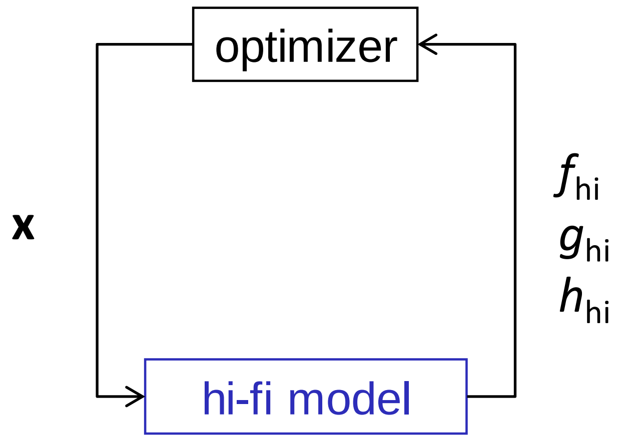
$$
\begin{aligned}
&\min_{x} f(x) \\
\text{s.t.} \quad &g(x) \leq 0 \\
&h(x) = 0
\end{aligned}
$$

3. Multifidelity Optimization

Design optimization formulation

$$\min_{x} f(x)$$
$$\text{s.t.} \quad g(x) \leq 0$$
$$h(x) = 0$$

Design variables	x
Objective	$f(x)$
Constraints	$g(x)$, $h(x)$

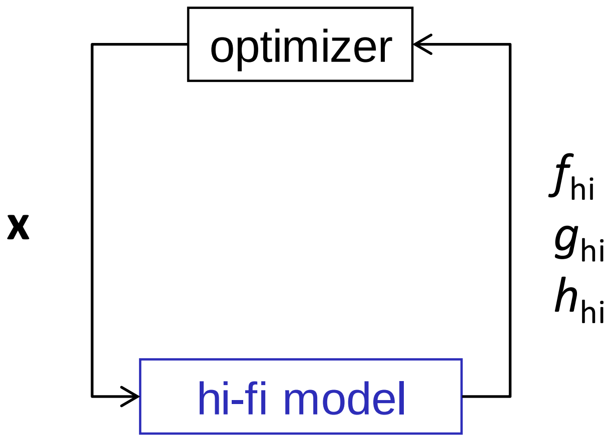


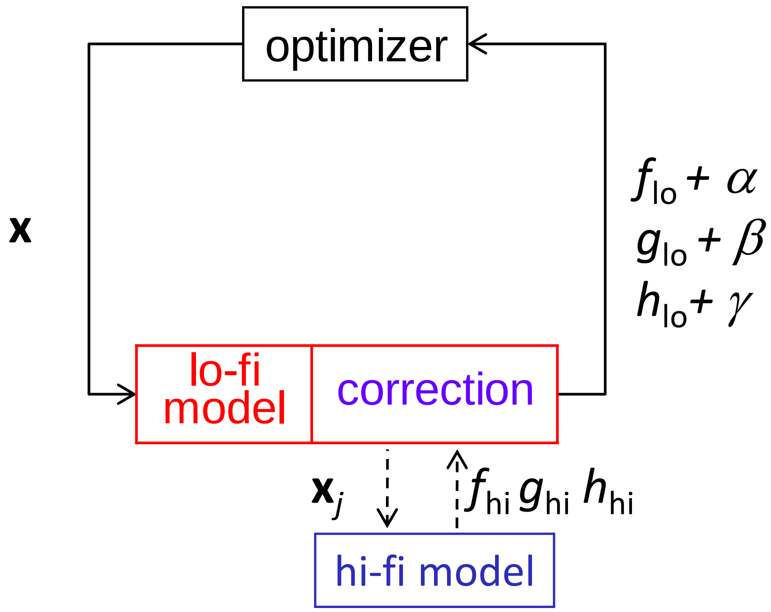
- Interested in optimization of systems governed by PDEs (constraints and objective evaluation is expensive)

Multifidelity optimization formulation

$$\min_{x} f(x)$$
$$\text{s.t.}\quad g(x) \le 0$$
$$h(x) = 0$$

Design variables	x
Objective	$f(x)$
Constraints	$g(x)$, $h(x)$





Multifidelity optimization

Defining the surrogate model

- Denote a surrogate model of $f_{\text{high}}(\mathbf{x})$ as $m(\mathbf{x})$
- The surrogate model could be:
  1. The **low-fidelity function** (reduced model)
  $$m(\mathbf{x}) = f_{\text{low}}(\mathbf{x}) \approx f_{\text{high}}(\mathbf{x})$$
  2. The **sum** of the low-fidelity function and an additive correction
  $$m(\mathbf{x}) = f_{\text{low}}(\mathbf{x}) + e(\mathbf{x}) \approx f_{\text{high}}(\mathbf{x})$$
  where $e(\mathbf{x})$ is calibrated to the difference $f_{\text{high}}(\mathbf{x})$- $f_{\text{low}}(\mathbf{x})$
  3. The **product** of a low-fidelity function and a multiplicative correction
  $$m(\mathbf{x}) = \beta_c(\mathbf{x}) f_{\text{low}}(\mathbf{x}) \approx f_{\text{high}}(\mathbf{x})$$
  where $\beta_c(\mathbf{x})$ is calibrated to the quotient $f_{\text{high}}(\mathbf{x}) / f_{\text{low}}(\mathbf{x})$
- Update the correction terms as the optimization algorithm proceeds and additional evaluations of $f_{\text{high}}(\mathbf{x})$ become available

Multifidelity optimization: Trust-region model management

- At iteration k, define a trust region centered on iterate $\mathbf{x}_k$ with size Δ_k

$$\mathcal{B}_k = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}_k\| \leq \Delta_k\}$$

- m_k is the surrogate model on the kth iteration

- Determine a trial step $\boldsymbol{s}_k$ at iteration k, by solving a subproblem of the form:

$$\min_{\mathbf{x}} \;\; f(\mathbf{x}) \quad \Longrightarrow \quad \boxed{\begin{array}{ll} \min\limits_{\mathbf{s}_k} & m_k(\mathbf{x}_k + \mathbf{s}_k) \\ \text{s.t.} & \|\mathbf{s}_k\| \leq \Delta_k \end{array}}$$

(unconstrained case)

Multifidelity optimization: Trust-region model management

- Evaluate the function at the trial point: $f_{\text{high}}(\mathbf{x}_k + \mathbf{s}_k)$

- Compute the ratio of the actual improvement in the function value to the improvement predicted by the surrogate model:

$$\rho_k = \frac{f_{\text{high}}(\mathbf{x}_k) - f_{\text{high}}(\mathbf{x}_k + \mathbf{s}_k)}{m_k(\mathbf{x}_k) - m_k(\mathbf{x}_k + \mathbf{s}_k)}$$

- Accept or reject the trial point and update trust region size according to (typical parameters):

$\rho^k \leq 0$	Reject step	$\Delta^{k+1} \equiv 0.5\Delta^k$
$0 < \rho^k \leq 0.1$	Accept step	$\Delta^{k+1} \equiv 0.5\Delta^k$
$0.1 < \rho^k < 0.75$	Accept step	$\Delta^{k+1} \equiv \Delta^k$
$0.75 \leq \rho^k$	Accept step	$\Delta^{k+1} \equiv 2\Delta^k$

Trust-Region Algorithm for Iteration k

1. Compute a step, $\mathbf{s}_k$, by solving the trust-region subproblem,

$$\begin{aligned} \min_{\mathbf{s}_k} \quad & m_k(\mathbf{x}_k + \mathbf{s}_k) \\ \text{s.t.} \quad & \|\mathbf{s}_k\| \leq \Delta_k. \end{aligned}$$

2. Evaluate $f_{\text{high}}(\mathbf{x}_k + \mathbf{s}_k)$.

3. Compute the ratio of actual improvement to predicted improvement,

$$\rho_k = \frac{f_{\text{high}}(\mathbf{x}_k) - f_{\text{high}}(\mathbf{x}_k + \mathbf{s}_k)}{m_k(\mathbf{x}_k) - m_k(\mathbf{x}_k + \mathbf{s}_k)}.$$

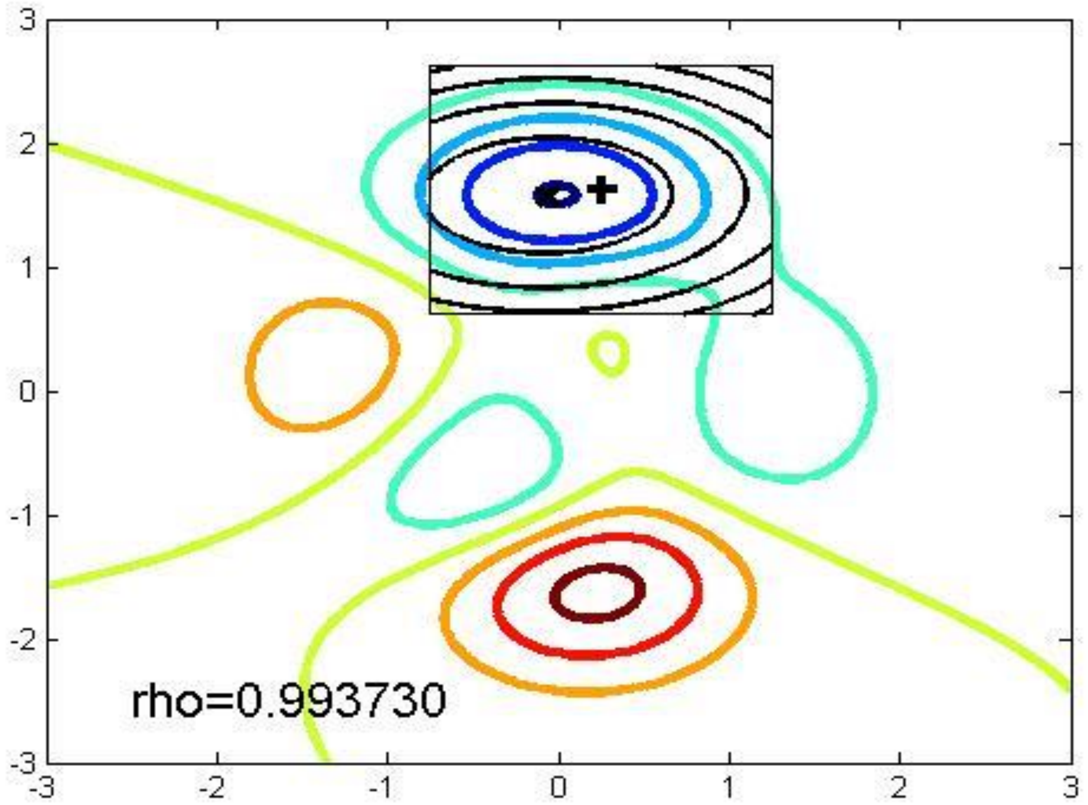
4. Accept or reject the trial point according to ρ_k,

$$\mathbf{x}_{k+1} = \begin{cases} \mathbf{x}_k + \mathbf{s}_k & \text{if } \rho_k > 0 \\ \mathbf{x}_k & \text{otherwise.} \end{cases}$$

5. Update the trust region size according to ρ_k,

$$\Delta_{k+1} = \begin{cases} \gamma_1 \Delta_k & \text{if } \rho_k \leq \eta_1 \\ \Delta_k & \text{if } \eta_1 < \rho_k < \eta_2 \\ \gamma_2 \Delta_k & \text{if } \rho_k \geq \eta_2. \end{cases}$$

Trust-Region Demonstration



Trust-region model management: Corrections and convergence

- Provably convergent to local minimum of high-fidelity function if surrogate is first-order accurate at center of trust region [Alexandrov et al., 2001]

- Additive correction: $a(\mathbf{x}) = f_{\text{high}}(\mathbf{x}) - f_{\text{low}}(\mathbf{x})$

  with surrogate constructed as

$$m_k(\mathbf{x}) = f_{\text{low}}(\mathbf{x}) + a(\mathbf{x}_k) + \nabla a(\mathbf{x}_k)^T(\mathbf{x} - \mathbf{x}_k)$$

- Multiplicative correction: $\beta(\mathbf{x}) = \dfrac{f_{\text{high}}(\mathbf{x})}{f_{\text{low}}(\mathbf{x})}$

  with surrogate constructed as

$$m_k(\mathbf{x}) = \left[\beta(\mathbf{x}_k) + \nabla\beta(\mathbf{x}_k)^T(\mathbf{x} - \mathbf{x}_k)\right] f_{\text{low}}(\mathbf{x})$$

- Only first-order corrections required to guarantee convergence; quasi-second-order corrections accelerate convergence [Eldred et al., 2004]

- Trust-region POD [Arian, Fahl, Sachs, 2000]

Trust-region model management: Derivative-free framework

- Derivative-free trust region approaches
  [Conn, Scheinberg, and Vicente, 2009]

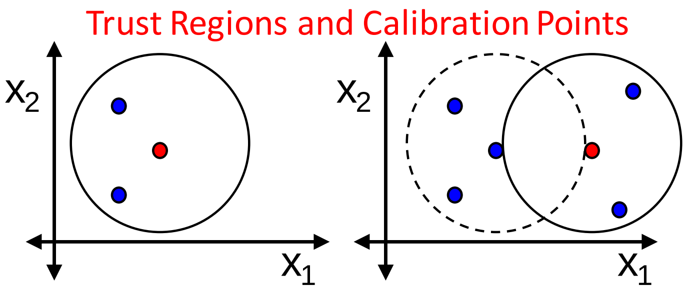
- Provably convergent under appropriate conditions if the surrogate model is "fully linear"

$$\left\| \nabla f_{high}(\mathbf{x}) - \nabla m_k(\mathbf{x}) \right\| \leq \kappa_g \Delta_k$$

$$\left| f_{high}(\mathbf{x}) - m_k(\mathbf{x}) \right| \leq \kappa_f \Delta_k^2$$

- Achieved through adaptive corrections or adaptive calibration e.g., radial basis function calibration with sample points chosen to make surrogate model fully linear by construction
  [Wild, Regis and Shoemaker, 2011; Wild and Shoemaker, 2013]

- Key: never need gradients wrt the high-fidelity model



4. Multifidelity Monte Carlo (MFMC)

Efficient uncertainty propagation leveraging multiple models

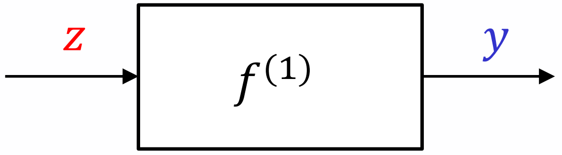
Ng & W., **Multifidelity approaches for optimization under uncertainty**, *IJNME*, 2014

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

Estimating QoI statistics

via Monte Carlo sampling

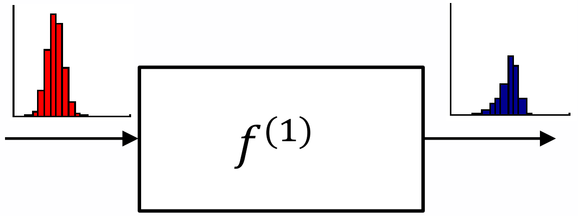
- uncertain **input** $z \in \mathcal{Z}$
- **output** quantity of interest $y \in \mathcal{Y}$
- high-fidelity **model**
  $f^{(1)}: \mathcal{Z} \to \mathcal{Y}$
  with cost $w_1 > 0$ ("truth")



- 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

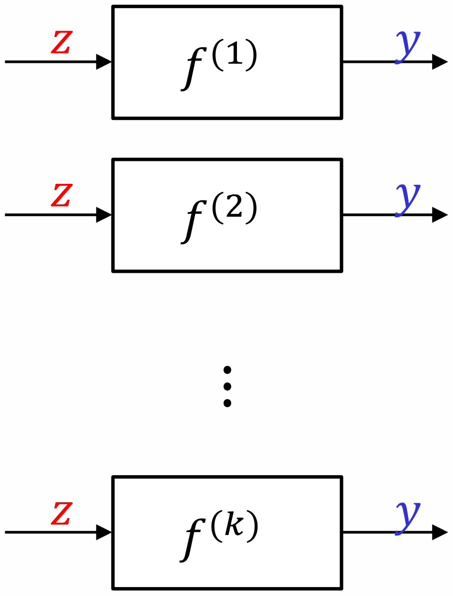
  $$f^{(2)}, \dots, f^{(k)}: \mathcal{Z} \to \mathcal{Y}$$

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

- m_i evaluations for model i, with

  $$m_1 \leq m_2 \leq \dots \leq m_k$$

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

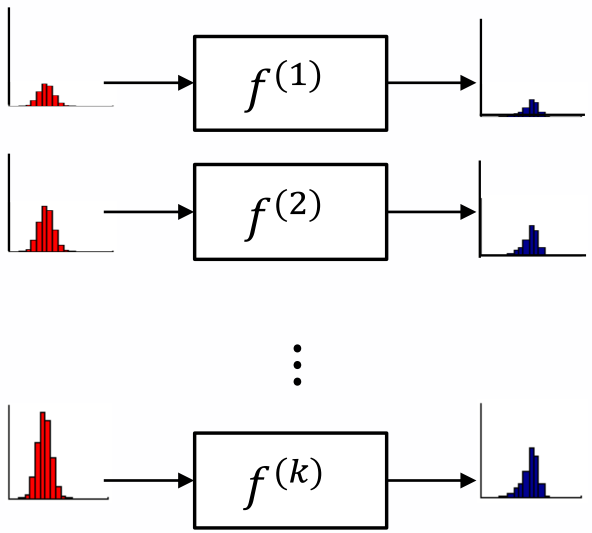


Multifidelity Monte Carlo

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

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

- 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

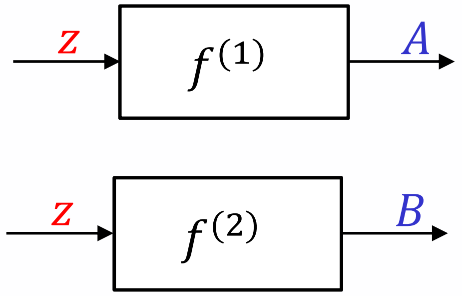
mean estimate using m_{i-1} evaluations of model i

Multifidelity Monte Carlo

Example for intuition: two models ($\boldsymbol{k = 2}$)

Ng & W., AIAA 2012, *IJNME* 2014, *J. Aircraft* 2015

random uncertain parameters



random output of high-fidelity model

random output of low-fidelity model

z_i = samples of Z
$a_i = f^{(1)}(z_i)$ = samples of A
$b_i = f^{(2)}(z_i)$ = samples of $B = a_i +$ error

s_A = statistics of A (e.g., mean, variance)
$\hat{s}_A$ = estimator of s_A

Monte Carlo variance reduction

Classical control variate approach

- Regular MC estimator for $s_A = \mathbb{E}[A]$ using n samples of A:

$$\bar{a}_n = \frac{1}{n}\sum_{i=1}^{n} a_i \qquad \mathrm{Var}[\bar{a}_n] = \frac{\sigma_A^2}{n}$$

Definitions:

$\sigma_A^2 = \mathrm{Var}[A]$

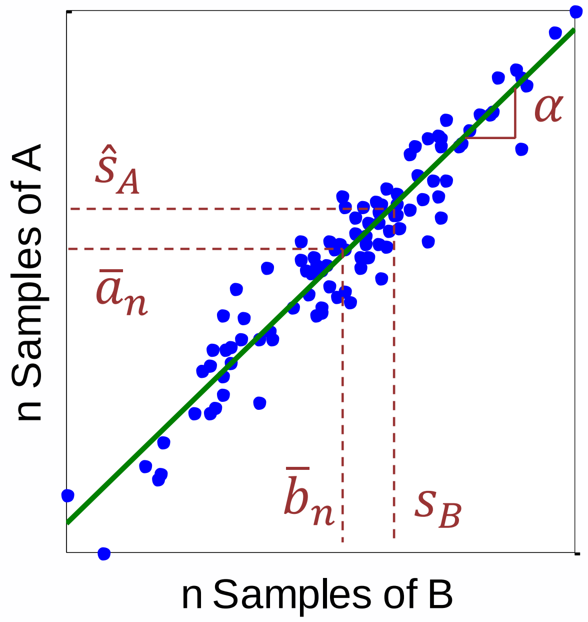
$\sigma_B^2 = \mathrm{Var}[B]$

$\rho_{AB} = \mathrm{Corr}[A, B]$

- Control variate estimator of s_A:
  - Additional random variable B with known $s_B = \mathbb{E}[B]$

$$\hat{s}_A = \bar{a}_n + \alpha\left(s_B - \bar{b}_n\right)$$

$$\mathrm{Var}[\hat{s}_A] = \frac{\sigma_A^2 + \alpha^2\sigma_B^2 - 2\alpha\rho_{AB}\sigma_A\sigma_B}{n}$$



- Minimize $\mathrm{Var}[\hat{s}_A]$ with respect to α

$$\mathrm{Var}[\hat{s}_A^*] = \underbrace{\left(1 - \rho_{AB}^2\right)}_{\leq 1}\frac{\sigma_A^2}{n}$$

Multifidelity Monte Carlo

Low-fidelity model as a control variate

Ng PhD 2013

Ng, W., 2012, 2014

- Multifidelity estimator of s_A based on control variate method:
  - A = random output of high-fidelity model
  - B = random output of low-fidelity model (s_B unknown)

Definitions:

$$\sigma_A^2 = \mathrm{Var}[A]$$

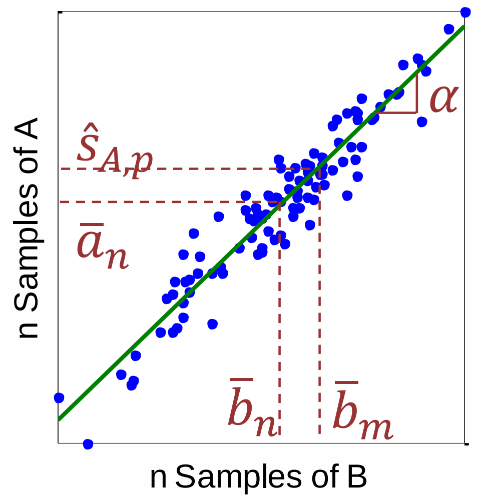
$$\sigma_B^2 = \mathrm{Var}[B]$$

$$\rho_{AB} = \mathrm{Corr}[A, B]$$

$$\hat{s}_{A,p} = \bar{a}_n + \alpha\left(\bar{b}_m - \bar{b}_n\right) \quad \text{with} \quad m \gg n$$

$$\mathrm{Var}\left[\hat{s}_{A,p}\right] = \frac{\sigma_A^2 + \alpha^2\sigma_B^2 - 2\alpha\rho_{AB}\sigma_A\sigma_B}{n} - \frac{\alpha^2\sigma_B^2 - 2\alpha\rho_{AB}\sigma_A\sigma_B}{m}$$

- Using difference $\left(\bar{b}_m - \bar{b}_n\right)$ as correction to $\bar{a}_n$
- Leveraging correlation between A and B
  - Correlation captured in α



Multifidelity Monte Carlo

optimally allocate computational budget across *k* models

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

- MFMC estimator

$$\underbrace{\hat{s}}_{\text{MFMC estimate for the mean}} = \underbrace{\bar{y}^{(1)}_{m_1}}_{\text{mean estimate using } m_1 \text{ evaluations of truth model}} + \sum_{i=2}^{k} \alpha_i \left(\underbrace{\bar{y}^{(i)}_{m_i}}_{\text{mean estimate using } m_i \text{ evaluations of model } i} - \underbrace{\bar{y}^{(i)}_{m_{i-1}}}_{\text{mean estimate using } m_{i-1} \text{ evaluations of model } i} \right)$$

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

- 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.)

Multifidelity Monte Carlo

We optimally balance the number of model evaluations to obtain the best multifidelity estimator given the computational budget

- Minimize the MSE of the MFMC estimator for a given computational budget p

- MFMC estimate $\hat{s}$ is unbiased; MSE is given by $\mathrm{Var}[\hat{s}]$

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

  - σ_i^2 is variance of $f^{(i)}(Z)$
  - ρ_i is correlation coefficient between $f^{(1)}(Z)$ and $f^{(i)}(Z)$

- Leads to optimization problem

$$\begin{array}{cc} \min\limits_{m \in \mathbb{R}^k, \alpha_2, \ldots, \alpha_k \in \mathbb{R}} & \mathrm{Var}[\hat{s}] \\ \text{such that} & m_{i-1} \leq m_i, i = 2, \ldots, k \\ & 0 \leq m_1 \\ & c(\hat{s}) = w^T m = p \end{array}$$

Balancing the number of information source (IS) evaluations

- Optimization problem has unique (analytic) solution if

$$\rho_1^2 > \rho_2^2 > \ \dots > \rho_k^2$$

and

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

- The costs/correlation ratio establishes a relationship between
  - preceding IS $f^{(i-1)}$,
  - current IS $f^{(i)}$, and
  - succeeding IS $f^{(i+1)}$

- If $f^{(1)}, \dots, f^{(k)}$ violate conditions, can construct a subset of IS's that satisfy conditions

$$\boxed{\text{IS } f^{(i-1)}} \longleftrightarrow \boxed{\text{IS } f^{(i)}} \longleftrightarrow \boxed{\text{IS } f^{(i+1)}}$$

The interactions between the IS's impact the behavior of the MFMC estimator; not the properties of the IS's alone

MFMC variance reduction

- Let $\bar{y}_n^{(1)}$ be the (benchmark) Monte Carlo estimator with computational budget p
- Ratio of MSE of MFMC estimator $\hat{s}$ and MSE of $\bar{y}_n^{(1)}$ is

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

- The MFMC estimator has lower MSE than the Monte Carlo estimator if

$$\sum_{i=1}^{k} \sqrt{\frac{w_i}{w_1} \left(\rho_i^2 - \rho_{i+1}^2 \right)} < 1$$

- Condition on the collective whole of the models (sum), not on properties of each model separately
- The interaction between the models is what drives the MFMC estimator, not model properties alone

Example: three information sources ($k = 3$)

- truth model $f^{(1)}$, cost w_1
- surrogate model $f^{(2)}$, cost w_2, correlation with truth ρ_2
- surrogate model $f^{(3)}$, cost w_3, correlation with truth ρ_3

- feasibility conditions

$$1 > \rho_2^2 > \rho_3^2, \qquad \frac{w_1}{w_2} > \frac{\rho_1^2 - \rho_2^2}{\rho_2^2 - \rho_3^2}, \qquad \frac{w_2}{w_3} > \frac{\rho_2^2 - \rho_3^2}{\rho_3^2}$$

- variance reduction (low $S \rightarrow$ low variance)

$$S(w_1, w_2, w_3, \rho_2, \rho_3) = \sum_{i=1}^{k} \sqrt{\frac{w_i}{w_1}\left(\rho_i^2 - \rho_{i+1}^2\right)}$$

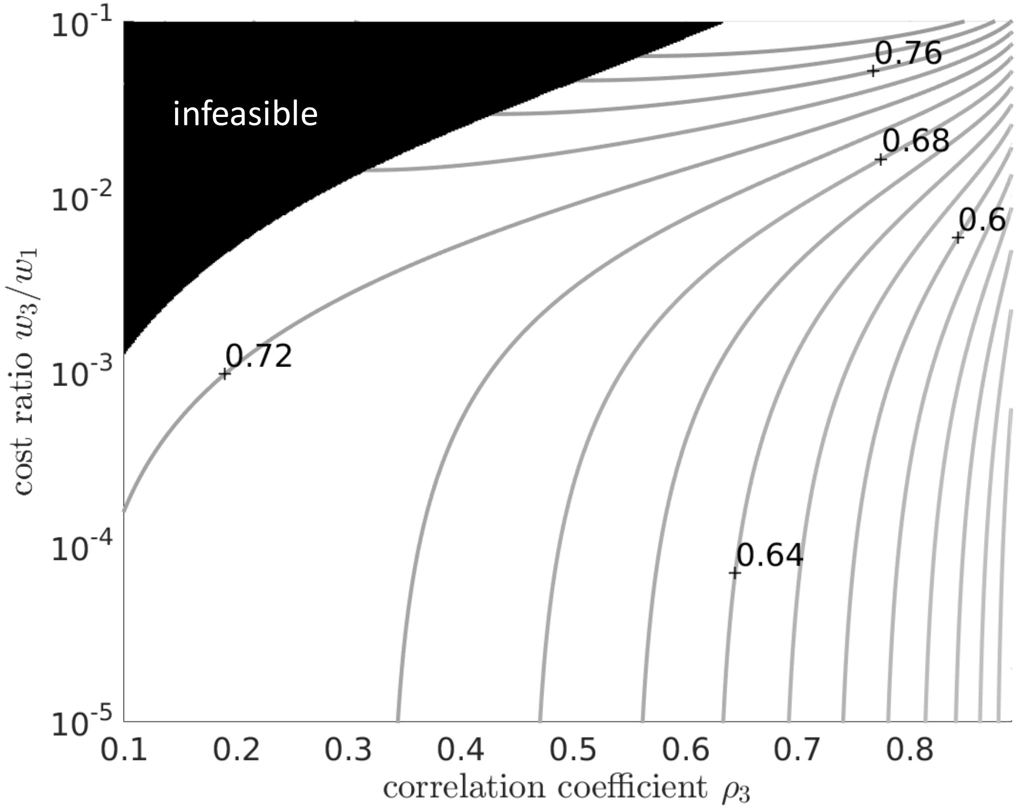
$$= \sqrt{1 - \rho_2^2} + \sqrt{\frac{w_2}{w_1}(\rho_2^2 - \rho_3^2)} + \sqrt{\frac{w_3}{w_1}\rho_3^2}$$

< 1 for the MFMC estimator to be efficient

Example: three information sources ($k = 3$)

$$S(w_1, w_2, w_3, \rho_2, \rho_3) = \sqrt{1 - \rho_2^2} + \sqrt{\frac{w_2}{w_1}(\rho_2^2 - \rho_3^2)} + \sqrt{\frac{w_3}{w_1}\rho_3^2}$$

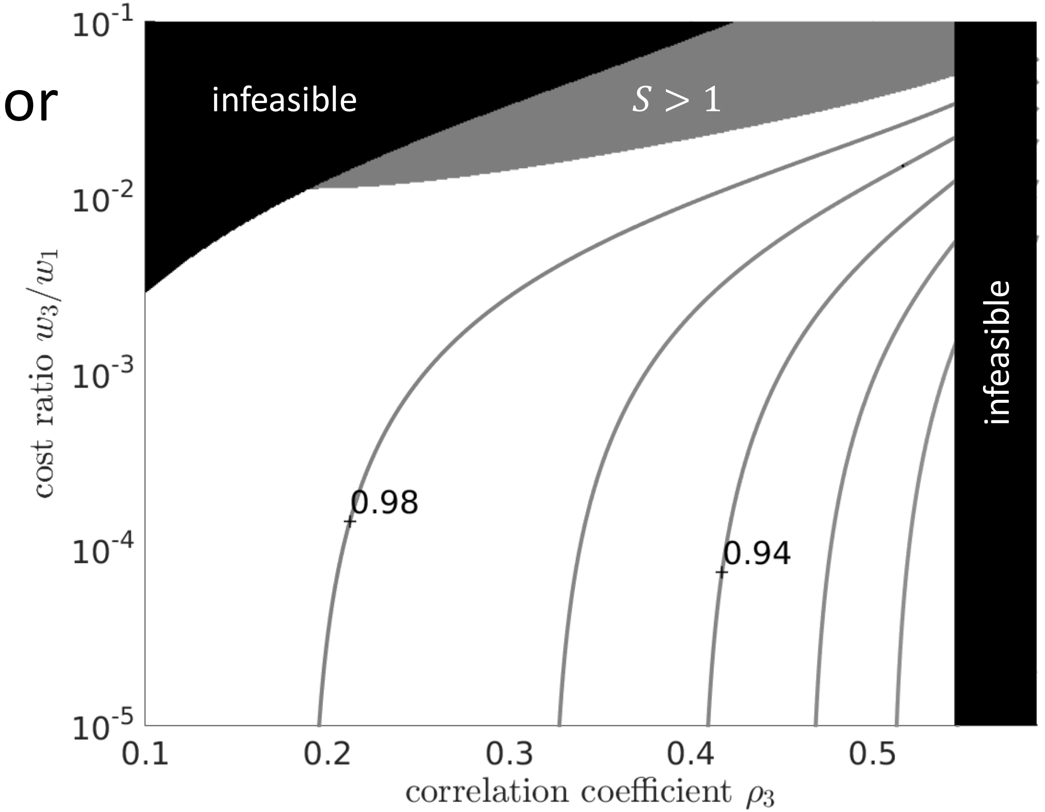
- Set $w_2/w_1 = 0.1, \;\; \rho_2 = 0.9$
- Using $f^{(1)}, f^{(2)}$ only $\rightarrow$ larger variance than MC estimator (S > 1)
- Vary w_3/w_1 and ρ_3; plot contours of S
- If costs w_3 high (> $0.01 w_1$): third term in S can dominate, increasing correlation ρ_3 can lead to larger S
- If w_3 low, then increase of ρ_3 always reduces S



Example: three information sources ($k = 3$)

$$S(w_1, w_2, w_3, \rho_2, \rho_3) = \sqrt{1 - \rho_2^2} + \sqrt{\frac{w_2}{w_1}(\rho_2^2 - \rho_3^2)} + \sqrt{\frac{w_3}{w_1}\rho_3^2}$$

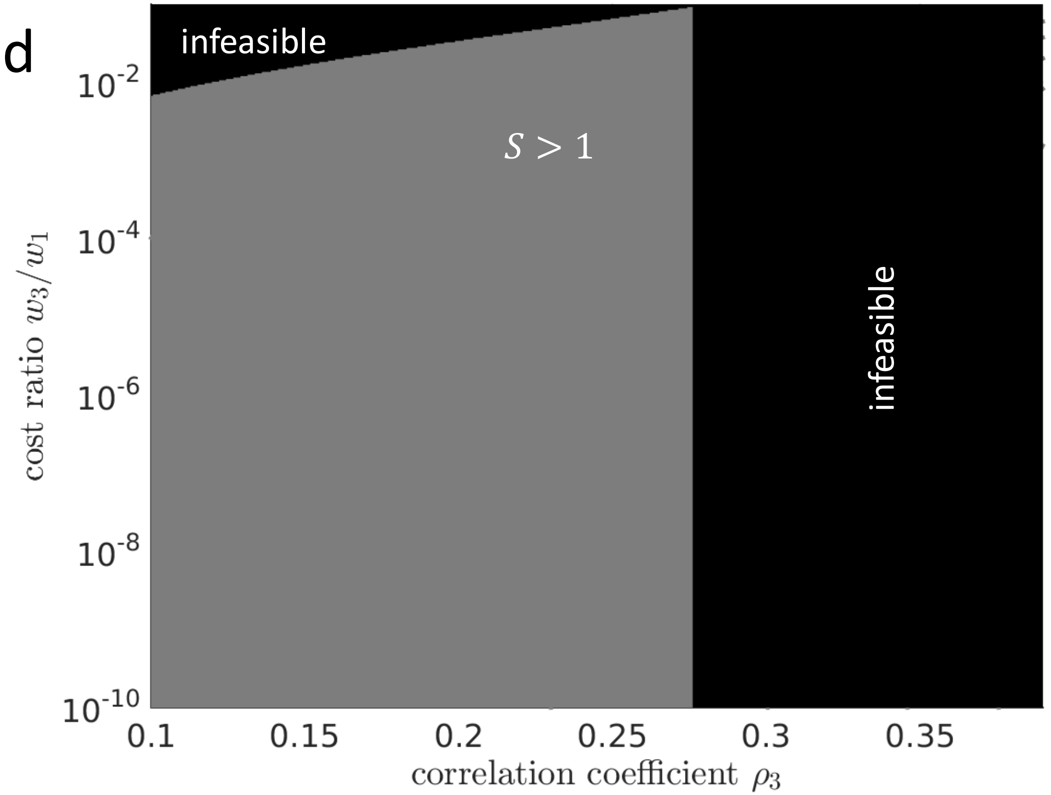
- Set $w_2/w_1 = 0.1$, $\rho_2 = 0.6$
- Increasing the correlation can violate feasibility condition: if $\rho_3 \approx \rho_2$ then denominator $\rho_2^2 - \rho_3^2$ in feasibility condition becomes small and condition is violated



Example: three information sources ($k = 3$)

$$S(w_1, w_2, w_3, \rho_2, \rho_3) = \sqrt{1 - \rho_2^2} + \sqrt{\frac{w_2}{w_1}(\rho_2^2 - \rho_3^2)} + \sqrt{\frac{w_3}{w_1}\rho_3^2}$$

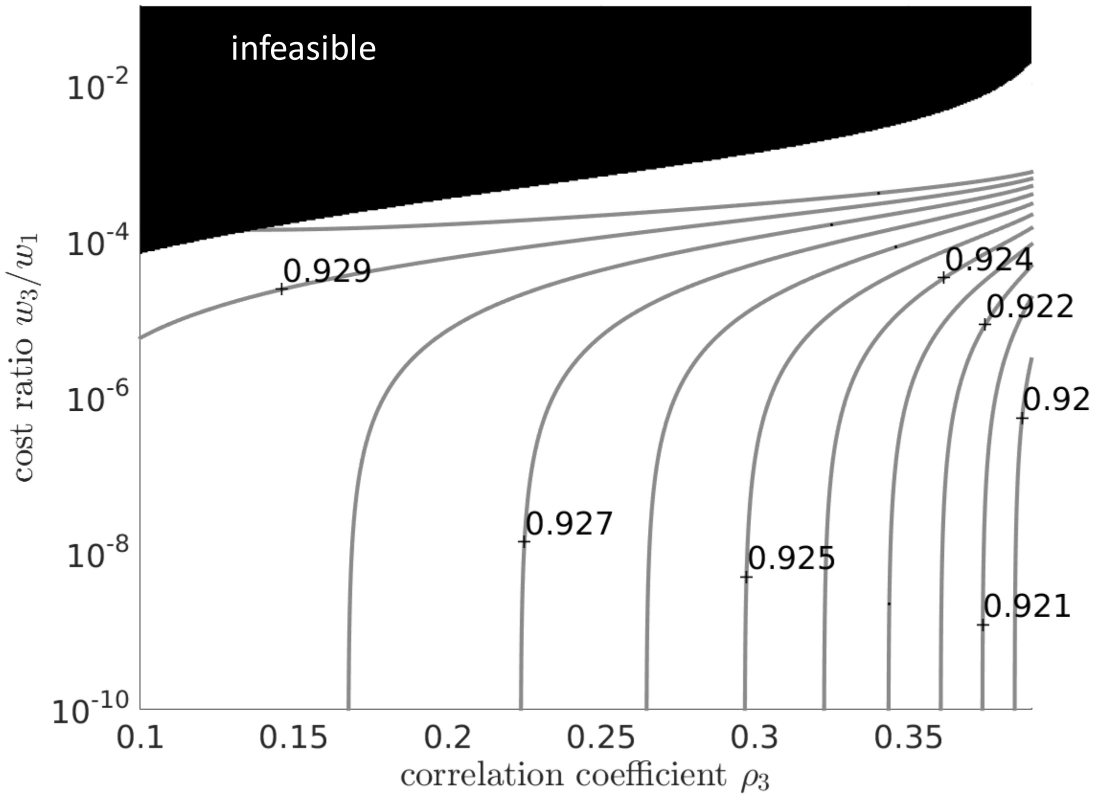
- Set $w_2/w_1 = 0.1$, $\rho_2 = 0.4$
- IS $f^{(2)}$ has high costs and low correlation
- Variance cannot be improved by adding a third IS
- Any third IS will lead either to a violation of the feasibility condition or to a higher variance than the MC estimator ($S > 1$)
- Have to remove/change IS $f^{(2)}$ to reduce variance



Example: three information sources ($k = 3$)

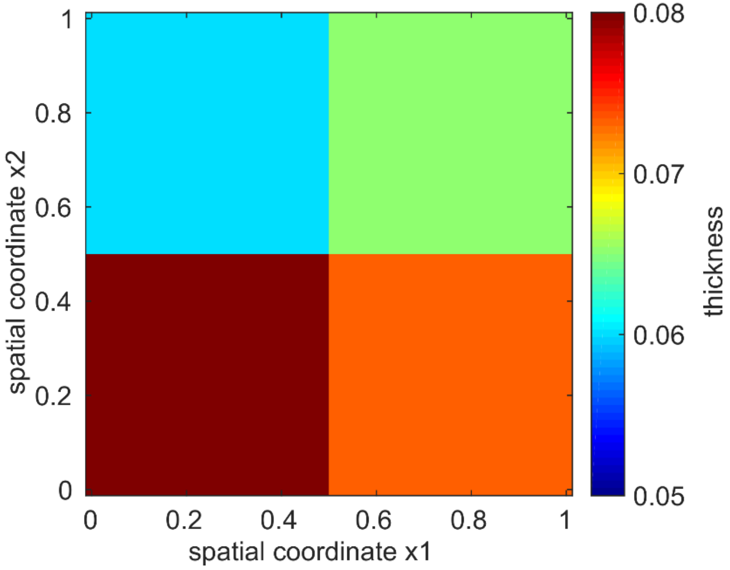
$$S(w_1, w_2, w_3, \rho_2, \rho_3) = \sqrt{1 - \rho_2^2} + \sqrt{\frac{w_2}{w_1}(\rho_2^2 - \rho_3^2)} + \sqrt{\frac{w_3}{w_1}\rho_3^2}$$

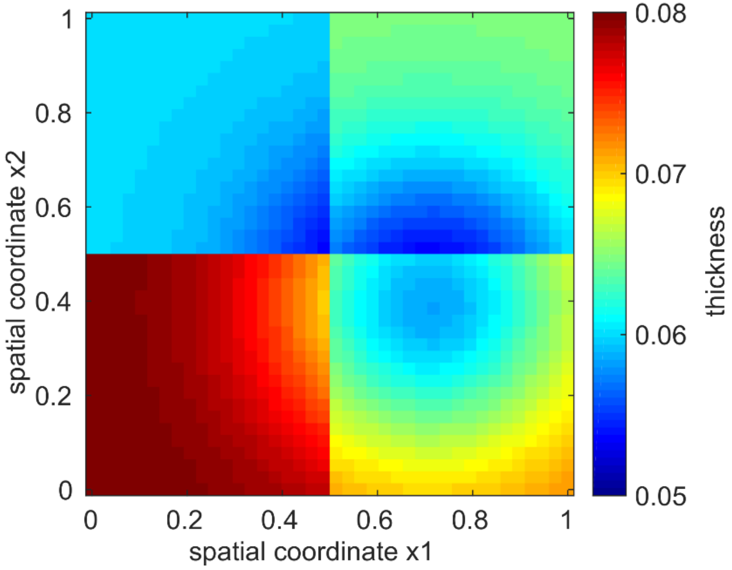
- Set $w_2/w_1 = 10^{-4}$, $\rho_2 = 0.4$
- Decreasing the costs of $f^{(2)}$ releases deadlock; adding a third model can improve the variance



Example: Locally damaged plate (multiple models)

- 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



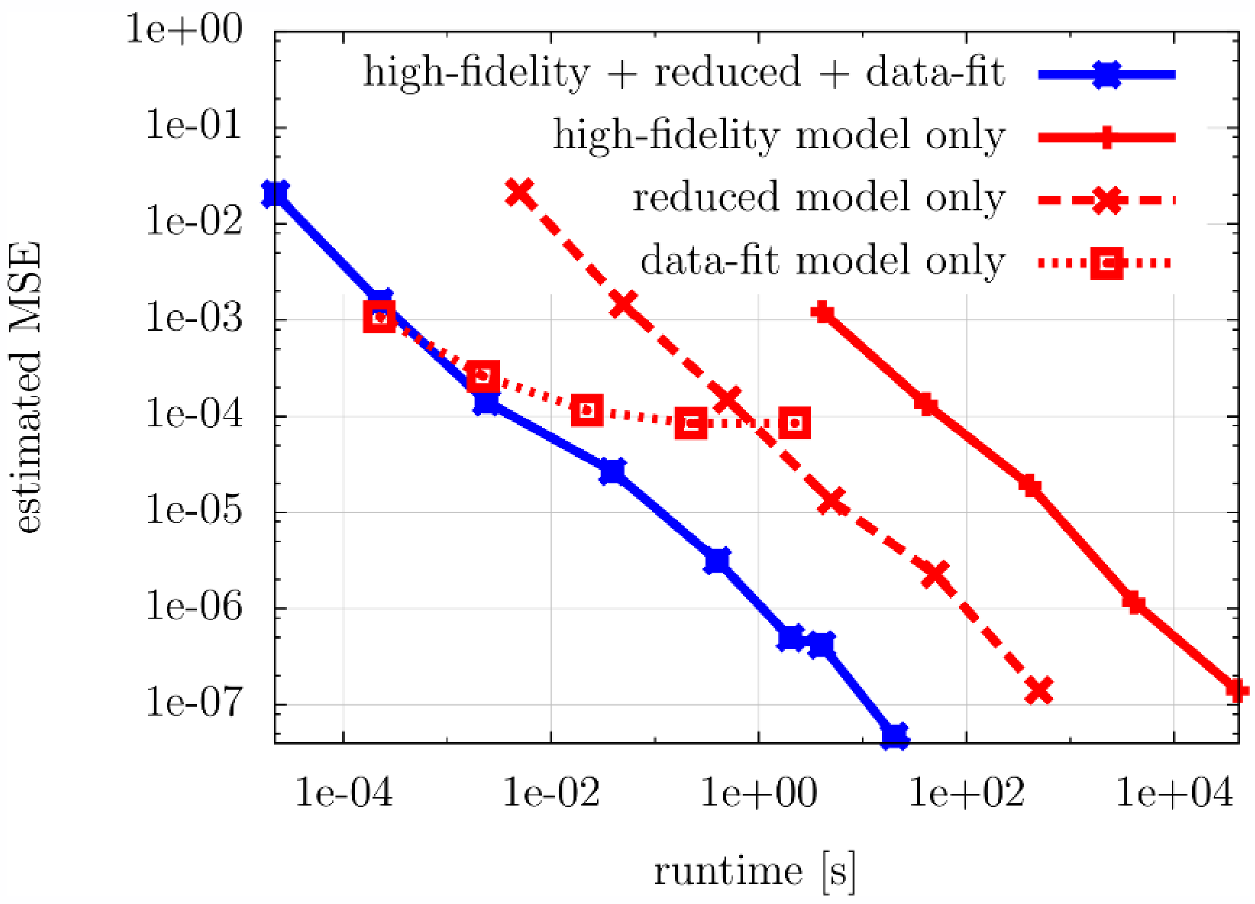


- Six models available
  - High-fidelity model: FEM, 300 dof
  - Reduced model: POD, 10 dof
  - Reduced model: POD, 2 dof
  - Reduced model: POD, 5 dof
  - Data-fit model: linear interpolation, 256 pts
  - Support vector machine: 256 pts

- 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



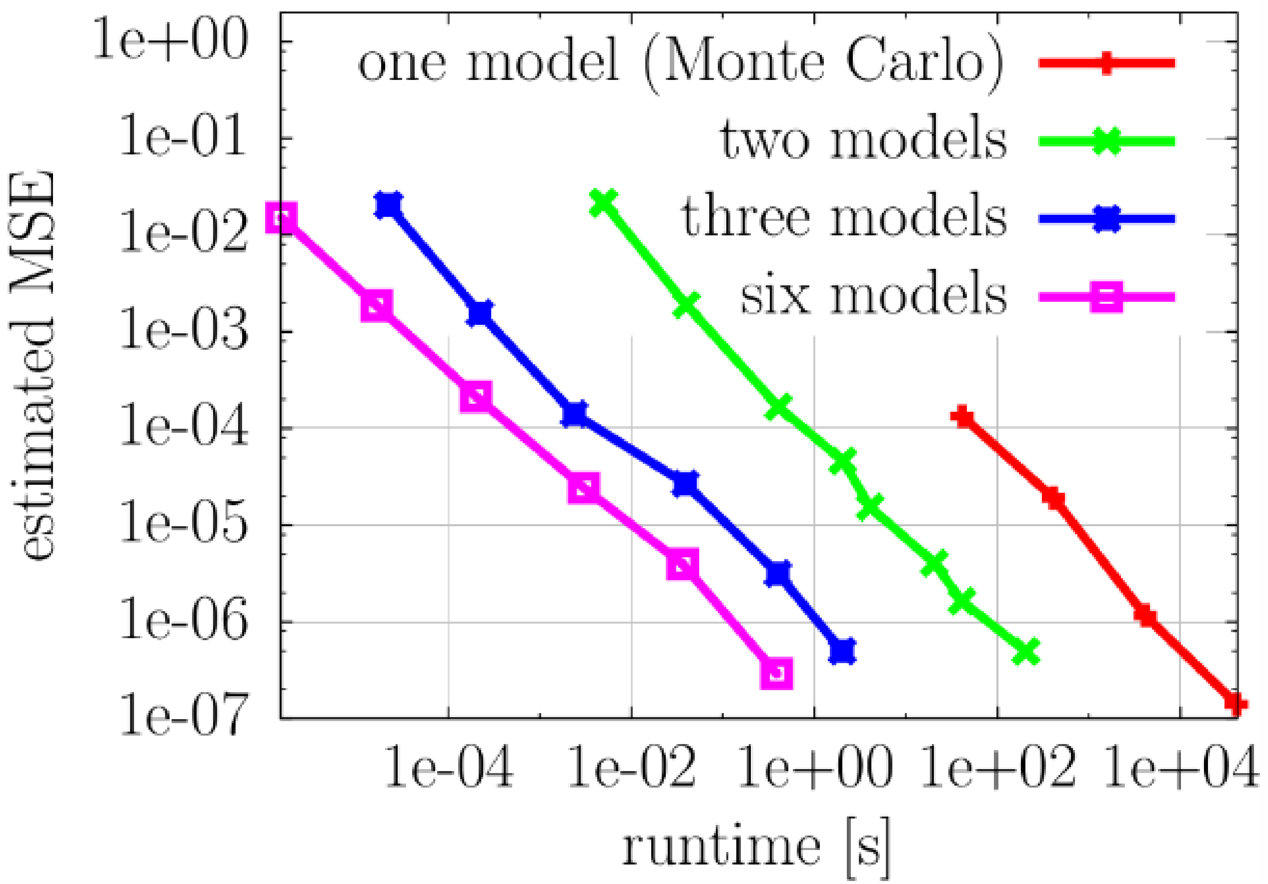
- Combine high-fidelity + reduced (POD, 10) + data-fit (linear interp, 256)
- Reduced and data-fit model lead to biased estimator, MFMC is unbiased

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.



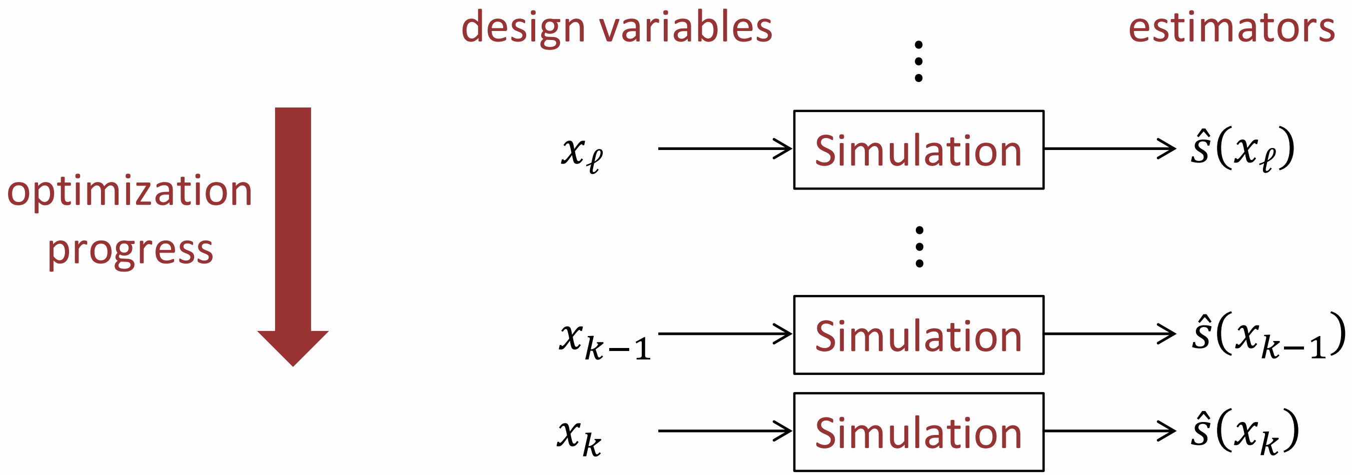
A broad view of multifidelity models

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

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

Cook, Jarrett, W., IJNME 2018

- 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

MFMC vs. MLMC

Multifidelity Monte Carlo

- control variate formulation
- applies to general models and information sources
- need to sample and learn the correlation between models
- optimal model allocations

Multilevel Monte Carlo [Heinrich 2001, Giles, 2008, Cliffe 2011]

- control variate formulation
- formulated for hierarchies of grids in specific problems (elliptic PDEs)
- leverage known error rates and cost rates

→ Clear connections between the two, especially as MLMC is expanded to consider a broader range of problems

→ Comparative analysis in Peherstorfer, Gunzburger, W., Convergence analysis of multifidelity Monte Carlo estimation, *Numerische Mathematik*, 2018.

5. Multifidelity Importance Sampling (MFIS)

Efficient estimation of low probability events, leveraging multiple models

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

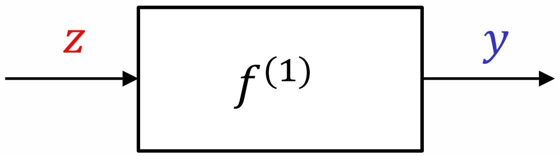
Peherstorfer, Kramer, W., **Combining multiple surrogate models to accelerate failure probability estimation with expensive high-fidelity models**, *J. Computational Physics*, 2017

Peherstorfer, Kramer, W. **Multifidelity preconditioning of the cross-entropy method for rare event simulation and failure probability estimation**. *SIAM/ASA J. Uncertainty Quantification*, 2018.

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")



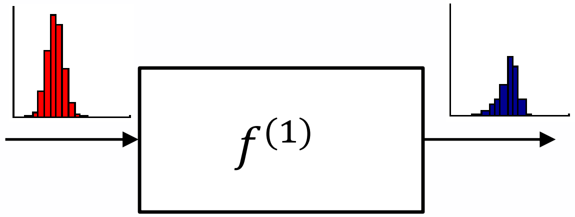
- define indicator function $I^{(1)}: \mathcal{Z} \to \mathcal{Y}$ as

$$I^{(1)}(z) = \begin{cases} 1, & \text{if } f^{(1)}(z) < 0 \\ 0, & \text{else} \end{cases}$$ ← 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 $z_1, \ldots, z_N$:

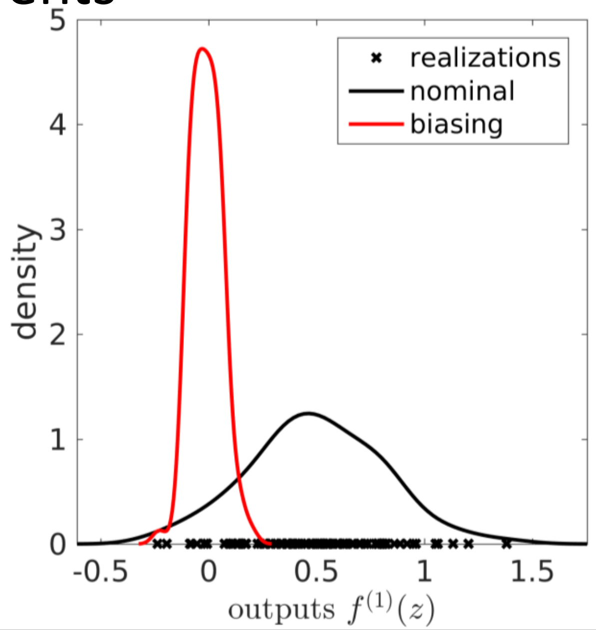
$$P_f^{\mathrm{MC}} = \frac{1}{N} \sum_{i=1}^{N} I^{(1)}(z_i)$$



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

$$P_f = \mathbb{E}_p\left[I^{(1)}(\mathrm{Z})\right] = \mathbb{E}_{q}\left[I^{(1)}(\hat{Z})w(\hat{Z})\right]$$

- Goal: construct a biasing density q such that

$$\mathrm{Var}_{q}\left[I^{(1)}(\hat{Z})w(\hat{Z})\right] < \mathrm{Var}_p\left[I^{(1)}(\mathrm{Z})\right]$$

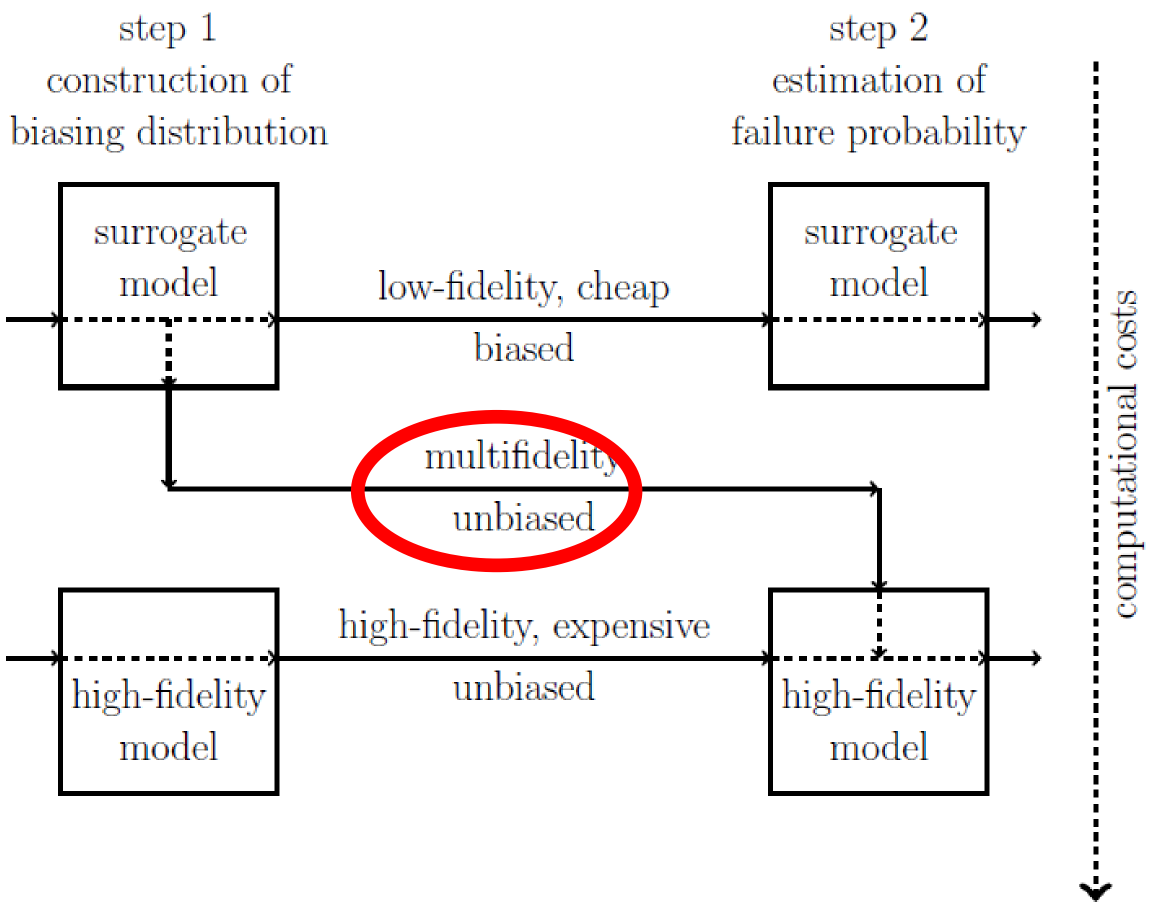
- Lower variance means fewer realizations of $\hat{Z}$ than of Z are necessary to achieve the same MSE $\longrightarrow$ fewer model evaluations

Multifidelity importance sampling (MFIS)

with two models

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

- We derive biasing distribution 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

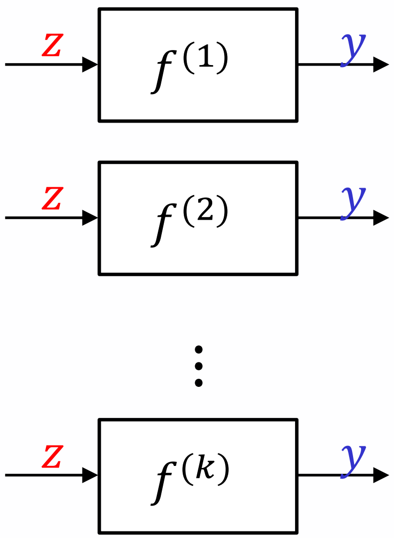
$$P_f = \mathbb{E}_q[P_f^{\mathrm{MFIS}}]$$

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]

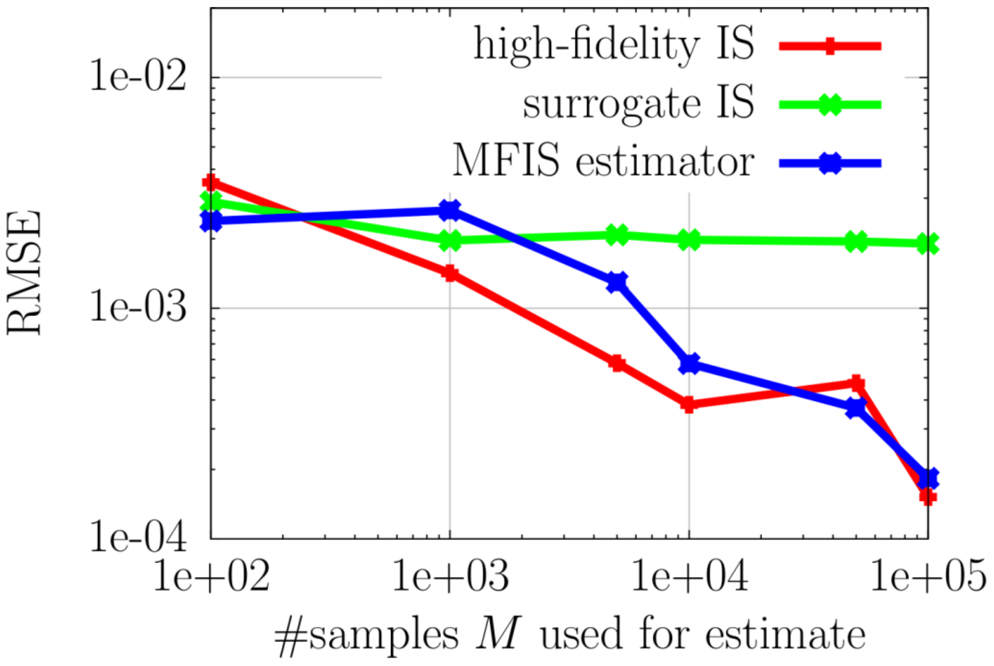
- Known that

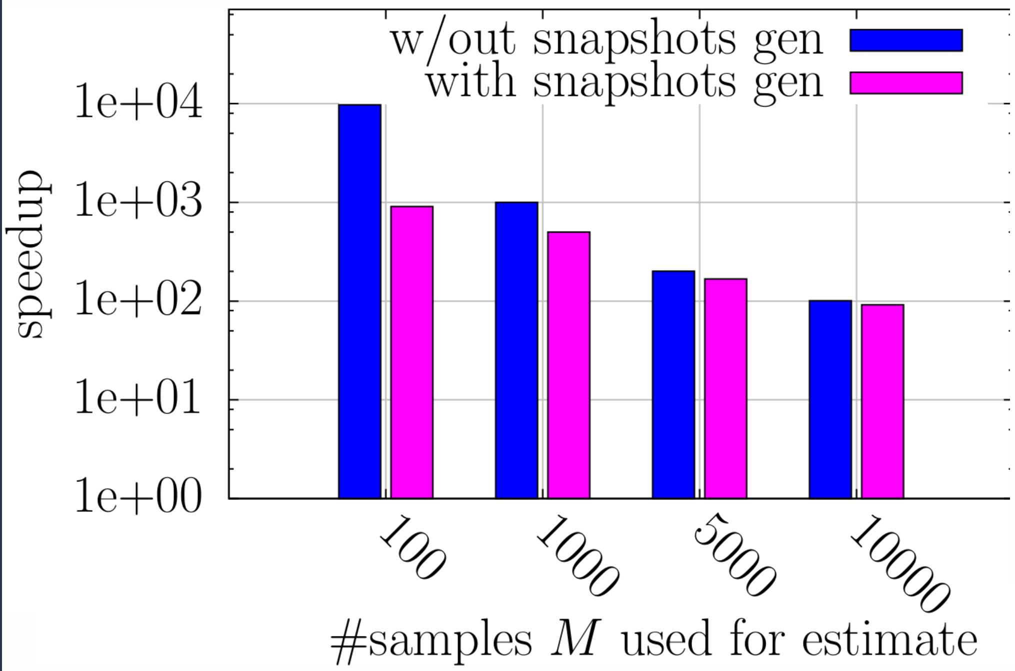
$$\frac{\mathrm{Var}\left[P_f^{\mathrm{Mixed}}\right]}{k\ -1} \leq \min_{i=2,\dots,k} \mathrm{Var}\left[I^{(1)} \frac{p}{q_i}\right]$$

- Our P_f^{Mixed} is up to factor $k-1$ as good as using the surrogate that minimizes variance

Locally damaged plate: MFIS

Estimate the probability that the deflection exceeds a critical value



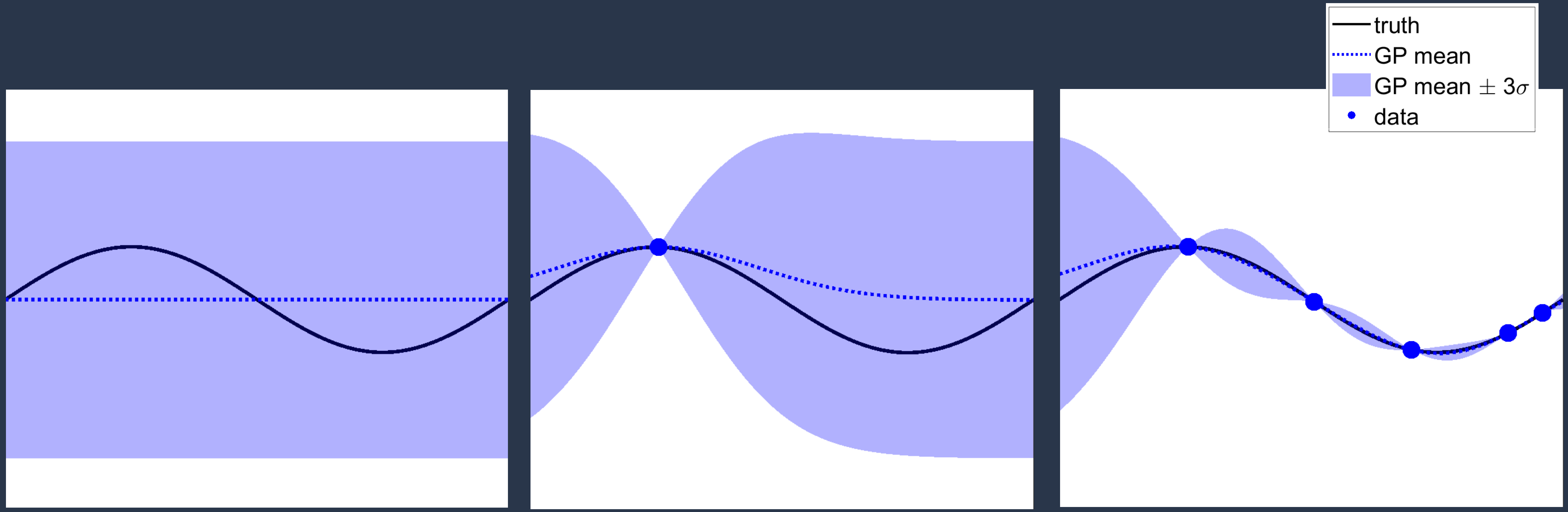


- Biasing density constructed from $N = 10^6$ realizations
- Using surrogate only leads to large bias
- MFIS leads to unbiased estimate of P_f
- If ROM available, speedup of up to 10^4, cf. high-fidelity

6. Bayesian Optimization and Multi-Information Source Optimization

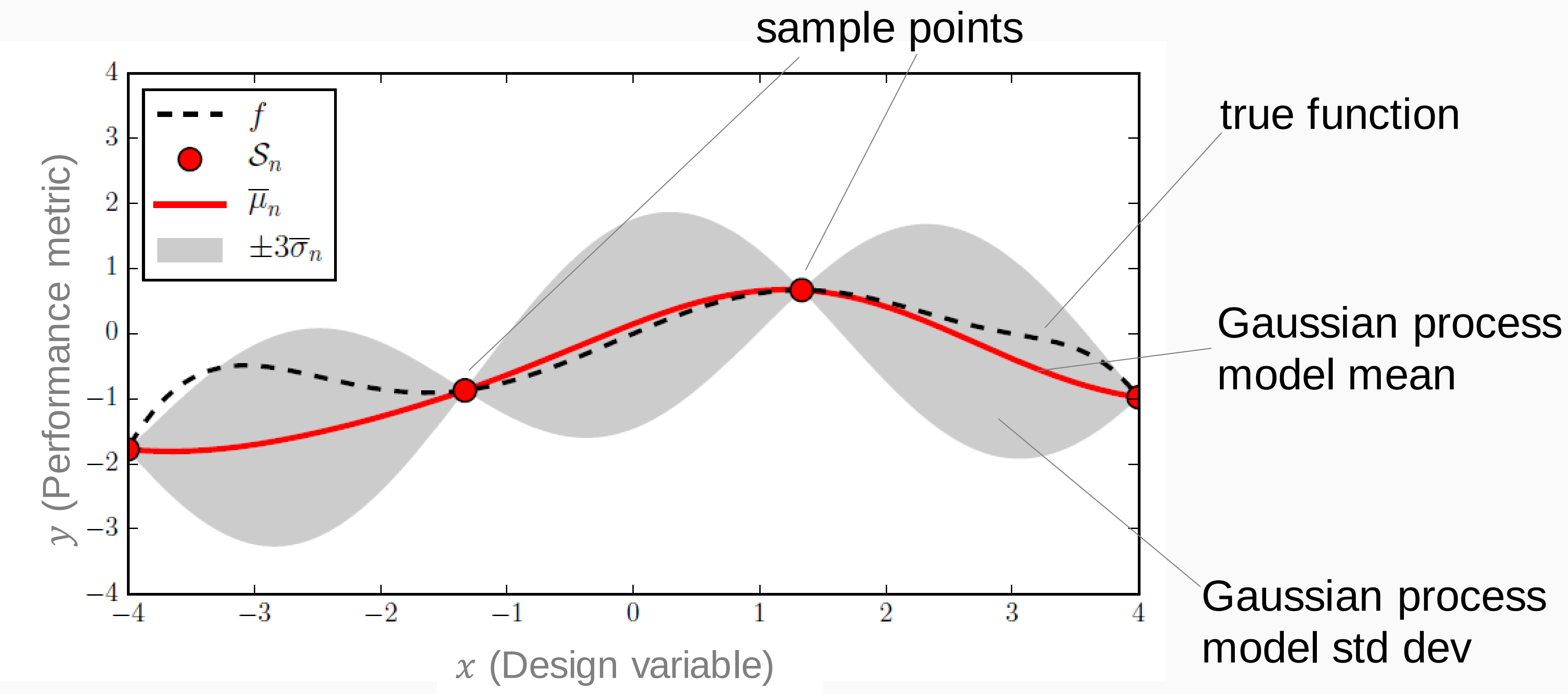
Optimal management of multiple sources of information

Bayesian optimization is a powerful and flexible foundation for multifidelity modeling



Components of Bayesian optimization

1. **Statistical model**
   uses data and Bayes' Theorem to compute
   posterior distribution



2. **Optimization via a value of information analysis**
   uses the posterior distribution and an acquisition
   function to decide what data to obtain next

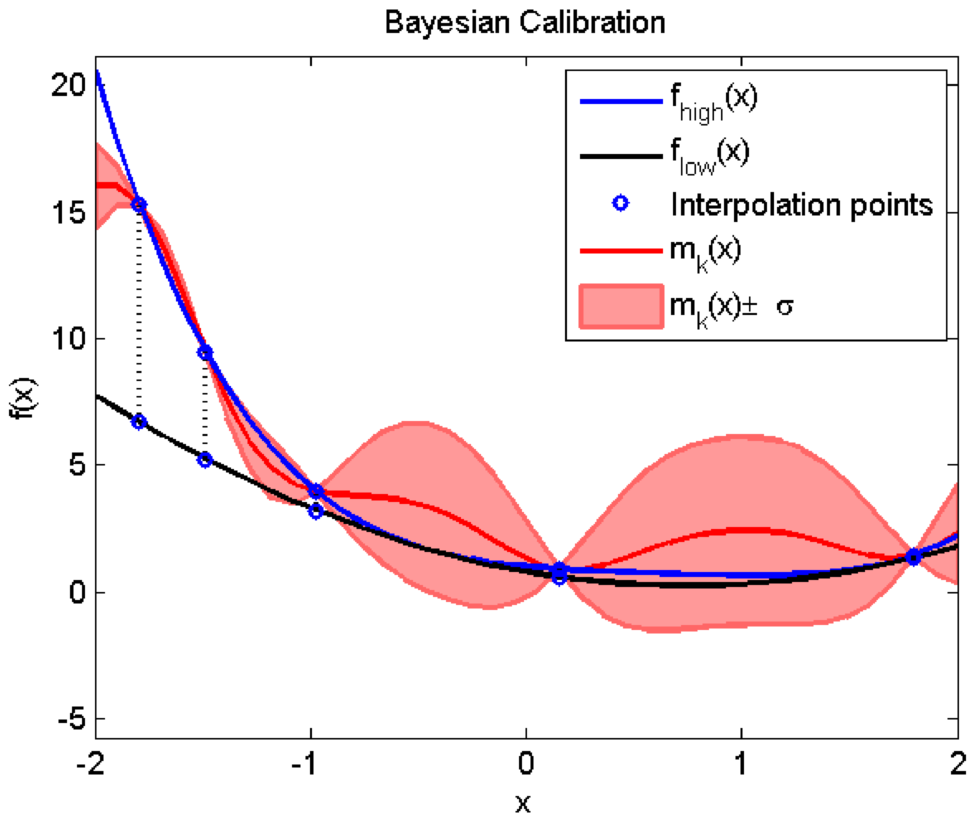
Bayesian Model Calibration

- Instead of modeling $f_{high}(\mathbf{x})$ directly using a Gaussian process, use a Gaussian process to capture the error between $f_{high}(\mathbf{x})$ and $f_{low}(\mathbf{x})$ (Kennedy and O'Hagan, 2000)

$$G(\mathbf{x}) = f_{\text{low}}(\mathbf{x}) + \mu + \beta^T \mathbf{x} + Z(\mathbf{x})$$

- Low-fidelity information can greatly improve convergence rates

- Model calibration also useful when high-fidelity analysis is black box – only need sample information



Efficient Global Optimization (Jones, 1998)

- Define f_{min} as the lowest observed value of the high-fidelity function

$$f_{\min} = \min_{j=1,\ldots,M} f_{\text{high}}(\mathbf{x}^j)$$

- Improvement would be finding a value of f_{high} that is lower than f_{min}

- Using Gaussian process model, improvement at point x is

$$I(x) = \max(f_{min} - G(x), 0)$$

- Maximum expected improvement is

$$E\left[I(\mathbf{x})\right] = \left(f_{\min} - m(\mathbf{x})\right) P\left(\frac{f_{\min} - m(\mathbf{x})}{\sigma(\mathbf{x})}\right) + \sigma p\left(\frac{f_{\min} - m(\mathbf{x})}{\sigma(\mathbf{x})}\right)$$

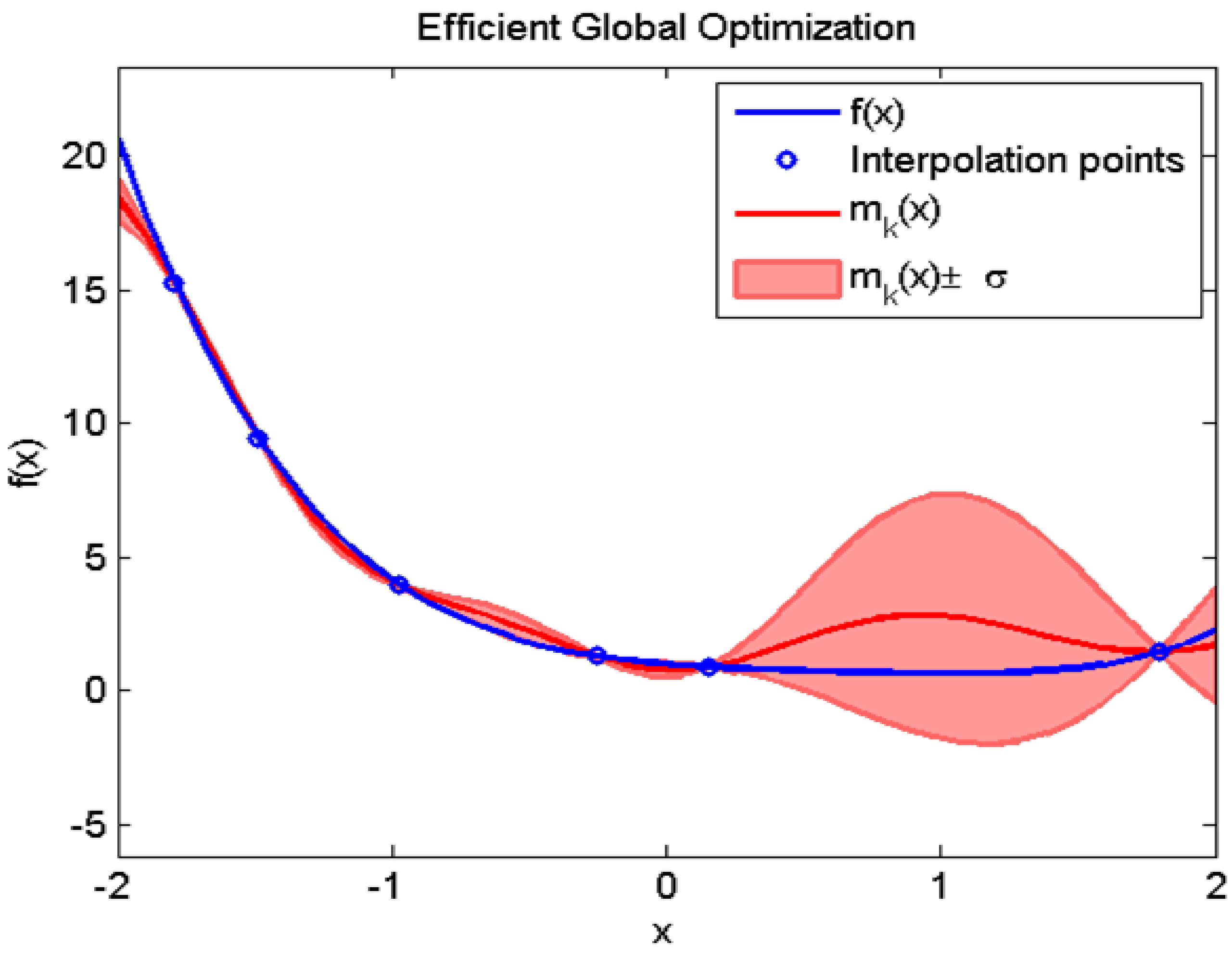
where $P(\cdot)$ is the standard normal cdf and p(·) is the standard normal pdf

EGO Algorithm

EGO algorithm from Jones (1998):

- Sample the high-fidelity function at *M* design points
- Find the point that globally maximizes the expected improvement in $f_{high}(\mathbf{x})$
- **do**
  - Calculate $f_{high}(\mathbf{x})$ at the point maximizing the expected improvement and construct a new Gaussian process model
  - (Optional) Validate the new Gaussian process model using cross-validation
  - Find the point that globally maximizes the expected improvement in $f_{high}(\mathbf{x})$

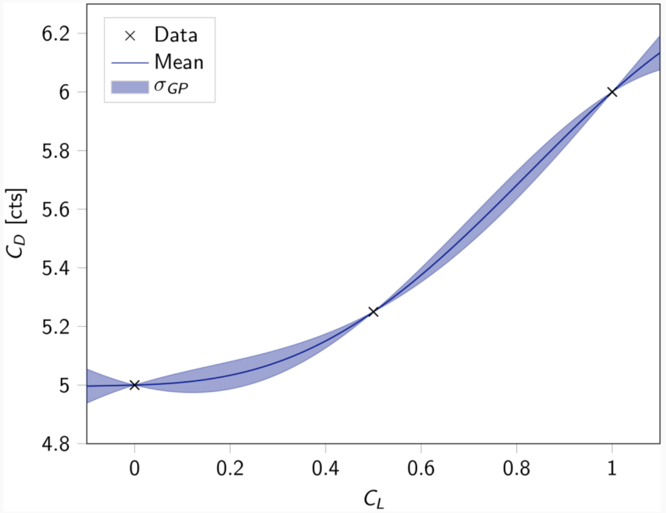
while the maximum expected improvement is larger than a tolerance



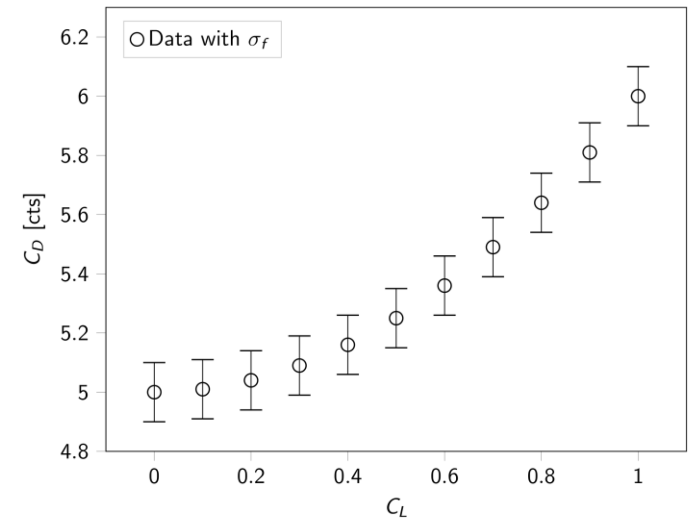
Multifidelity surrogates can also embed estimation of information source fidelity

A particularly important source of uncertainty in early-stage design

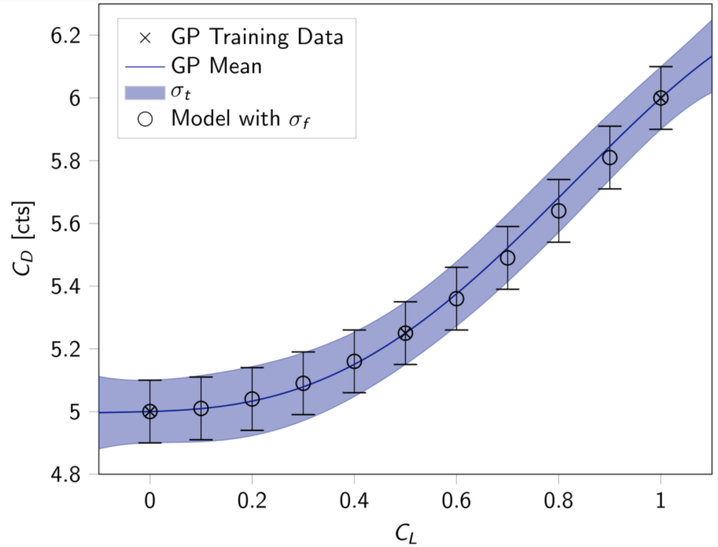
[Lam, Allaire, Willcox, AIAA 2015]



A standard Gaussian process model encodes **uncertainty due to training point locations**



Our **fidelity function** encodes confidence in underlying info source (model discrepancy)



Our **multifidelity surrogate**

- combines training uncertainty and fidelity

$$\sigma_t = \sqrt{\sigma_{GP}^2 + \sigma_f^2}$$

- fuses information from multiple sources

Multi-information source optimization with general model discrepancies

misoKG samples to maximize expected information gain per unit cost

[Poloczek, Wang, Frazier, NIPS 2016]

Two kinds of uncertainty

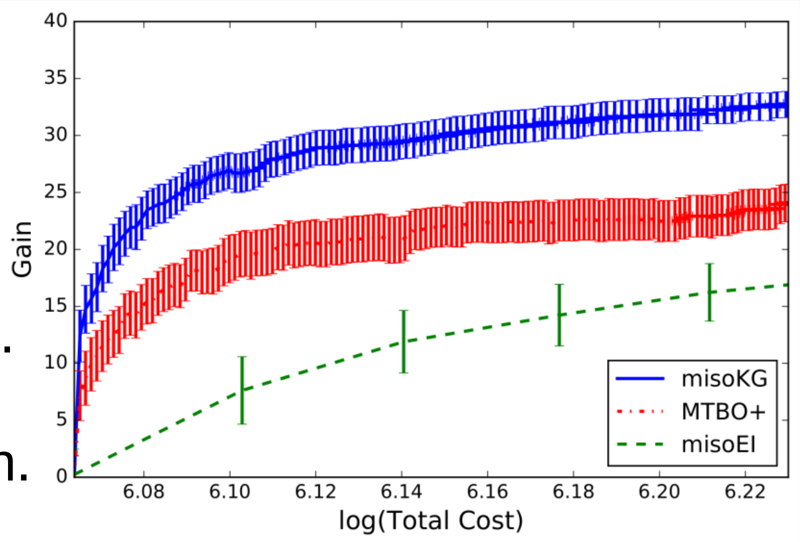
I) **Model discrepancy**
The internal model of each information source deviates from reality. This results in an *unknown bias*, e.g., due to an incomplete implementation of physical laws.

II) **Noise and numerical errors**
when observing the output of an information source.

→ More general than traditional “multifidelity”; no hierarchy that limits correlations. Indeed, misoKG exploits correlations among information sources.

- **GP regression:** a unified treatment of the various (latent) quantities
- **Value of information analysis** to decide **optimally** in a single, coordinated step what information source to query and for what design. Balances expected information gain and cost for every sampling decision.

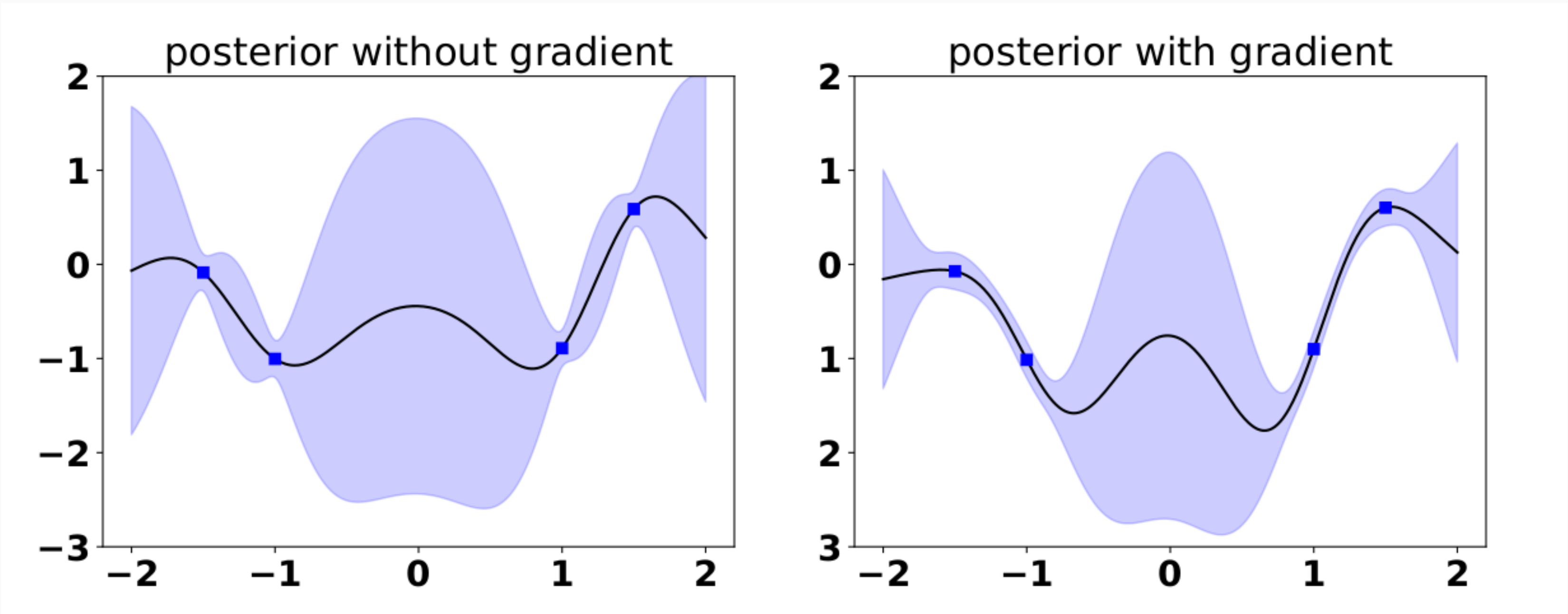
Assemble-To-Order (8d)



Gaussian process regression naturally accommodates derivative observations

Exploiting availability of adjoints gives better scalability

[Wu, Poloczek, Wilson, Frazier, NIPS 2017]



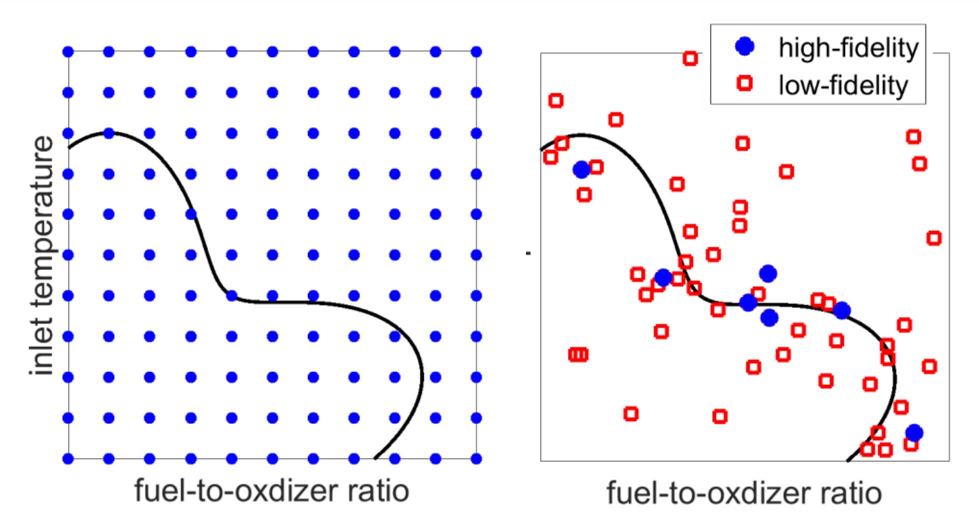
- Value of information analysis must account for the expected information gain and the cost of sampling both $f(x)$ and $\nabla f(x)$
- Derivative observations can be incomplete and/or noisy

Multifidelity stability boundary characterization

- Combines many cheap ROM evaluations and few expensive FOM evaluations
- New Bayesian optimization formulation that adaptively chooses when and where in parameter space to evaluate a ROM

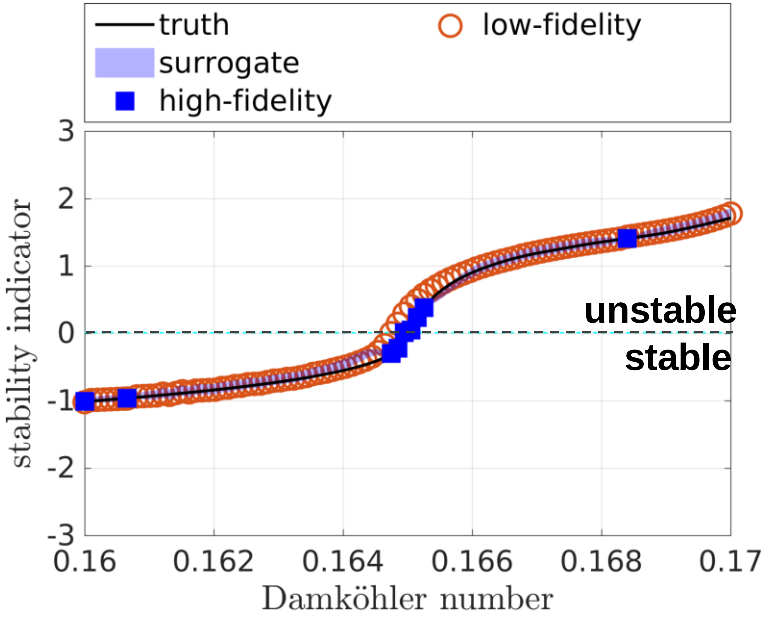
[Marques, Lam, W. NIPS 2018]

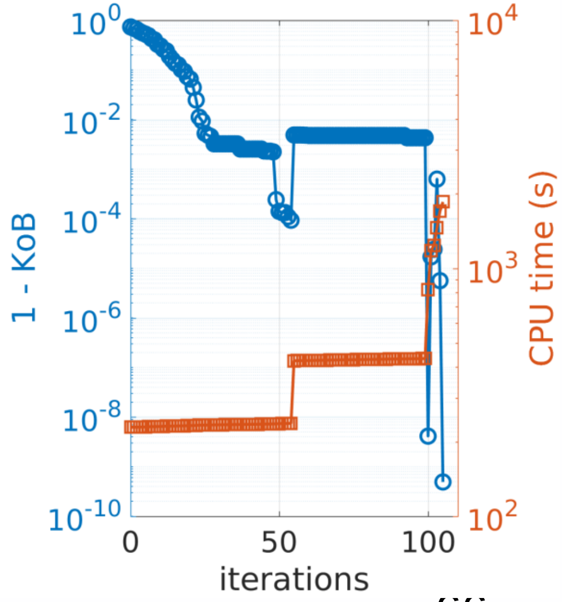
“Brute force” grid search **Multi-fidelity approach**



- Rigorous multi-fidelity framework combines accuracy of HFM with computational efficiency of ROM
- Use Bayesian optimization formulation: Gaussian process model + new acquisition function to drive adaptive sampling based on Knowledge of Boundary (KoB)

Tubular reactor test problem (LCO instability): algorithm explores using ROM + handful of carefully placed HFM evaluations





7. Conclusions

Multifidelity strategies for the outer loop

Leverage approximate models

but maintain guarantees on outer-loop result

- Multifidelity optimization (MFO) via trust region model management
- Multifidelity Monte Carlo (MFMC): a control variate formulation for estimating means
- Multifidelity Importance Sampling (MFIS): an importance sampling formulation for estimating probabilities
- Bayesian optimization and Multi-Information Source Management (MISO)
- Many more topics to cover!
  - Many methods to construct multifidelity surrogates
  - Bayesian optimization in a variety of settings
  - MFMC for global sensitivity analysis
  - Optimization under uncertainty (control variate, importance sampling formulations)
  - Adaptive delayed acceptance MCMC

References

N. M. Alexandrov, J. E. D. Jr, R. M. Lewis, and V. Torczon. A trust-region framework for managing the use of approximation models in optimization. *Structural Optimization*, 15(1):16-23, 1998.

N. M. Alexandrov, R. M. Lewis, C. R. Gumbert, L. L. Green, and P. A. Newman. Approximation and model management in aerodynamic optimization with variable-fidelity models. *Journal of Aircraft*, 38(6):1093-1101, 2001.

E. Arian, M. Fahl, and E. Sachs. Trust-region proper orthogonal decomposition models by optimization methods. In Proceedings of the 41st IEEE Conference on Decision and Control, pp. 3300-3305, Las Vegas, NV, 2002.

J. F. M. Barthelemy and R. T. Haftka. Approximation concepts for optimum structural design: a review. *Structural Optimization*, 5(3):129-144, 1993.

A. Basudhar and S. Missoum. An improved adaptive sampling scheme for the construction of explicit boundaries. *Structural and Multidisciplinary Optimization*, 42(4):517{529, 2010.

A. J. Booker, J. E. Dennis, P. D. Frank, D. B. Serani, V. Torczon, and M. W. Trosset. A rigorous framework for optimization of expensive functions by surrogates. *Structural Optimization*, 17(1):1-13, 1999.

S. Boyaval. A fast Monte Carlo method with a reduced basis of control variates applied to uncertainty propagation and Bayesian estimation. *Computer Methods in Applied Mechanics and Engineering*, 2012.

S. Boyaval and T. Lelievre. A variance reduction method for parametrized stochastic differential equations using the reduced basis paradigm. *Communications in Mathematical Sciences*, 8(3):735-762, 2010.

Chaudhuri, A., Lam, R. and Willcox, K., Multifidelity uncertainty propagation via adaptive surrogates in coupled multidisciplinary systems. *AIAA Journal*, Vol. 56, No. 1, pp. 235-249, 2018.

S. Choi, J. J. Alonso, I. M. Kroo, and M. Wintzer. Multifidelity Design Optimization of Low-Boom Supersonic Jets. *Journal of Aircraft*, 45(1):106-118, Jan. 2008.

References

J. A. Christen and C. Fox. Markov chain Monte Carlo using an approximation. *Journal of Computational and Graphical Statistics*, 14(4):795-810, 2005.

K. A. Clie, M. B. Giles, R. Scheichl, and A. L. Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Computing and Visualization in Science*, 14(1):3-15, 2011.

A. Conn, N. Gould, and P. Toint. *Trust Region Methods*. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics, Philadelphia, PA.

Conn, A., Scheinberg, K., and Vicente, L. (2009). Global convergence of general derivative-free trust-region algorithms to first- and second-order critical points. *SIAM Journal of Optimization*, 20(1):387-415.

Cook, L., Jarrett, J. and Willcox, K., Generalized information reuse for optimization under uncertainty with non-sample average estimators, *International Journal for Numerical Methods in Engineering*, Volume 115, Issue 12, pp. 1457-1476, September 2018.

T. Cui, C. Fox, and M. J. O'Sullivan. Bayesian calibration of a large-scale geothermal reservoir model by a new adaptive delayed acceptance Metropolis Hastings algorithm. *Water Resources Research*, 47(10):1-26, 2011.

T. Cui, Y. M. Marzouk, and K. Willcox. Data-driven model reduction for the Bayesian solution of inverse problems. *International Journal for Numerical Methods in Engineering*, 102(5):966-990, 2015.

C. Dribusch, S. Missoum, and P. Beran. A multifidelity approach for the construction of explicit decision boundaries: application to aeroelasticity. *Structural and Multidisciplinary Optimization*, 42(5):693-705, Nov. 2010.

Eldred, M. and Dunlavy, D. (2006). Formulations for surrogate-based optimization with data-fit, multifidelity, and reduced-order models. AIAA-2006-7117, 11th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Portsmouth, Virginia.

Eldred, M., Giunta, A., and Collis, S. (2004). Second-order corrections for surrogate-based optimization with model hierarchies. In Proceedings of the 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference. AIAA 2004-4457.

References

M. S. Eldred, L. W. T. Ng, M. F. Barone, and S. P. Domino. Multifidelity uncertainty quantification using spectral stochastic discrepancy models. In R. Ghanem, D. Higdon, and H. Owhadi, editors, *Handbook of Uncertainty Quantification*, pages 1-45, 2016.

D. Higdon, H. Lee, and Z. Bi. A Bayesian approach to characterizing uncertainty in inverse problems using coarse and fine-scale information. *IEEE Transactions on Signal Processing*, 50(2):389-399, 2002.

Jones, D., “A taxonomy of global optimization methods based on response surfaces,” *Journal of Global Optimization*, 21, 345-383, 2001.

Jones, D., Schonlau, M., and Welch, W. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13:455-492.

Keane, A. J. (2003). Wing optimization using design of experiment, response surface, and data fusion methods. *Journal of Aircraft*, 40(4):741-750.

Kennedy, M. and O'Hagan, A. (2001). Bayesian calibration of computer models. *Journal of the Royal Statistical Society*, 63(2):425-464.

March, A. and Willcox, K., "Convergent multifidelity optimization using Bayesian model calibration." *Structural and Multidisciplinary Optimization*, Vol. 46, No.1, 2012, pp. 93-109.

March, A. and Willcox, K., “A Provably Convergent Multifidelity Optimization Algorithm not Requiring High-Fidelity Derivatives,” *AIAA Journal*, Vol. 50, No. 5, 2012, pp.1079-1089.

L. W. Ng and K. Willcox. Multifidelity approaches for optimization under uncertainty. *International Journal for Numerical Methods in Engineering*, 100(10):746-772, 2014.

L. W. Ng and K. Willcox. Monte-Carlo information-reuse approach to aircraft conceptual design optimization under uncertainty. *Journal of Aircraft*, 53(2):427-438, 2016.

References

Peherstorfer, B., Willcox, K. and Gunzburger, M. Optimal model management for multifidelity Monte Carlo estimation. *SIAM Journal on Scientific Computing*, 38(5):A3163-A3194, 2016.

Peherstorfer, B., Kramer, B. and Willcox, K., Multifidelity preconditioning of the cross-entropy method for rare event simulation and failure probability estimation. *SIAM/ASA Journal on Uncertainty Quantification*, Vol. 6, No. 2, pp. 737-761, 2018.

Peherstorfer, B., Gunzburger, M. and Willcox, K., Convergence analysis of multifidelity Monte Carlo estimation, *Numerische Mathematik*, 139(3):683-707, 2018.

Peherstorfer, B., Willcox, K. and Gunzburger, M., Survey of multifidelity methods in uncertainty propagation, inference, and optimization, *SIAM Review*, Vol. 60, No. 3, pp. 550-591, 2018.

Peherstorfer, B., Kramer, B. and Willcox, K., Combining multiple surrogate models to accelerate failure probability estimation with expensive high-fidelity models, *Journal of Computational Physics*, Vol. 341, pp. 61-75, 2017.

Peherstorfer, B., Cui, T., Marzouk, Y. and Willcox, K., Multifidelity importance sampling, *Computer Methods in Applied Mechanics and Engineering* , Vol. 300, pp. 490-509, 2016.

Perdikaris, P., Venturi, D., Royset, J.O. and Karniadakis, G.E. Multi-fidelity modelling via recursive co-kriging and Gaussian-Markov random fields. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 471(2179), 2015.

References

Qian, E., Peherstorfer, B., O'Malley, D., Vesselinov, V. and Willcox, K., Multifidelity Monte Carlo estimation of variance and sensitivity indices, *SIAM/ASA Journal on Uncertainty Quantification*, Vol. 6, No. 2, pp. 683-706, 2018.

Robinson, T., Eldred, M., Willcox, K., and Haimes, R. (2008). Surrogate-based optimization using multifidelity models with variable parameterization and corrected space mapping. *AIAA Journal*, 46(11):2814-2822.

Sacks, J., Welch, W., Mitchell, T., and Wynn, H. (1989). Design and analysis of computer experiments. *Statistical Science*, 4(4):409-435.

Sobieszczanski-Sobieski, J., Altus, T., Phillips, M., and Sandusky, R. (2003). Bi-level integrated system synthesis for concurrent and distributed processing. *AIAA Journal*, 41(10):1996-2003.

Wild, S. and Shoemaker, C. A. Global convergence of radial basis function trust region derivative-free algorithms. *SIAM Journal of Optimization*, 21(3):761-781, 2011.