Multifidelity importance sampling methods for rare event simulation

Benjamin Peherstorfer
University of Wisconsin-Madison

Karen Willcox and Boris Kramer
MIT

Max Gunzburger
Florida State University

July 2017

Introduction

High-fidelity model with costs $w_1 \gg 0$

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

Random variable Z, estimate

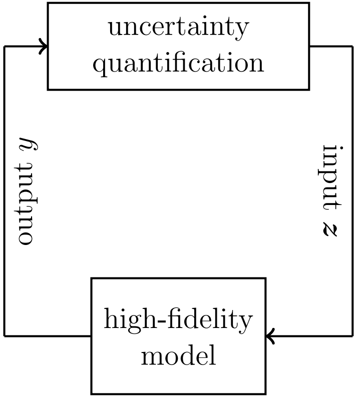
$$s = \mathbb{E}[f^{(1)}(Z)]$$

Monte Carlo estimator of s with samples $Z_1, \dots, Z_n$

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

Computational costs high

- Many evaluations of high-fidelity model
- Typically $10^3 - 10^6$ evaluations
- Intractable if $f^{(1)}$ expensive



Surrogate models

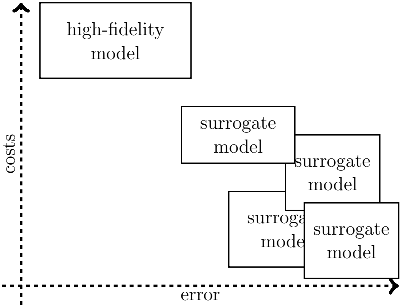
Given is a high-fidelity model $f^{(1)} : \mathcal{D} \to \mathcal{Y}$

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

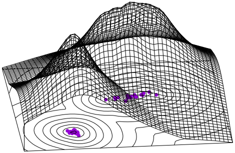
Additionally, often have surrogate models

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

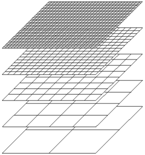
- Approximate high-fidelity $f^{(1)}$
- Often orders of magnitudes cheaper



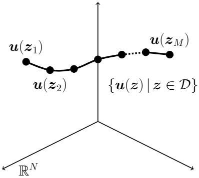
Examples of surrogate models



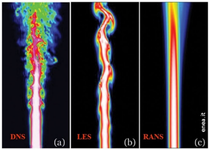
data-fit models, response surfaces, machine learning



coarse-grid approximations



reduced basis, proper orthogonal decomposition



simplified models, linearized models

Surrogate models in uncertainty quantification

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

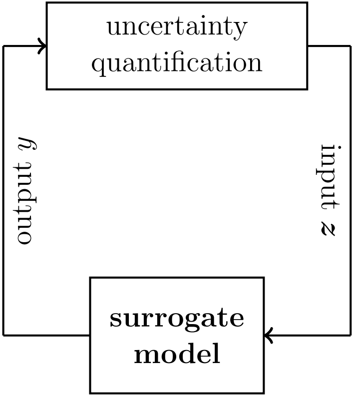
- Costs of uncertainty quantification reduced
- Often orders of magnitude speedups

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

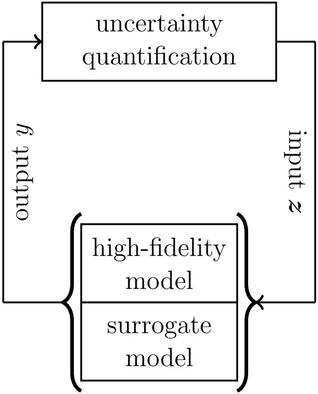
- Leverage surrogate models for speedup
- Recourse to high-fidelity for accuracy

Multifidelity speeds up computations

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

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



[P., Willcox, Gunzburger, ***Survey of multifidelity methods in uncertainty propagation, inference, and optimization***; available online as technical report, MIT, 2016]

Our approach: Multifidelity methods

Combine high-fidelity and surrogate models

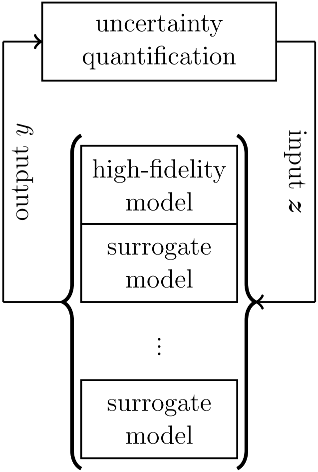
- Leverage surrogate models for speedup
- Recourse to high-fidelity for accuracy

Multifidelity speeds up computations

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

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



[P., Willcox, Gunzburger, ***Survey of multifidelity methods in uncertainty propagation, inference, and optimization***; available online as technical report, MIT, 2016]

Multifidelity Monte Carlo (MFMC)

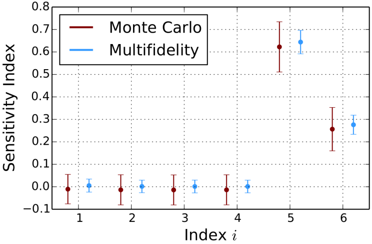
MFMC use control variates for var reduction

- Derives control variates from surrogates
- Uses number of samples that minimize error

[P., Willcox, Gunzburger: Optimal model management for multi-fidelity Monte Carlo estimation. SIAM Journal on Scientific Computing, 2016]

Multifidelity sensitivity analysis

- Identify the parameters of model with largest influence on quantity of interest
- Elizabeth Qian (MIT)/Earth Science (LANL)



Figures: Elizabeth Qian (MIT)

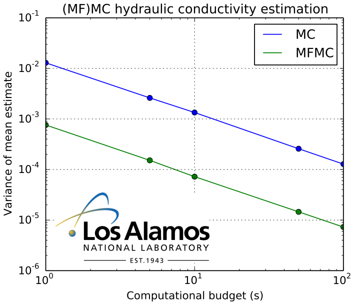
Asymptotic analysis of MFMC

[P., Gunzburger, Willcox: Convergence analysis of multifidelity Monte Carlo estimation, submitted. 2016]

MFMC with information reuse

[Ng, Willcox: Monte Carlo Information-Reuse Approach to Aircraft Conceptual Design Optimization Under Uncertainty. 2015]

MFMC with optimally-adapted surrogates



Multifidelity rare event simulation based on importance sampling

with Karen Willcox and Boris Kramer

MFCE: Problem setup

Threshold $0 < t \in \mathbb{R}$ and random variable

$$Z \sim p$$

Estimate rare event probability

$$P_f = \mathbb{P}_p[f^{(1)} \leq t]$$

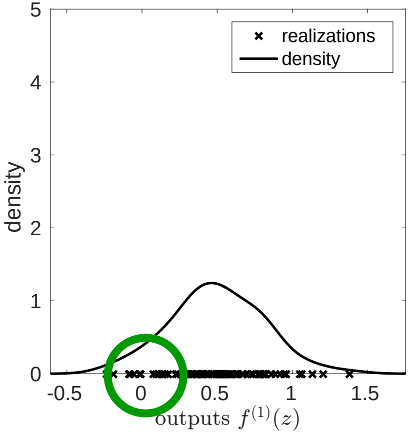
Can be reformulated to estimating $\mathbb{E}$

$$P_f = \mathbb{E}_p[\mathbb{I}_t^{(1)}]$$

with indicator function

$$\mathbb{I}_t^{(1)}(\boldsymbol{z}) = \begin{cases} 1, & f^{(1)}(\boldsymbol{z}) \leq t\,, \\ 0, & f^{(1)}(\boldsymbol{z}) > t \end{cases}$$

If $P_f \ll 1$, very unlikely that we hit $f^{(1)} \leq t$



MFCE: Importance sampling

Consider a biasing density q with

$$\mathrm{supp}(p) \subseteq \mathrm{supp}(q)$$

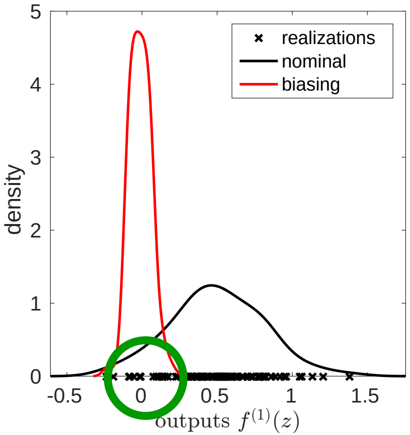
Reformulate estimation problem

$$P_f = \mathbb{E}_p[\mathbb{I}_t^{(1)}] = \mathbb{E}_q\left[\mathbb{I}_t^{(1)} \frac{p}{q}\right]$$

Goal is constructing suitable q with

$$\mathrm{Var}_q\left[\mathbb{I}_t^{(1)} \frac{p}{q}\right] \ll \mathrm{Var}_p[\mathbb{I}_t^{(1)}]$$

$\Rightarrow$ **Use surrogate models**



MFCE: Literature review

Two-fidelity approaches

- Switch between models [Li, Xiu et al., 2010, 2011, 2014]
- Reduced basis models with error estimators [Chen and Quarteroni, 2013]
- Kriging models and importance sampling [Dubourg et al., 2013]
- Subset method with machine-learning-based models [Bourinet et al., 2011], [Papadopoulos et al., 2012]
- Surrogates and importance sampling [P., Cui, Marzouk, Willcox, 2016]

Multilevel methods for rare event simulation

- Variance reduction via control variates [Elfverson et al., 204, 2016], [Fagerlund et al., 2016]
- Subset method with coarse-grid approximations [Ullmann and Papaioannou, 2015]

Combining multiple general types of surrogates

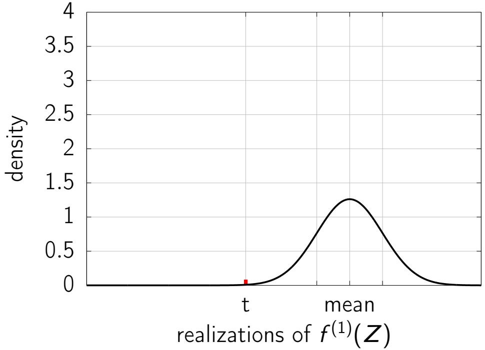
- Importance sampling + control variates [P., Kramer, Willcox, 2017]

MFCE: Direct sampling of surrogate models

Directly sampling surrogate models to construct biasing density

- Reduces costs per sample
- Number of samples to construct biasing density remains the same
- Works well for probabilities $> 10^{-5}$

$\Rightarrow$ Insufficient for *very* rare event probabilities in range of 10^{-9}



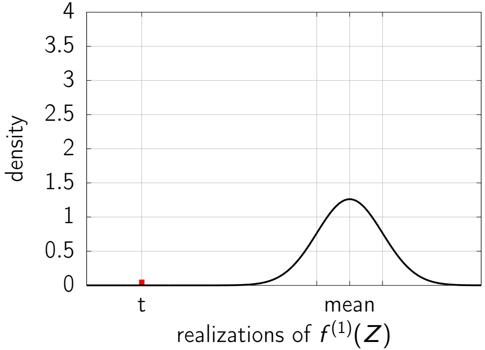
[P., Cui, Marzouk, Willcox, 2016], [P., Kramer, Willcox, 2017]

MFCE: Direct sampling of surrogate models

Directly sampling surrogate models to construct biasing density

- Reduces costs per sample
- Number of samples to construct biasing density remains the same
- Works well for probabilities $> 10^{-5}$

$\Rightarrow$ Insufficient for *very* rare event probabilities in range of 10^{-9}



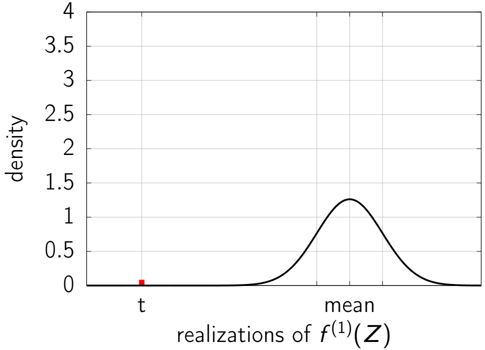
[P., Cui, Marzouk, Willcox, 2016], [P., Kramer, Willcox, 2017]

MFCE: Construct biasing density iteratively

Threshold t controls "rareness" of event

$$\mathbb{I}_t^{(1)}(\boldsymbol{z}) = \begin{cases} 1\,, & f^{(1)}(\boldsymbol{z}) \leq t\,, \\ 0\,, & f^{(1)}(\boldsymbol{z}) > t \end{cases}$$

Cross-entropy method constructs densities iteratively $t_1 > t_2 > \cdots > t$



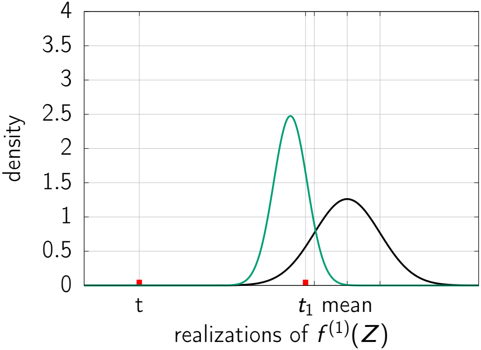
[Rubinstein, 1999], [Rubinstein, 2001]

MFCE: Construct biasing density iteratively

Threshold t controls "rareness" of event

$$\mathbb{I}_t^{(1)}(\boldsymbol{z}) = \begin{cases} 1\,, & f^{(1)}(\boldsymbol{z}) \leq t\,, \\ 0\,, & f^{(1)}(\boldsymbol{z}) > t \end{cases}$$

Cross-entropy method constructs densities iteratively $t_1 > t_2 > \cdots > t$



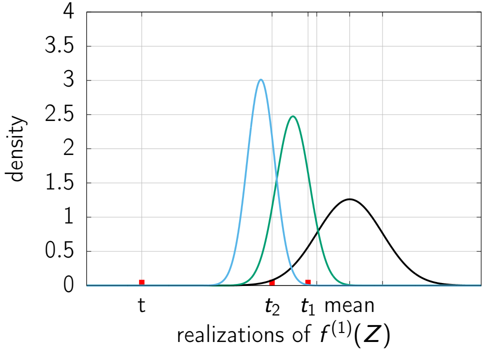
[Rubinstein, 1999], [Rubinstein, 2001]

MFCE: Construct biasing density iteratively

Threshold t controls "rareness" of event

$$\mathbb{I}_t^{(1)}(\boldsymbol{z}) = \begin{cases} 1\,, & f^{(1)}(\boldsymbol{z}) \leq t\,, \\ 0\,, & f^{(1)}(\boldsymbol{z}) > t \end{cases}$$

Cross-entropy method constructs densities iteratively $t_1 > t_2 > \cdots > t$



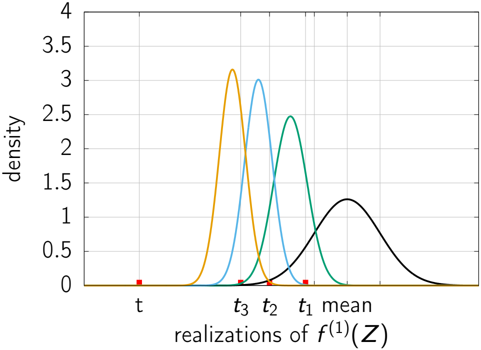
[Rubinstein, 1999], [Rubinstein, 2001]

MFCE: Construct biasing density iteratively

Threshold t controls "rareness" of event

$$\mathbb{I}_t^{(1)}(\boldsymbol{z}) = \begin{cases} 1\,, & f^{(1)}(\boldsymbol{z}) \leq t\,, \\ 0\,, & f^{(1)}(\boldsymbol{z}) > t \end{cases}$$

Cross-entropy method constructs densities iteratively $t_1 > t_2 > \cdots > t$



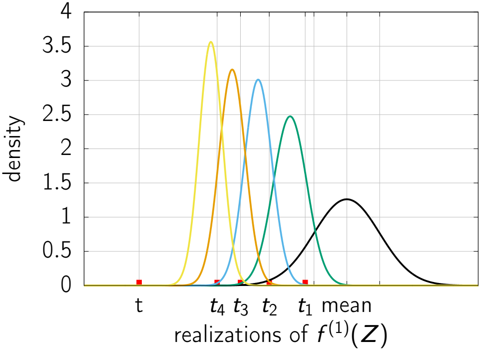
[Rubinstein, 1999], [Rubinstein, 2001]

MFCE: Construct biasing density iteratively

Threshold t controls "rareness" of event

$$\mathbb{I}_t^{(1)}(\boldsymbol{z}) = \begin{cases} 1\,, & f^{(1)}(\boldsymbol{z}) \leq t\,, \\ 0\,, & f^{(1)}(\boldsymbol{z}) > t \end{cases}$$

Cross-entropy method constructs densities iteratively $t_1 > t_2 > \cdots > t$



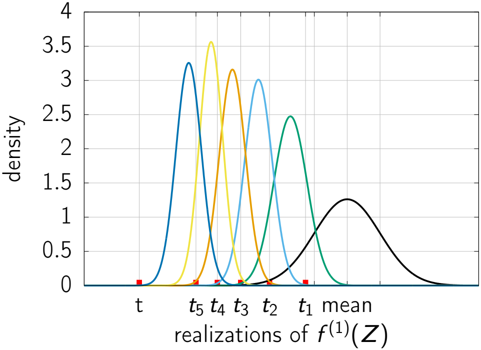
[Rubinstein, 1999], [Rubinstein, 2001]

MFCE: Construct biasing density iteratively

Threshold t controls "rareness" of event

$$\mathbb{I}_t^{(1)}(\boldsymbol{z}) = \begin{cases} 1\,, & f^{(1)}(\boldsymbol{z}) \leq t\,, \\ 0\,, & f^{(1)}(\boldsymbol{z}) > t \end{cases}$$

Cross-entropy method constructs densities iteratively $t_1 > t_2 > \cdots > t$



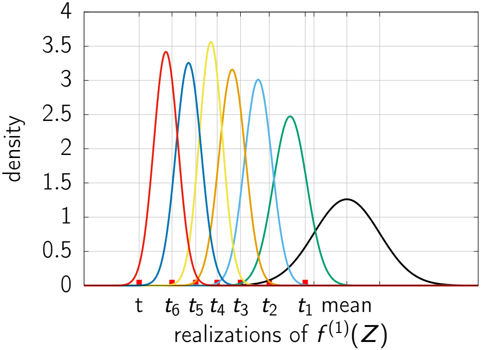
[Rubinstein, 1999], [Rubinstein, 2001]

MFCE: Construct biasing density iteratively

Threshold t controls "rareness" of event

$$\mathbb{I}_t^{(1)}(\boldsymbol{z}) = \begin{cases} 1\,, & f^{(1)}(\boldsymbol{z}) \leq t\,, \\ 0\,, & f^{(1)}(\boldsymbol{z}) > t \end{cases}$$

Cross-entropy method constructs densities iteratively $t_1 > t_2 > \cdots > t$



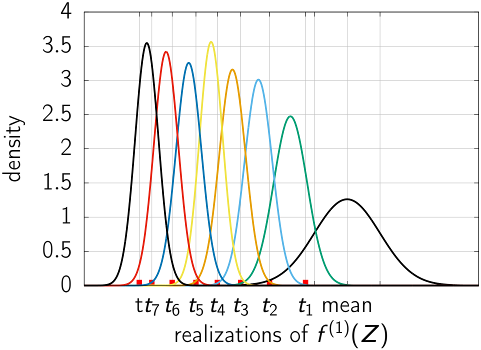
[Rubinstein, 1999], [Rubinstein, 2001]

MFCE: Construct biasing density iteratively

Threshold t controls "rareness" of event

$$\mathbb{I}_t^{(1)}(\boldsymbol{z}) = \begin{cases} 1\,, & f^{(1)}(\boldsymbol{z}) \leq t\,, \\ 0\,, & f^{(1)}(\boldsymbol{z}) > t \end{cases}$$

Cross-entropy method constructs densities iteratively $t_1 > t_2 > \cdots > t$



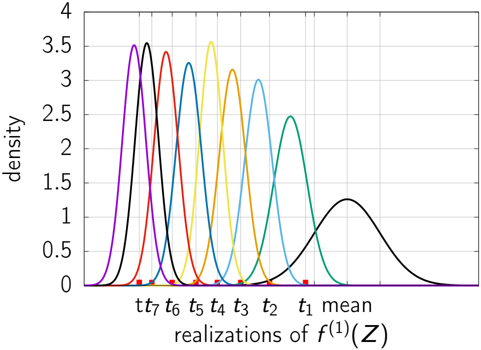
[Rubinstein, 1999], [Rubinstein, 2001]

MFCE: Construct biasing density iteratively

Threshold t controls "rareness" of event

$$\mathbb{I}_t^{(1)}(\boldsymbol{z}) = \begin{cases} 1\,, & f^{(1)}(\boldsymbol{z}) \leq t\,, \\ 0\,, & f^{(1)}(\boldsymbol{z}) > t \end{cases}$$

Cross-entropy method constructs densities iteratively $t_1 > t_2 > \cdots > t$



[Rubinstein, 1999], [Rubinstein, 2001]

MFCE: Cross-entropy method – in each step

Need to find biasing density in each step $i = 1, \dots, T$

- Optimal biasing density that reduces variance to 0

$$q_i(\boldsymbol{z}) \propto \mathbb{I}^{(1)}_{t_i}(\boldsymbol{z}) p(\boldsymbol{z})$$

$\Rightarrow$ Unknown normalizing constant (quantity we want to estimate)

- Find $q_{\boldsymbol{v}_i} \in \mathcal{Q} = \{q_{\boldsymbol{v}} : \boldsymbol{v} \in \mathcal{P}\}$ with min Kullback-Leibler distance to q_i

$$\min_{\boldsymbol{v}_i \in \mathcal{P}} D_{\mathrm{KL}}(q_i || q_{\boldsymbol{v}_i})$$

- Reformulate as (independent of normalizing constant of q_i)

$$\max_{\boldsymbol{v}_i \in \mathcal{P}} \mathbb{E}_p[\mathbb{I}^{(1)}_{t_i} \log(q_{\boldsymbol{v}_i})]$$

- Solve approximately by replacing $\mathbb{E}_p$ with Monte Carlo estimator

$$\max_{\boldsymbol{v}_i \in \mathcal{P}} \frac{1}{m} \sum_{i=1}^{m} \mathbb{I}^{(1)}_{t_i}(Z_i) \log(q_{\boldsymbol{v}_i}(Z_i)) \,, \qquad Z_1, \dots, Z_m \sim p$$

$\Rightarrow$ **Optimization problems affected by rareness of event** $\mathbb{I}^{(1)}_{t_i}(Z)$

[Rubinstein, 1999], [Rubinstein, 2001]

MFCE: Iterative cross-entropy procedure

Reuse $q_{\boldsymbol{v}_{i-1}}$ as biasing density for constructing $q_{\boldsymbol{v}_i}$

$$\max_{\boldsymbol{v}_i \in \mathcal{P}} \mathbb{E}_{q_{\boldsymbol{v}_{i-1}}}\left[\mathbb{I}^{(1)}_{t_i} \frac{p}{q_{\boldsymbol{v}_{i-1}}} \log(q_{\boldsymbol{v}_i})\right]$$

- Choose $t_1 \gg t$, solve for $\boldsymbol{v}_1 \in \mathcal{P}$ with

$$\max_{\boldsymbol{v}_1 \in \mathcal{P}} \mathbb{E}_p[\mathbb{I}^{(1)}_{t_1} \log(q_{\boldsymbol{v}_1})]$$

- Select $t_2 < t_1$, solve for $\boldsymbol{v}_2 \in \mathcal{P}$ with

$$\max_{\boldsymbol{v}_2 \in \mathcal{P}} \mathbb{E}_{q_{\boldsymbol{v}_1}}\left[\mathbb{I}^{(1)}_{t_2} \frac{p}{q_{\boldsymbol{v}_1}} \log(q_{\boldsymbol{v}_2})\right]$$

- Repeat until threshold t is reached and parameter $\boldsymbol{v}_*$ is obtained
- Reweighed optimization problems have same optimum as original problems

Once $q_{\boldsymbol{v}_*}$ has been found, cross-entropy estimator of P_f is

$$\hat{P}^{\mathrm{CE}}_t = \frac{1}{m} \sum_{i=1}^{m} \mathbb{I}^{(1)}_t(Z^*_i) \frac{p(Z^*_i)}{q_{\boldsymbol{v}_*}(Z^*_i)}, \qquad Z^*_1, \dots, Z^*_m \sim q_{\boldsymbol{v}_*}$$

MFCE: Iterative cross-entropy procedure

Reuse $q_{\boldsymbol{v}_{i-1}}$ as biasing density for constructing $q_{\boldsymbol{v}_i}$

$$\max_{\boldsymbol{v}_i \in \mathcal{P}} \mathbb{E}_{q_{\boldsymbol{v}_{i-1}}}\left[\mathbb{I}^{(1)}_{t_i} \frac{p}{q_{\boldsymbol{v}_{i-1}}} \log(q_{\boldsymbol{v}_i})\right]$$

- Choose $t_1 \gg t$, solve for $\boldsymbol{v}_1 \in \mathcal{P}$ with

$$\max_{\boldsymbol{v}_1 \in \mathcal{P}} \mathbb{E}_p[\mathbb{I}^{(1)}_{t_1} \log(q_{\boldsymbol{v}_1})]$$

- Select $t_2 < t_1$, solve for $\boldsymbol{v}_2 \in \mathcal{P}$ with

$$\max_{\boldsymbol{v}_2 \in \mathcal{P}} \mathbb{E}_{q_{\boldsymbol{v}_1}}\left[\mathbb{I}^{(1)}_{t_2} \frac{p}{q_{\boldsymbol{v}_1}} \log(q_{\boldsymbol{v}_2})\right]$$

- Repeat until threshold t is reached and parameter $\boldsymbol{v}_*$ is obtained
- Reweighed optimization problems have same optimum as original problems

Once $q_{\boldsymbol{v}_*}$ has been found, cross-entropy estimator of P_f is

$$\hat{P}^{\text{CE}}_t = \frac{1}{m} \sum_{i=1}^{m} \mathbb{I}^{(1)}_t(Z^*_i) \frac{p(Z^*_i)}{q_{\boldsymbol{v}_*}(Z^*_i)}\,, \qquad Z^*_1, \dots, Z^*_m \sim q_{\boldsymbol{v}_*}$$

MFCE: Costs of the cross-entropy method

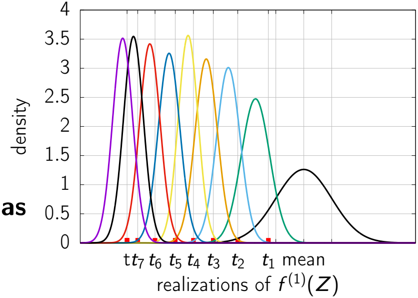
Step t_i defined by quantile parameter ρ

- Quantile parameter typically $\rho \in [10^{-2}, 10^{-1}]$
- Step t_i is ρ quantile corresponding to $q_{\boldsymbol{v}_{i-1}}$

Number of steps T

- Introduce minimal step size $\delta > 0$
- Number steps T is bounded as

$$T \leq \frac{t_1 - t}{\delta}$$



Costs of cross-entropy method bounded as

$$c(\hat{P}_t^{\mathrm{CE}}) \leq \frac{(t_1 - t)\, m w_1}{\delta}$$

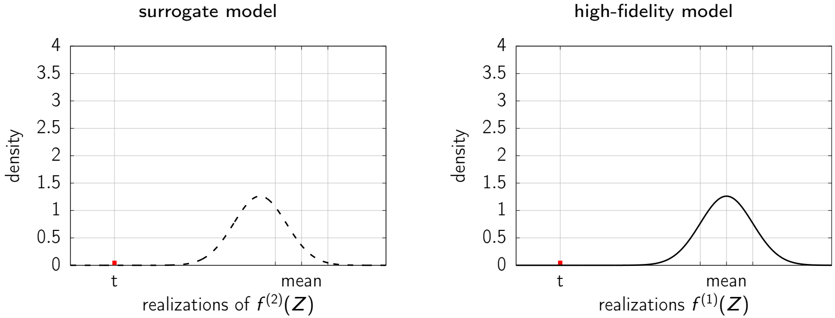
$\Rightarrow$ **Critically depends on t_1 and thus on density used in first step**

MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method



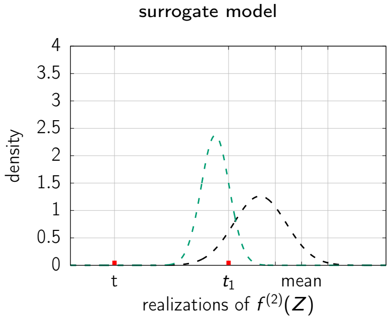
- Find biasing density $q_{\boldsymbol{v}_*}^{(k)}$ with surrogate $f^{(k)}$
- Start with $q_{\boldsymbol{v}_*}^{(k)}$ to find biasing density $q_{\boldsymbol{v}_*}^{(k-1)}$ with $f^{(k-1)}$
- Repeat until $q_{\boldsymbol{v}_*}^{(1)}$ is found with $f^{(1)}$

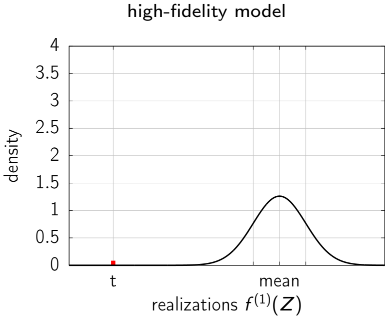
MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method





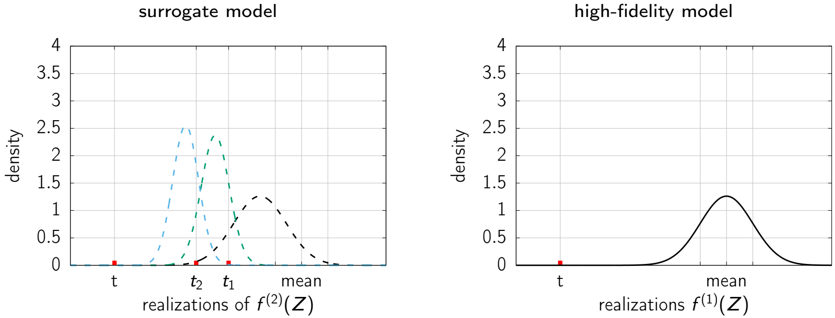
- Find biasing density $q_{\boldsymbol{v}_*}^{(k)}$ with surrogate $f^{(k)}$
- Start with $q_{\boldsymbol{v}_*}^{(k)}$ to find biasing density $q_{\boldsymbol{v}_*}^{(k-1)}$ with $f^{(k-1)}$
- Repeat until $q_{\boldsymbol{v}_*}^{(1)}$ is found with $f^{(1)}$

MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method



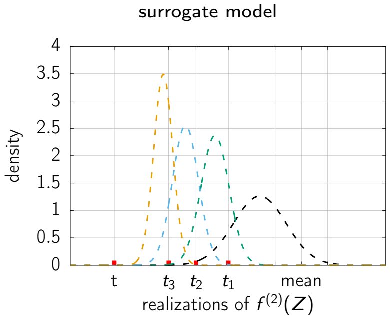
- Find biasing density $q^{(k)}_{\boldsymbol{v}_*}$ with surrogate $f^{(k)}$
- Start with $q^{(k)}_{\boldsymbol{v}_*}$ to find biasing density $q^{(k-1)}_{\boldsymbol{v}_*}$ with $f^{(k-1)}$
- Repeat until $q^{(1)}_{\boldsymbol{v}_*}$ is found with $f^{(1)}$

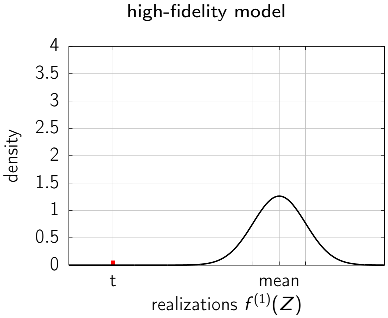
MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method





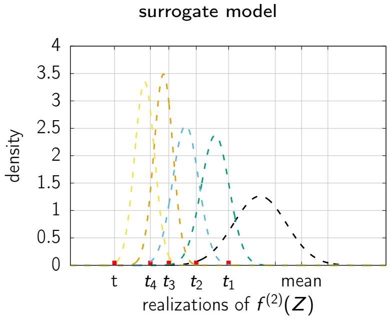
- Find biasing density $q^{(k)}_{\boldsymbol{v}_*}$ with surrogate $f^{(k)}$
- Start with $q^{(k)}_{\boldsymbol{v}_*}$ to find biasing density $q^{(k-1)}_{\boldsymbol{v}_*}$ with $f^{(k-1)}$
- Repeat until $q^{(1)}_{\boldsymbol{v}_*}$ is found with $f^{(1)}$

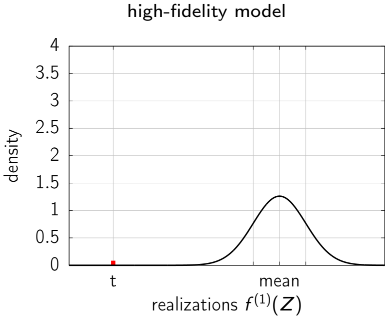
MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method





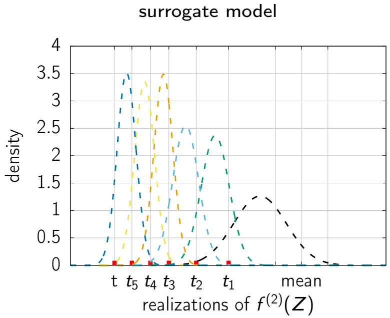
- Find biasing density $q^{(k)}_{\boldsymbol{v}_*}$ with surrogate $f^{(k)}$
- Start with $q^{(k)}_{\boldsymbol{v}_*}$ to find biasing density $q^{(k-1)}_{\boldsymbol{v}_*}$ with $f^{(k-1)}$
- Repeat until $q^{(1)}_{\boldsymbol{v}_*}$ is found with $f^{(1)}$

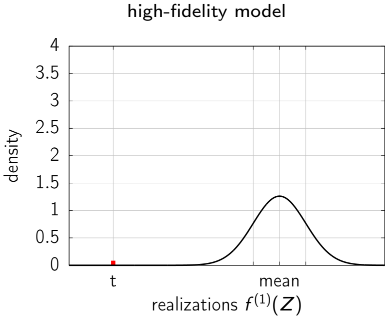
MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method





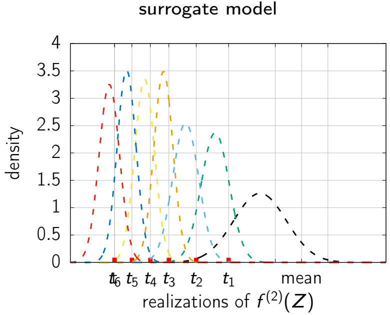
- Find biasing density $q^{(k)}_{\boldsymbol{v}_*}$ with surrogate $f^{(k)}$
- Start with $q^{(k)}_{\boldsymbol{v}_*}$ to find biasing density $q^{(k-1)}_{\boldsymbol{v}_*}$ with $f^{(k-1)}$
- Repeat until $q^{(1)}_{\boldsymbol{v}_*}$ is found with $f^{(1)}$

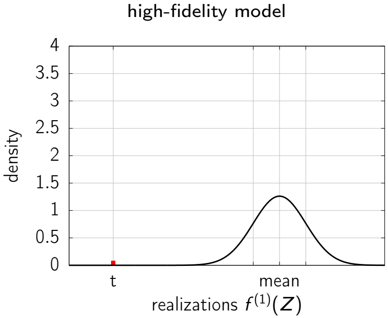
MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method





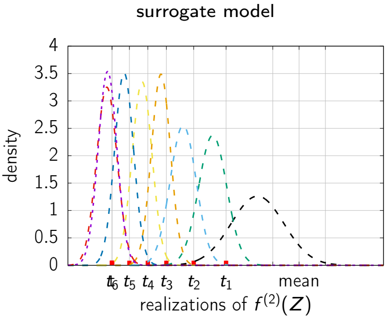
- Find biasing density $q^{(k)}_{\boldsymbol{v}_*}$ with surrogate $f^{(k)}$
- Start with $q^{(k)}_{\boldsymbol{v}_*}$ to find biasing density $q^{(k-1)}_{\boldsymbol{v}_*}$ with $f^{(k-1)}$
- Repeat until $q^{(1)}_{\boldsymbol{v}_*}$ is found with $f^{(1)}$

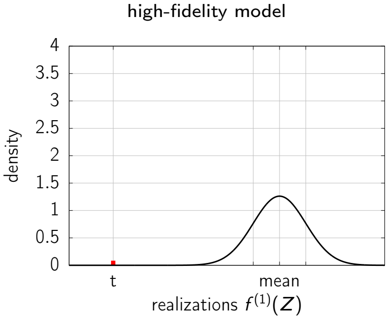
MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method





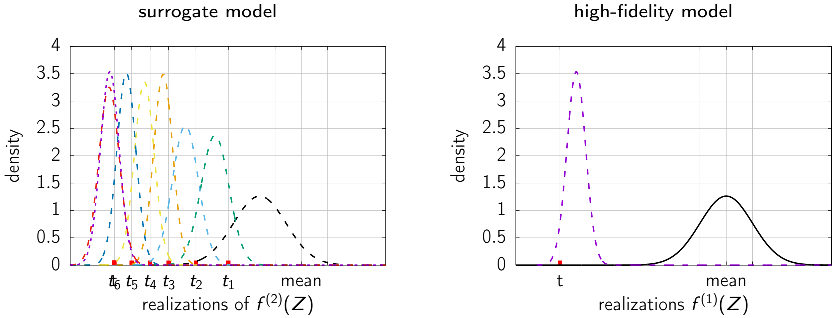
- Find biasing density $q_{\boldsymbol{v}_*}^{(k)}$ with surrogate $f^{(k)}$
- Start with $q_{\boldsymbol{v}_*}^{(k)}$ to find biasing density $q_{\boldsymbol{v}_*}^{(k-1)}$ with $f^{(k-1)}$
- Repeat until $q_{\boldsymbol{v}_*}^{(1)}$ is found with $f^{(1)}$

MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method



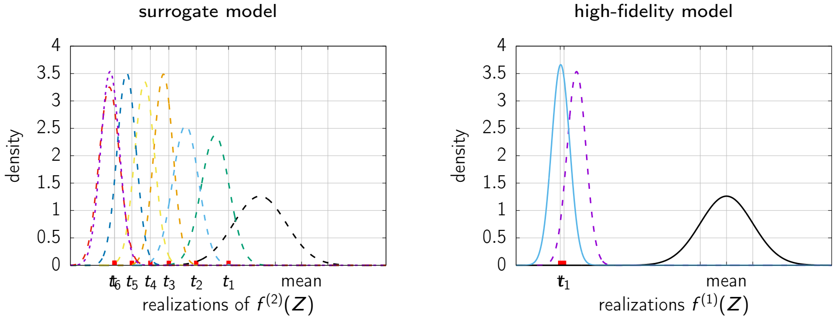
- Find biasing density $q_{\boldsymbol{v}_*}^{(k)}$ with surrogate $f^{(k)}$
- Start with $q_{\boldsymbol{v}_*}^{(k)}$ to find biasing density $q_{\boldsymbol{v}_*}^{(k-1)}$ with $f^{(k-1)}$
- Repeat until $q_{\boldsymbol{v}_*}^{(1)}$ is found with $f^{(1)}$

MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method



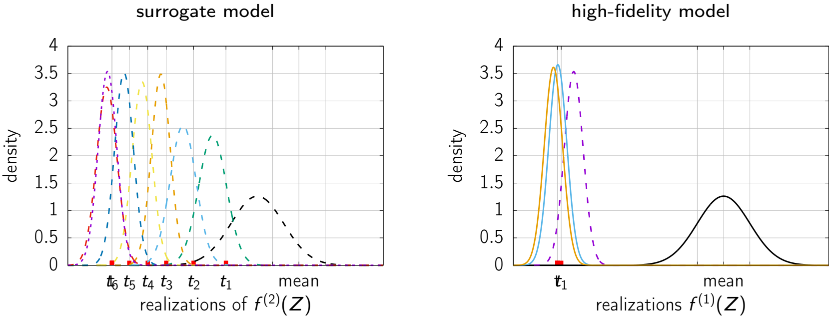
- Find biasing density $q^{(k)}_{\boldsymbol{v}_*}$ with surrogate $f^{(k)}$
- Start with $q^{(k)}_{\boldsymbol{v}_*}$ to find biasing density $q^{(k-1)}_{\boldsymbol{v}_*}$ with $f^{(k-1)}$
- Repeat until $q^{(1)}_{\boldsymbol{v}_*}$ is found with $f^{(1)}$

MFCE: Multifidelity-preconditioned cross-entropy method

Our approach: Use surrogates to reduce t_1 and thus #iters with $f^{(1)}$

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

Multifidelity-preconditioned cross-entropy (MFCE) method



- Find biasing density $q^{(k)}_{\boldsymbol{v}_*}$ with surrogate $f^{(k)}$
- Start with $q^{(k)}_{\boldsymbol{v}_*}$ to find biasing density $q^{(k-1)}_{\boldsymbol{v}_*}$ with $f^{(k-1)}$
- Repeat until $q^{(1)}_{\boldsymbol{v}_*}$ is found with $f^{(1)}$

MFCE: Analysis for two models

Consider high-fidelity $f^{(1)}$ and low-fidelity model $f^{(2)}$

- Let $q^{(2)}_{\boldsymbol{v}_*}$ be the biasing density found with $f^{(2)}$
- Set $t_1^{(1)}$ to be the ρ-quantile corresponding to $q^{(2)}_{\boldsymbol{v}_*}$
- Set t_p to be the ρ-quantile corresponding to the nominal density p
- Need to show

$$\mathbb{P}_{q^{(2)}_{\boldsymbol{v}_*}}\left[f^{(1)} \leq t_1^{(1)}\right] - \mathbb{P}_p\left[f^{(1)} \leq t_1^{(1)}\right] \geq 0$$

to have $t_1^{(1)} \leq t_p$

Make two assumptions

- A "local" bound $0 < \alpha$ on the error $|f^{(1)}(z) - f^{(2)}(z)|$
- Lipschitz continuous distribution functions $F_q^{(i)}(t) = \mathbb{P}_q[f^{(i)} \leq t]$

Proposition 2 in [P., Kramer, Willcox, 2017]

$$\mathbb{P}_{q^{(2)}_{\boldsymbol{v}_*}}\left[f^{(1)} \leq t_1^{(1)}\right] - \mathbb{P}_p\left[f^{(1)} \leq t_1^{(1)}\right] \gtrsim -\alpha$$

MFCE: Numerical examples: Heat transfer

Consider

$$-\nabla \cdot (a(\omega,\xi)\nabla u(\omega,\xi)) = 1\,, \qquad \xi \in (0,1)\,, \tag{1}$$

$$u(\omega,0) = 0\,, \tag{2}$$

$$\partial_n u(\omega,1) = 0\,, \tag{3}$$

- Coefficient a is given as

$$a(\omega,\xi) = \mathrm{e}^{z_1(\omega)}\mathrm{e}^{-0.5\frac{|\xi-0.5|}{0.0225}} + \mathrm{e}^{z_2(\omega)}\mathrm{e}^{-0.5\frac{|\xi-0.8|}{0.0225}}$$

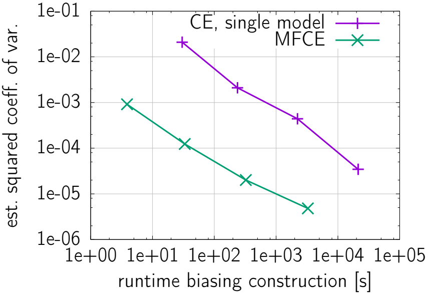
- Random vector $Z = [z_1, z_2]$ normally distributed
- System response

$$f(Z(\omega)) = u(\omega,1)$$

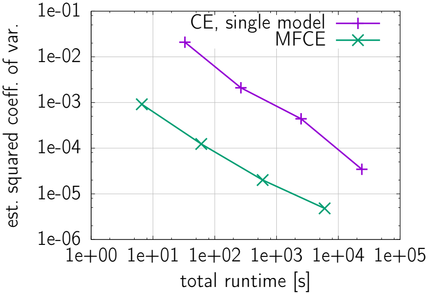
- Discretize with varying mesh width $h \in \{2^{-8}, 2^{-7}, \ldots, 2^{-4}\}$
- Obtain models $f^{(1)}, \ldots, f^{(5)}$

Goal is to estimate $P_f = \mathbb{P}_p[f^{(1)} \leq 1.4]$ with reference $P_f \approx 6 \times 10^{-9}$

MFCE: Numerical examples: Speedup



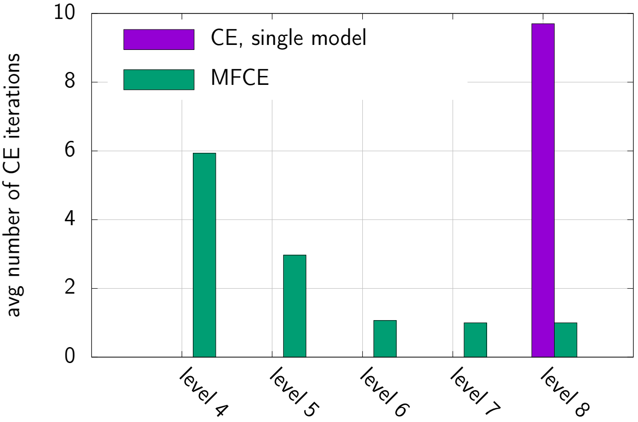
(a) biasing runtime



(b) total runtime

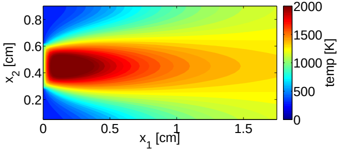
- Single-fidelity approach uses $f^{(1)}$
- MFCE uses $f^{(1)}, \dots, f^{(5)}$
- MFCE reduces runtime by almost 2 orders of magnitude

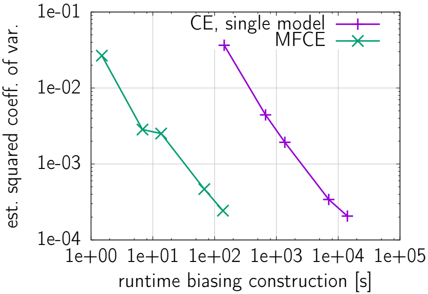
MFCE: Numerical examples: Number of iterations



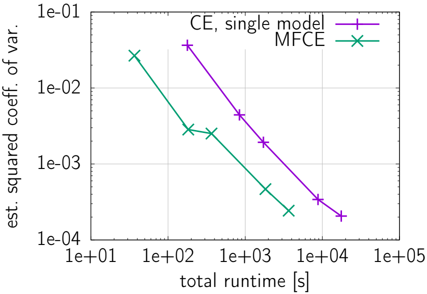
- Number of iterations averaged over 30 runs
- MFCE performs most iterations with coarse-grid models

MFCE: Numerical examples: Reacting flow





(a) biasing runtime



(b) total runtime

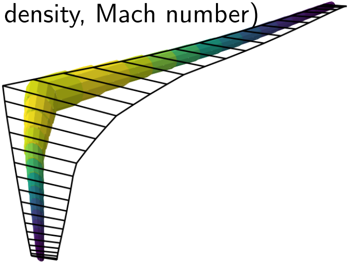
Reacting-flow problem

- Three models: data-fit, projection-based, and high-fidelity model
- Estimate probability that temperature is below threshold, ref $P_f \approx 10^{-6}$
- MFCE reduces iterations with high-fidelity model from ≈ 5 to ≈ 1

MFCE: Apply to OpenAeroStruct

Consider baseline UAV definition in OpenAeroStruct

- Design variables are thickness and position of control points
- Uncertain flight conditions (angle of attack, air density, Mach number)
- Output is fuel burn



Estimate 10^{-6}-quantile γ (value-at-risk)

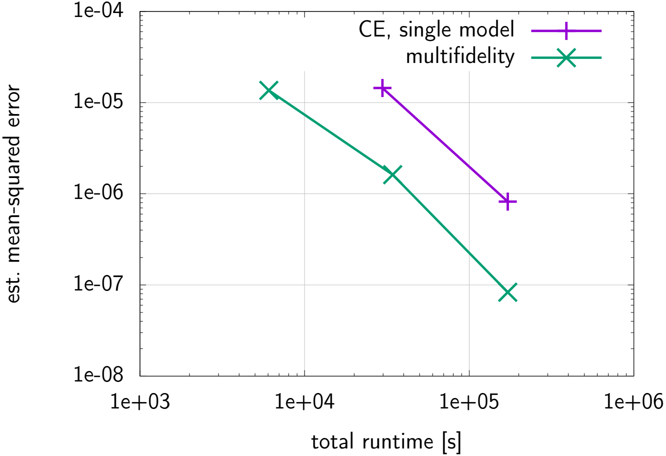
$$\mathbb{P}_p[f^{(1)} \leq \gamma] = 10^{-6}$$

Derive a data-fit surrogate

- Take a $3 \times 3 \times 3$ equidistant grid in stochastic domain
- Evaluate high-fidelity model at those 27 points
- Derive linear interpolant of output

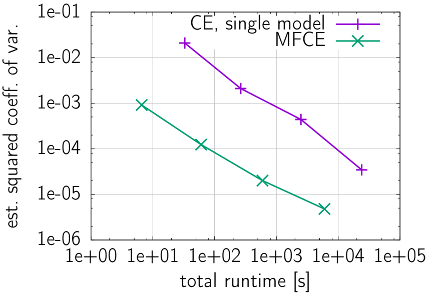
Apply multifidelity pre-conditioned cross-entropy method

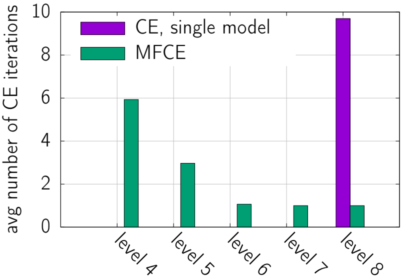
MFCE: OpenAeroStruct: Value-at-Risk computation



- Computing 10^{-6}-quantile for a fixed design point
- Multifidelity approach achieves up to one order of magnitude speedup

Conclusions





Multifidelity rare event simulation

- Leverage surrogate models for runtime speedup
- Recourse to high-fidelity model for accuracy guarantees

Our references

1. P., Cui, Marzouk, Willcox: *Multifidelity Importance Sampling*. Computer Methods in Applied Mechanics and Engineering, 300:490-509, 2016.
2. P., Kramer, Willcox: *Combining multiple surrogate models to accelerate failure probability estimation with expensive high-fidelity models*. J. of Computational Physics, 341:61-75, 2017.
3. P., Kramer, Willcox: *Multifidelity preconditioning of the cross-entropy method for rare event simulation and failure probability estimation*. *submitted*, 2017.


PostDoc position available
• data-driven model reduction
• uncertainty propagation
• Bayesian inverse problems
https://pehersto.engr.wisc.edu
peherstorfer@wisc.edu