Multifidelity importance sampling

Benjamin Peherstorfer[a,*], Tiangang Cui[a], Youssef Marzouk[a], Karen Willcox[a]

[a]*Department of Aeronautics & Astronautics, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*

Abstract

Estimating statistics of model outputs with the Monte Carlo method often requires a large number of model evaluations. This leads to long runtimes if the model is expensive to evaluate. Importance sampling is one approach that can lead to a reduction in the number of model evaluations. Importance sampling uses a biasing distribution to sample the model more efficiently, but generating such a biasing distribution can be difficult and usually also requires model evaluations. A different strategy to speed up Monte Carlo sampling is to replace the computationally expensive high-fidelity model with a computationally cheap surrogate model; however, because the surrogate model outputs are only approximations of the high-fidelity model outputs, the estimate obtained using a surrogate model is in general biased with respect to the estimate obtained using the high-fidelity model. We introduce a multifidelity importance sampling (MFIS) method, which combines evaluations of both the high-fidelity and a surrogate model. It uses a surrogate model to facilitate the construction of the biasing distribution, but relies on a small number of evaluations of the high-fidelity model to derive an unbiased estimate of the statistics of interest. We prove that the MFIS estimate is unbiased even in the absence of accuracy guarantees on the surrogate model itself. The MFIS method can be used with any type of surrogate model, such as projection-based reduced-order models and data-fit models. Furthermore, the MFIS method is applicable to black-box models, i.e., where only inputs and the corresponding outputs of the high-fidelity and the surrogate model are available but not the details of the models themselves. We demonstrate on nonlinear and time-dependent problems that our MFIS method achieves speedups of up to several orders of magnitude compared to Monte Carlo with importance sampling that uses the high-fidelity model only.

Keywords: Monte Carlo method, importance sampling, surrogate modeling, multifidelity methods

1. Introduction

Models of real-world engineering systems often depend on multiple inputs such as material properties and forcing terms. In most cases, these inputs are not known exactly due to measurement errors, noise, or small perturbations in the manufacturing processes, and thus they are modeled as random variables with a distribution that accounts for these uncertainties. The output of the model then also becomes a random variable, and one is typically interested in the statistics of the output. The Monte Carlo method evaluates the model at samples drawn from the distribution of the inputs, and estimates the statistics from the obtained outputs. Monte Carlo estimators converge slowly with the number of samples and thus many model evaluations are necessary to achieve a given error tolerance. A common remedy to reduce the runtime of the Monte Carlo method is to replace the high-fidelity model with a computationally cheap surrogate model with a lower fidelity; however, this leads in general to an estimator with a bias, see, e.g., [1] for a discussion. We therefore propose to speed up the Monte Carlo method by combining, instead of replacing, the high-fidelity model with the surrogate model. Even in the absence of accuracy guarantees on the surrogate model itself, our multifidelity estimator is unbiased with respect to the statistics of the high-fidelity model

---

*Corresponding author
Email addresses:* `pehersto@mit.edu` (Benjamin Peherstorfer), `+1-617-253-7831` (Benjamin Peherstorfer)
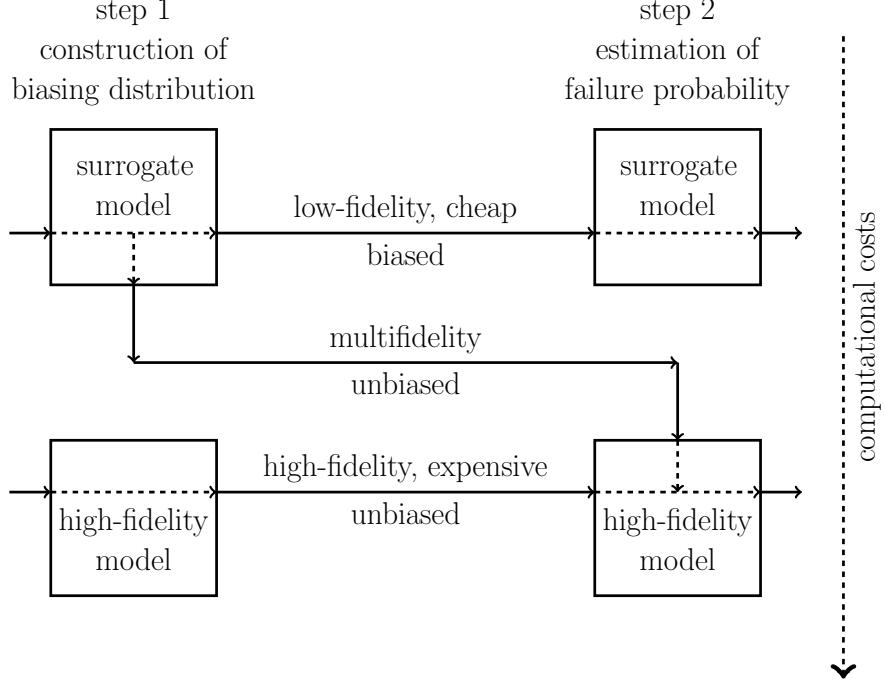
Figure 1: The proposed multifidelity importance sampling uses the surrogate model to construct a biasing distribution and the high-fidelity model to derive the actual estimate. This multifidelity approach leads to an unbiased estimator of the failure probability independent of the accuracy of the surrogate model.

outputs. A surrogate model that poorly approximates the high-fidelity model might not yield computational savings, but the unbiasedness of the multifidelity estimator is guaranteed.

We develop our multifidelity approach in the context of the computation of failure probabilities [2, 3] with the Monte Carlo method and importance sampling [4, 5, 6]. Failure probabilities are typically small and thus the Monte Carlo method requires many model evaluations to derive an accurate estimate. Importance sampling is a variance reduction technique that aims to reduce the number of required model evaluations by using a problem-dependent sampling strategy. The computation of failure probabilities is only one application of importance sampling. For example, importance sampling is also used to estimate statistics over distributions that cannot be sampled directly. We refer to, e.g., [4] for additional applications of importance sampling, and therefore potential applications of our multifidelity approach.

In its simplest form, the Monte Carlo method with importance sampling consists of two steps. In the first step, a biasing distribution is generated. In the second step, samples are drawn from the biasing distribution to derive an estimate of the failure probability. The key challenge is to construct an effective biasing distribution that leads to many samples for which a failure event occurs [4, 6]. Several approaches for constructing a biasing distribution have been presented. A common approach is the cross-entropy method [7, 8]. This iterative optimization algorithm fits a standard distribution with respect to the Kullback-Leibler divergence to inputs that correspond to failures of the system. There are also adaptive methods that break with the splitting of importance sampling into two steps and instead build the biasing distribution while the model is sampled and while the estimate is derived, see, e.g., [9, 10, 11, 12, 13]. A general overview of importance sampling and methods for constructing biasing distributions is given in [4, 6].

We propose to use the surrogate model to construct the biasing distribution in the first step of the Monte Carlo method with importance sampling and the high-fidelity model to derive the estimate in the second step. This is a multifidelity approach because both models are used to compute an estimate of the failure probability, see Figure 1.

Past work has used surrogate models to speed up the Monte Carlo method. The control variates method

provides a framework for variance reduction, using additive corrections from an auxiliary random variable that is correlated to the high-fidelity model outputs [14]. One approach based on the control variate method is the multi-level Monte Carlo method [15, 16], which combines the outputs of a hierarchy of coarse grid solutions using control variates. The multifidelity Monte Carlo method introduced in [17] allows arbitrary surrogate models that do not necessarily stem from coarse grid discretizations. In [18] the multifidelity Monte Carlo method is used in the context of optimization where intermediate results of the iterative optimization process are used instead of a surrogate model. In [19], statistics of high-fidelity models are estimated, where a residual-based indicator is introduced to decide whether a surrogate model is evaluated or recourse to the high-fidelity model is necessary. All these approaches, however, do not explicitly target the speedup of the estimation of failure probabilities or importance sampling in general. In [1], a multifidelity approach for computing failure probabilities is presented, which is extended in [20] with importance sampling and a cross-entropy method to handle rare events. The multifidelity approaches in [1, 20] adaptively switch between the high-fidelity model and the surrogate model. An analysis shows unbiasedness of the corresponding estimators if tolerance parameters are chosen well. The approach for computing failure probabilities presented in [21] uses *a posteriori* error estimators of the surrogate model to decide whether the outputs of the surrogate or the high-fidelity model should be used. Surrogate models with rigorous and efficient *a posteriori* error estimators are currently available only for a limited class of problems [22, 23].

In contrast, we use the surrogate model to construct the biasing distribution only, and thus do not rely on the availability of *a priori* error bounds or *a posteriori* error estimators. We show that the estimator of our multifidelity approach is an unbiased estimator of the failure probability independent of the accuracy of the surrogate model, although the speedup increases as the approximation quality of the surrogate model improves. The numerical results demonstrate runtime savings of orders of magnitude compared to the Monte Carlo method with importance sampling that uses the high-fidelity model only. Since our method requires only evaluations of the high-fidelity and the surrogate model, it applies to black-box models. Furthermore, our multifidelity approach is applicable to arbitrary surrogate models, including projection-based reduced-order models, data-fit and response surface models, and simplified physics models.

The following Section 2 gives a brief overview of the estimation of failure probabilities with the Monte Carlo method with importance sampling. Section 3 introduces the multifidelity approach and proves the unbiasedness of the multifidelity estimator. The numerical results in Section 4 confirm the unbiasedness of our estimator and demonstrate the runtime savings on linear and nonlinear problems. Finally, Section 5 draws conclusions.

2. Monte Carlo method with importance sampling

This section introduces the Monte Carlo method with importance sampling for the estimation of failure probabilities. The failure probability setting is defined in Section 2.1, and importance sampling is discussed in Section 2.2.

2.1. Failure probability estimation with the Monte Carlo method

We consider an input-output system with $d \in \mathbb{N}$ inputs and $d' \in \mathbb{N}$ outputs, which is modeled by the high-fidelity model $s : \mathbb{R}^d \to \mathbb{R}^{d'}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with the sample space Ω, the σ-algebra $\mathcal{F}$, and the probability distribution $\mathbb{P}$. The uncertainties in the inputs are represented by the probability distribution $\mathbb{P}$ with probability density function (PDF) p. Let $Z : \Omega \to \mathbb{R}^d$ be a random variable. The inputs to the high-fidelity model s are the components $\boldsymbol{z} = [z_1, \ldots, z_d]^T \in \mathbb{R}^d$ of a realization $Z(\omega)$ of Z for $\omega \in \Omega$. The outputs $s(\boldsymbol{z}) \in \mathbb{R}^{d'}$ are the realization $s(Z(\omega))$ of the random variable $s(Z)$.

We are interested in the failure probability of the system described by the high-fidelity model. Let $g : \mathbb{R}^{d'} \to \mathbb{R}$ be the limit state function. Without loss of generality, the system fails for an input $\boldsymbol{z}$ if $g(s(\boldsymbol{z})) < 0$. This leads to the failure domain

$$\mathcal{G} = \left\{ \boldsymbol{z} \in \mathbb{R}^d \, : \, g(s(\boldsymbol{z})) < 0 \right\},$$

which is a subset of $\mathbb{R}^d$ and an element of the σ-algebra $\mathcal{F}$. For inputs in the safe domain,

$$\left\{ \boldsymbol{z} \in \mathbb{R}^d \, : \, g(s(\boldsymbol{z})) \geq 0 \right\}, \tag{1}$$

the system does not fail. The failure boundary is the level set

$$\left\{ \boldsymbol{z} \in \mathbb{R}^d \, : \, g(s(\boldsymbol{z})) = 0 \right\}.$$

With the failure domain $\mathcal{G}$ we define the failure probability $P_f \in \mathbb{R}$ of the system as

$$P_f = \mathbb{E}_p[\mathbb{I}_\mathcal{G}(Z)], \tag{2}$$

where $\mathbb{I}_\mathcal{G} : \mathbb{R}^d \to \{0,1\}$ is the indicator function of the failure domain defined as

$$\mathbb{I}_\mathcal{G}(\boldsymbol{z}) = \begin{cases} 1, & \boldsymbol{z} \in \mathcal{G}, \\ 0, & \text{else} \end{cases} \tag{3}$$

and where

$$\mathbb{E}_p[\mathbb{I}_\mathcal{G}(Z)] = \int_{\mathbb{R}^d} \mathbb{I}_\mathcal{G}(\boldsymbol{z}) p(\boldsymbol{z}) \mathrm{d}\boldsymbol{z} \tag{4}$$

is the expectation of the random variable $\mathbb{I}_\mathcal{G}(Z)$. The variance of $\mathbb{I}_\mathcal{G}(Z)$ with respect to the PDF p is

$$\mathrm{var}_p[\mathbb{I}_\mathcal{G}(Z)] = \int_{\mathbb{R}^d} \left(\mathbb{I}_\mathcal{G}(\boldsymbol{z}) - \mathbb{E}_p[\mathbb{I}_\mathcal{G}(Z)]\right)^2 p(\boldsymbol{z}) \mathrm{d}\boldsymbol{z}. \tag{5}$$

The Monte Carlo method is often used to estimate expectations such as (2). The Monte Carlo method draws $M \in \mathbb{N}$ independent and identically distributed (i.i.d.) samples $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_M \in \mathbb{R}^d$ from the distribution of Z, i.e., M realizations of the random variable Z, and computes the Monte Carlo estimate

$$P_f^{\mathrm{MC}}(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_M) = \frac{1}{M} \sum_{i=1}^{M} \mathbb{I}_\mathcal{G}(\boldsymbol{z}_i). \tag{6}$$

We note that we distinguish between the estimate $P_f^{\mathrm{MC}}(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_M)$, which is a scalar value, and the Monte Carlo estimator $P_f^{\mathrm{MC}}(Z)$, which is a random variable. The Monte Carlo estimator $P_f^{\mathrm{MC}}(Z)$ is an unbiased estimator of the failure probability of the system [4].

The root mean square error (RMSE) of the Monte Carlo estimator $P_f^{\mathrm{MC}}(Z)$ is

$$e(P_f^{\mathrm{MC}}) = \sqrt{\frac{\mathrm{var}_p[\mathbb{I}_\mathcal{G}(Z)]}{M}}, \tag{7}$$

see [5]. With (7), it follows that for a threshold parameter $0 < \epsilon \in \mathbb{R}$, an RMSE $e(P_f^{\mathrm{MC}}) \leq \epsilon$ is achieved with

$$M = \left\lceil \frac{\mathrm{var}_p[\mathbb{I}_\mathcal{G}(Z)]}{\epsilon^2} \right\rceil$$

samples, where $\lceil \cdot \rceil$ ensures that M is an integer. In the following, we hide $\lceil \cdot \rceil$ and write

$$M \simeq \frac{\mathrm{var}_p[\mathbb{I}_\mathcal{G}(Z)]}{\epsilon^2}$$

instead. Let $0 < b \in \mathbb{R}$ denote the costs of one high-fidelity model evaluation. The costs to achieve an RMSE $e(P_f^{\mathrm{MC}}) \leq \epsilon$ are

$$c(P_f^{\mathrm{MC}}) = Mb \simeq \frac{\mathrm{var}_p[\mathbb{I}_\mathcal{G}(Z)]}{\epsilon^2} b.$$

Thus, the costs depend on the variance of the integrand $\mathbb{I}_\mathcal{G}$. This becomes especially problematic if the variance $\mathrm{var}_p[\mathbb{I}_\mathcal{G}(Z)]$ is large.

4

2.2. Importance sampling

Variance reduction methods aim to reduce the RMSE (7) by transforming (2) to integrate a function with a lower variance than the original function $\mathbb{I}_{\mathcal{G}}$. Importance sampling is one such variance reduction method that has been successfully used to estimate failure probabilities, see, e.g., [4, 6, 24].

Importance sampling introduces a biasing distribution with PDF q and a random variable $Z' : \Omega \to \mathbb{R}^d$ with distribution with PDF q. A realization of Z' is denoted as $\boldsymbol{z}' = [z_1', \dots, z_d']^T \in \mathbb{R}^d$. In the following, the distribution of Z is the nominal distribution and the corresponding PDF p is the nominal PDF. The distribution of Z' is the biasing distribution and q the biasing PDF. The biasing PDF q is constructed such that $\operatorname{supp}(p) \subseteq \operatorname{supp}(q)$, where

$$\operatorname{supp}(p) = \{ \boldsymbol{z} \in \mathbb{R}^d : p(\boldsymbol{z}) > 0 \} ,$$

denotes the support of the PDF p. Let $w : \operatorname{supp}(q) \to \mathbb{R}$ be the weight function with $w(\boldsymbol{z}') = p(\boldsymbol{z}')/q(\boldsymbol{z}')$. The weight $w(\boldsymbol{z}')$ is the importance weight of a realization $\boldsymbol{z}' = Z'(\omega)$. Because $\operatorname{supp}(p) \subseteq \operatorname{supp}(q)$ holds, the failure probability (2) equals the expectation of the random variable $\mathbb{I}_{\mathcal{G}}(Z')$ weighted with the random variable $w(Z')$. That is, we have

$$P_f = \mathbb{E}_q[\mathbb{I}_{\mathcal{G}}(Z')w(Z')] . \tag{8}$$

The expectation (8) is approximated with the Monte Carlo method, with samples $\boldsymbol{z}_1', \dots, \boldsymbol{z}_M'$ drawn from the biasing distribution. Thus, the importance sampling estimate of P_f with samples $\boldsymbol{z}_1', \dots, \boldsymbol{z}_M'$ is

$$P_f^{\text{IS}}(\boldsymbol{z}_1', \dots, \boldsymbol{z}_M') = \frac{1}{M} \sum_{i=1}^{M} \mathbb{I}_{\mathcal{G}}(\boldsymbol{z}_i')w(\boldsymbol{z}_i') . \tag{9}$$

Therefore, the Monte Carlo method with importance sampling consists of two steps. In step one, the biasing distribution is generated. In step two, the importance sampling estimate (9) is calculated with samples drawn from the biasing distribution. The importance sampling estimator $P_f^{\text{IS}}(Z')$ is an unbiased estimator of P_f because $\operatorname{supp}(p) \subseteq \operatorname{supp}(q)$ [4].

If $\mathbb{E}_q[\mathbb{I}_{\mathcal{G}}(Z')^2 w(Z')^2] < \infty$, then the RMSE of the importance sampling estimator is

$$e(P_f^{\text{IS}}) = \sqrt{\frac{\operatorname{var}_q[\mathbb{I}_{\mathcal{G}}(Z')w(Z')]}{M}} . \tag{10}$$

If the variance $\operatorname{var}_q[\mathbb{I}_{\mathcal{G}}(Z')w(Z')]$ is smaller than $\operatorname{var}_p[\mathbb{I}_{\mathcal{G}}(Z)]$, then the RMSE of the importance sampling estimator is smaller than the RMSE of the Monte Carlo estimator, for the same number of samples M. Using (10), it follows that

$$M \simeq \frac{\operatorname{var}_q[\mathbb{I}_{\mathcal{G}}(Z')w(Z')]}{\epsilon^2}$$

samples are required to achieve an RMSE $e(P_f^{\text{IS}}) \leq \epsilon$. The costs to achieve $e(P_f^{\text{IS}}) \leq \epsilon$ are split into the costs of generating the biasing distribution $c_{\text{bias}} = bN$ from $N \in \mathbb{N}$ samples and the costs of evaluating the high-fidelity model at the M samples drawn from the biasing distribution, i.e.,

$$c(P_f^{\text{IS}}) = c_{\text{bias}} + Mb \simeq Nb + \frac{\operatorname{var}_q[\mathbb{I}_{\mathcal{G}}(Z')w(Z')]}{\epsilon^2} b . \tag{11}$$

3. Multifidelity importance sampling

Our multifidelity importance sampling (MFIS) method aims to reduce the runtime of the Monte Carlo method with importance sampling by combining outputs of the high-fidelity model with outputs of a surrogate model. The surrogate model replaces the high-fidelity model during the construction of the biasing distribution. If the surrogate model leads to an accurate approximation of the limit state function, then the following MFIS method often leads to significant reductions in runtime compared to the Monte Carlo

method. If the surrogate model provides a poor approximation of the limit state function, then the runtime of the MFIS method might be even larger than the runtime of the Monte Carlo method; however, the unbiasedness of the MFIS estimator is guaranteed.

Similarly to the Monte Carlo method with importance sampling, our MFIS method consists of two steps. In step one, the biasing distribution is constructed using a surrogate model only. Surrogate models are discussed in Section 3.1. Details on the construction of the biasing distribution with a surrogate model are given in Section 3.2. In step two, the MFIS method derives an estimate of P_f by evaluating the high-fidelity model at samples drawn from the biasing distribution. We first define the MFIS estimator in Section 3.3 and then prove that it is an unbiased estimator of the failure probability P_f. The MFIS method is summarized in Section 3.4, and its computational costs are discussed in Section 3.5.

3.1. Surrogate models

In the following, we have a surrogate model $s_r : \mathbb{R}^d \to \mathbb{R}^{d'}$ that provides approximate outputs of the high-fidelity model s. Surrogate models are often orders of magnitude cheaper to evaluate than the high-fidelity model s and still provide accurate approximations of the high-fidelity model outputs. Many techniques to construct surrogate models of high-fidelity models are available. There are data-fit models [25, 26, 27], which use only inputs and outputs of the high-fidelity model to construct s_r, and projection-based reduced-order models [22, 28], which typically require access to the internal structure of the high-fidelity model s. Our MFIS method is applicable to any type of surrogate model.

Let

$$\mathcal{G}_r = \{\boldsymbol{z} : g(s_r(\boldsymbol{z})) < 0\} \tag{12}$$

be the failure domain induced by the surrogate model, and let $\mathbb{I}_{\mathcal{G}_r}$ be the indicator function of $\mathcal{G}_r$ following (3). To assess the quality of a surrogate model, consider the symmetric difference $(\mathcal{G}_r \setminus \mathcal{G}) \cup (\mathcal{G} \setminus \mathcal{G}_r)$ of the failure domains $\mathcal{G}_r$ and $\mathcal{G}$, which contains inputs $\boldsymbol{z} \in \mathbb{R}^d$ for which either the surrogate or the high-fidelity model predict a failure but not both. The corresponding indicator function is $\mathbb{I}_{(\mathcal{G}_r \setminus \mathcal{G}) \cup (\mathcal{G} \setminus \mathcal{G}_r)}(\boldsymbol{z}) = (\mathbb{I}_{\mathcal{G}_r}(\boldsymbol{z}) - \mathbb{I}_{\mathcal{G}}(\boldsymbol{z}))^2$. We assess here the quality of a surrogate model by the expectation

$$\mathbb{E}_p[\mathbb{I}_{(\mathcal{G}_r \setminus \mathcal{G}) \cup (\mathcal{G} \setminus \mathcal{G}_r)}(Z)], \tag{13}$$

which is the probability that a realization $\boldsymbol{z}$ of Z is in the symmetric difference of the failure domains $\mathcal{G}_r$ and $\mathcal{G}$, and thus that the surrogate model predicts a failure but the high-fidelity model does not or vice versa. We emphasize that the error (13) is different from the maximal error

$$\max_{\boldsymbol{z} \in \mathbb{R}^d} |s_r(\boldsymbol{z}) - s(\boldsymbol{z})|, \tag{14}$$

which is usually considered in surrogate modeling [22, 28]. In our setting, a surrogate model could achieve an error (13) equal to zero even if there is an input that leads to large maximal error (14). We therefore explicitly allow here surrogate models that have large maximal errors.

3.2. Step one: Constructing biasing distribution with the surrogate model

In the first step of the MFIS method, the PDF q of the biasing distribution is constructed. The aim is to construct a biasing distribution so that samples drawn from the biasing distribution lie in the failure domain $\mathcal{G}$ with higher probability than samples drawn from the nominal distribution [4, 6].

To construct the biasing PDF q, we evaluate the surrogate model s_r at many samples drawn from the nominal distribution with PDF p. We can afford to evaluate the surrogate model at many samples because evaluating the surrogate model is computationally cheap. The PDF q of the biasing distribution is then constructed as a mixture model of normal distributions fitted to the samples in $\mathcal{G}_r$. Note that other density estimation techniques could be used instead as well [29, 30, 31]. It is reasonable to expect that q is a suitable biasing distribution for the failure domain $\mathcal{G}$ because if the surrogate model approximates the high-fidelity model well with respect to (13), the failure domain of the surrogate model $\mathcal{G}_r$ and the failure domain $\mathcal{G}$ of the high-fidelity model have a large overlap. We now discuss the construction of q in more detail.

Let $N \in \mathbb{N}$ and let $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N \in \mathbb{R}^d$ be samples drawn from the nominal distribution, i.e., $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N$ are realizations of the random variable Z. Let $N' \in \mathbb{N}$ be the number of samples in the failure domain $\mathcal{G}_r$. Note that $N' = 0$ is a valid case and will be discussed below. To check if a sample $\boldsymbol{z}_i$ is in $\mathcal{G}_r$, the surrogate model s_r is first evaluated at $\boldsymbol{z}_i$ and then the limit state function g is evaluated at the surrogate model output $s_r(\boldsymbol{z}_i)$. If $g(s_r(\boldsymbol{z}_i)) < 0$, then $\boldsymbol{z}_i \in \mathcal{G}_r$, otherwise $\boldsymbol{z}_i \notin \mathcal{G}_r$, see (12). Let $\boldsymbol{l} = [l_1, \ldots, l_{N'}]^T \in \mathbb{N}^{N'}$ be the indices of the samples in the failure domain and let

$$\mathcal{Z} = \{\boldsymbol{z}_{l_1}, \ldots, \boldsymbol{z}_{l_{N'}}\} \subset \mathcal{G}_r \tag{15}$$

be the set of these samples.

We first consider the case where $N' > 0$, i.e., where at least one of the samples $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N$ is in the failure domain $\mathcal{G}_r$. We define $\phi(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ to be the PDF of the d-dimensional multivariate normal distribution with mean $\boldsymbol{\mu} \in \mathbb{R}^d$ and covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$. A mixture model of $k \in \mathbb{N}$ normal distributions is a linear combination of k PDFs $\phi(\boldsymbol{z}; \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1), \ldots, \phi(\boldsymbol{z}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$,

$$\phi_{\mathrm{MM}}(\boldsymbol{z}; \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_k, \boldsymbol{\pi}) = \sum_{i=1}^k \pi_i \phi(\boldsymbol{z}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i), \tag{16}$$

where $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k \in \mathbb{R}^d$ and $\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_k \in \mathbb{R}^{d \times d}$ are the means and covariance matrices, respectively, and where $\boldsymbol{\pi} = [\pi_1, \ldots, \pi_k]^T \in \mathbb{R}^k$ are the mixture coefficients with $\sum_{i=1}^k \pi_i = 1$. Note that we discuss the selection of the number k of normal distributions in Section 4.1. Since $N' > 0$, we can now fit the parameters $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_k \in \mathbb{R}^{d \times d}$ and $\boldsymbol{\pi}$ of the mixture model ϕ_{MM} to the samples (15) in the failure domain $\mathcal{G}_r$. In the following, we use a maximum likelihood approach to fit the parameters of the mixture model, i.e., we solve the optimization problem

$$\max_{\{\boldsymbol{\mu}_i\}_1^k, \{\boldsymbol{\Sigma}_i\}_1^k, \boldsymbol{\pi} \in \mathbb{R}^k} \prod_{i=1}^{N'} \phi_{\mathrm{MM}}(\boldsymbol{z}_{l_i}; \boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_k, \boldsymbol{\pi}) \tag{17}$$

with the constraint $\sum_{i=1}^k \pi_i = 1$. Problem (17) is a high-dimensional nonconvex optimization problem, which is in general computationally expensive to solve; however, this particular problem (17) is a standard optimization problem in machine learning, which is well analyzed and for which efficient algorithms exist that compute approximate solutions [32, 33, 34]. We employ the expectation-maximization (EM) algorithm to solve (17) [32]. To cope with local maxima due to the non-convexity, we follow the usual approach and run the EM algorithm several times with random start parameters and then select the parameters that lead to the largest likelihood. We emphasize that the EM algorithm is a standard algorithm in machine learning. Its properties are discussed in detail in, e.g., the text books [32, 33, 34].

If $N' = 0$, then (15) is the empty set, which means that the outputs of the surrogate model at the samples $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N$ did not provide any information about the failure domain. In this case, we set the biasing PDF to the nominal PDF. As we will discuss in the proof of Theorem 1, this leads to an MFIS estimator that is equal to the Monte Carlo estimator and therefore no runtime savings can be gained with our MFIS method over the Monte Carlo method in this situation. A different approach is to consider $N' = 0$ as an indication that the number of samples N is set too low. A natural remedy is then to draw more samples, i.e., increase N, until the number of samples N' in the failure domain is larger than 0 and a mixture model of normal distributions can be constructed with the EM algorithm. In this way, N' could be used to guide the adaptive selection of the number of samples N.

Algorithm 1 summarizes the construction of the biasing distribution. Inputs are the surrogate model s_r, the nominal PDF p, the number N of samples to draw from the nominal distribution, and the number k of normal distributions in the mixture model ϕ_{MM}. The algorithm first evaluates the surrogate model at N samples drawn from the nominal distribution. If none of these samples is in the failure domain $\mathcal{G}_r$, the nominal PDF p is returned as the PDF of the biasing distribution. If samples are in $\mathcal{G}_r$, a mixture model ϕ_{MM} is fitted to these samples with the EM algorithm. The algorithm then returns the fitted mixture model.

Algorithm 1 Biasing distribution

1: **procedure** BIASINGDIST(s_r, p, N, k)
2:     Draw N samples $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N$ from nominal distribution with PDF p
3:     Initialize the set $\mathcal{Z} = \{\}$
4:     **for** $i = 1, \ldots, N$ **do**
5:         Evaluate surrogate model s_r at sample $\boldsymbol{z}_i$
6:         Evaluate limit state function g at surrogate model output $s_r(\boldsymbol{z}_i)$
7:         **if** $g(s_r(\boldsymbol{z}_i)) < 0$ **then**
8:             Add sample $\boldsymbol{z}_i$ to set $\mathcal{Z}$
9:         **end if**
10:     **end for**
11:     **if** $\mathcal{Z}$ is empty **then**
12:         **return** p
13:     **end if**
14:     Use EM algorithm to fit mixture ϕ_{MM} of k normal distributions to samples in $\mathcal{Z}$
15:     **return** ϕ_{MM}
16: **end procedure**

3.3. Step two: Evaluating the MFIS estimator with the high-fidelity model

In step two of the MFIS method, $M \in \mathbb{N}$ samples $\boldsymbol{z}_1', \ldots, \boldsymbol{z}_M'$ are drawn from the biasing distribution that was computed using Algorithm 1. The MFIS estimate of the failure probability is then defined as

$$P_f^{\mathrm{MFIS}}(\boldsymbol{z}_1', \ldots, \boldsymbol{z}_M') = \frac{1}{M} \sum_{i=1}^{M} \mathbb{I}_{\mathcal{G}}(\boldsymbol{z}_i') w(\boldsymbol{z}_i') \,. \tag{18}$$

The function $\mathbb{I}_{\mathcal{G}}$ in (18) is the indicator function with respect to the failure domain $\mathcal{G}$ of the high-fidelity model. Thus, the evaluation of $\mathbb{I}_{\mathcal{G}}$ at sample $\boldsymbol{z}_i'$ entails a high-fidelity model evaluation. The value of the indicator function $\mathbb{I}_{\mathcal{G}}$ at sample $\boldsymbol{z}_i'$ is $\mathbb{I}_{\mathcal{G}}(\boldsymbol{z}_i') = 1$ if the limit state function $g(s(\boldsymbol{z}_i')) < 0$ at the high-fidelity model output $s(\boldsymbol{z}_i')$, otherwise $\mathbb{I}_{\mathcal{G}}(\boldsymbol{z}_i') = 0$, see (3).

We emphasize that the surrogate model s_r is not used in (18). This allows us to show that the MFIS estimator (18) is unbiased.

Theorem 1. *Let q be the density of a biasing distribution constructed with Algorithm 1, and let Z' be the corresponding random variable. The MFIS estimator* (18) *is an unbiased estimator of the failure probability P_f.*

Proof. Let us consider the construction of the biasing distribution in Algorithm 1. We distinguish between two cases. In case one, the set $\mathcal{Z}$ in line 11 of Algorithm 1 is empty. Thus, none of the samples drawn from the nominal distribution fall in the failure domain $\mathcal{G}_r$ of the surrogate model, and the nominal PDF p is returned in line 12. In this case, the MFIS estimator $P_f^{\mathrm{MFIS}}(Z)$ becomes the Monte Carlo estimator (6) and is thus unbiased.

In case two, the set $\mathcal{Z}$ is not empty and Algorithm 1 returns a mixture of normal distributions. Since a mixture of normal distributions has global support, i.e., $q(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \in \mathbb{R}^d$, we obtain

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

Thus, the transformation

$$\mathbb{E}_p[\mathbb{I}_{\mathcal{G}}(Z)] = \int_{\mathbb{R}^d} \mathbb{I}_{\mathcal{G}}(\boldsymbol{z}) p(\boldsymbol{z}) \mathrm{d}\boldsymbol{z} = \int_{\mathbb{R}^d} \mathbb{I}_{\mathcal{G}}(\boldsymbol{z}') q(\boldsymbol{z}') \frac{p(\boldsymbol{z}')}{q(\boldsymbol{z}')} \mathrm{d}\boldsymbol{z}' = \mathbb{E}_q[\mathbb{I}_{\mathcal{G}_r}(Z') w(Z')]$$

from (2) to (8) holds [4] and therefore (18) is an unbiased estimator of P_f. This shows that $P_f^{\mathrm{MFIS}}(Z)$ is an unbiased estimator of the failure probability P_f. $\qquad\square$

Corollary 1. *Consider the same setting as in Theorem 1 but additionally assume that $\mathbb{E}_q[\mathbb{I}_{\mathcal{G}}(Z')^2 w(Z')^2] < \infty$ holds. Then, the RMSE of the MFIS estimator is*

$$e(P_f^{MFIS}) = \sqrt{\frac{\text{var}_q[\mathbb{I}_{\mathcal{G}}(Z')w(Z')]}{M}} \, . \tag{19}$$

Proof. This follows with the same arguments as for the importance sampling estimator discussed in Section 2.2. □

We emphasize that Theorem 1 holds even in the absence of accuracy guarantees on the surrogate model itself (e.g., via an estimate or bound on the error (13)).

3.4. The MFIS method

Algorithm 2 summarizes the two steps of the MFIS method. Inputs are the high-fidelity model s, the surrogate model s_r, the number of samples N to construct the biasing distribution, the number of samples M to use in the MFIS estimator, and the number k of normal distributions in the mixture model. Algorithm 2 first calls BIASINGDIST in Algorithm 1 to construct the biasing distribution using the surrogate model only. Then, the MFIS estimate is computed by evaluating the high-fidelity model at samples $z'_1, \ldots, z'_M$ drawn from the biasing distribution. Finally, the estimate $P_f^{\text{MFIS}}(z'_1, \ldots, z'_M)$ is returned.

3.5. Error and cost analysis of the MFIS method

With the RMSE $e(P_f^{\text{MFIS}})$ of the MFIS estimator P_f^{MFIS} given in (19), we find that

$$M \simeq \frac{\text{var}_q[\mathbb{I}_{\mathcal{G}}(Z')w(Z')]}{\epsilon^2} \tag{20}$$

samples are required to obtain an RMSE $e(P_f^{\text{MFIS}}) \leq \epsilon$, for a threshold parameter $0 < \epsilon \in \mathbb{R}$. Similarly to the costs of the importance sampling estimator, the costs $c(P_f^{\text{MFIS}})$ of the MFIS estimator to achieve $e(P_f^{\text{MFIS}}) \leq \epsilon$ are split into the costs of generating the biasing distribution and the costs of evaluating the high-fidelity model at the samples drawn from the biasing distribution.

We first consider the costs of generating the biasing distribution, i.e., the costs of Algorithm 1. In Algorithm 1 only the surrogate model is used, and thus the costs are independent of the costs b of evaluating the high-fidelity model s. Evaluating the limit state function, drawing samples from the nominal distribution, and fitting the mixture model typically introduces only a small amount of costs compared to the costs of the surrogate model evaluations. We therefore consider only the costs of the surrogate model evaluations here. Algorithm 1 requires N surrogate model evaluations. Let $0 < b_r \in \mathbb{R}$ denote the cost of one surrogate model evaluation. The costs of Algorithm 1, i.e., for the construction of the biasing distribution, are then

$$c_{\text{bias}}^{\text{MFIS}} = b_r N \, .$$

Second, the costs of evaluating the high-fidelity model at the samples of the biasing distribution are Mb.

Overall, the costs of the MFIS estimator to achieve an RMSE $e(P_f^{\text{MFIS}}) \leq \epsilon$ are therefore

$$c(P_f^{\text{MFIS}}) = c_{\text{bias}}^{\text{MFIS}} + bM \tag{21}$$
$$\simeq b_r N + b\frac{\text{var}_q[\mathbb{I}_{\mathcal{G}}(Z')w(Z')]}{\epsilon^2} \, ,$$

where we used (20).

We now compare the costs $c(P_f^{\text{MFIS}})$ of the MFIS estimator to the costs of the Monte Carlo estimator $c(P_f^{\text{MC}})$ for deriving an estimator with an RMSE below or equal to the threshold parameter ϵ. Let

$$\gamma = \frac{\text{var}_p[\mathbb{I}_{\mathcal{G}}(Z)]}{\text{var}_q[\mathbb{I}_{\mathcal{G}}(Z')w(Z')]} \, ,$$

9

Algorithm 2 Multifidelity importance sampling

1: **procedure** MULTIIS(s, s_r, p, N, M, k)
2:     Call BIASINGDIST(s_r, p, N, k) and obtain biasing distribution with PDF q
3:     Draw M samples $\boldsymbol{z}'_1, \ldots, \boldsymbol{z}'_M$ from biasing distribution
4:     Compute importance weights

$$\boldsymbol{w} = \left[\frac{p(\boldsymbol{z}'_1)}{q(\boldsymbol{z}'_1)}, \ldots, \frac{p(\boldsymbol{z}'_M)}{q(\boldsymbol{z}'_M)}\right]^T \in \mathbb{R}^d$$

5:     Evaluate high-fidelity model at samples $\boldsymbol{z}'_1, \ldots, \boldsymbol{z}'_M$ from the biasing distribution
6:     Obtain multifidelity importance sampling estimate $P_f^{\mathrm{MFIS}}(\boldsymbol{z}'_1, \ldots, \boldsymbol{z}'_M)$ via (18)
7:     **return** $P_f^{\mathrm{MFIS}}(\boldsymbol{z}'_1, \ldots, \boldsymbol{z}'_M)$
8: **end procedure**
---

be the ratio that represents the reduction of the variance achieved by the biasing distribution. Note that γ serves here as a theoretical quantity that is useful for analyzing our method but that is typically unavailable in practice. We write $c(P_f^{\mathrm{MFIS}})$ in terms of the costs of the Monte Carlo estimator as

$$c(P_f^{\mathrm{MFIS}}) = c_{\mathrm{bias}}^{\mathrm{MFIS}} + bM$$
$$\simeq b_r N + b \frac{\mathrm{var}_q[\mathbb{I}_{\mathcal{G}}(Z')w(Z')]}{\epsilon^2}$$
$$\simeq b_r N + b \frac{1}{\gamma} \frac{\mathrm{var}_p[\mathbb{I}_{\mathcal{G}}(Z)]}{\epsilon^2}$$
$$\simeq b_r N + \frac{1}{\gamma} c(P_f^{\mathrm{MC}}) \, .$$

Thus, the costs $c(P_f^{\mathrm{MFIS}})$ of the MFIS estimator are lower than the costs $c(P_f^{\mathrm{MC}})$ of the Monte Carlo estimator if the inequality

$$b_r N < \left(1 - \frac{1}{\gamma}\right) c(P_f^{\mathrm{MC}}) \tag{22}$$

holds. Further transforming (22) into

$$\frac{b_r N}{c(P_f^{\mathrm{MC}})} < 1 - \frac{1}{\gamma}$$

makes more explicit that a larger variance reduction, i.e., a larger γ, compensates higher costs $b_r N$ of generating the biasing distribution; however, the costs $b_r N$ of generating the biasing distribution must not exceed the costs $c(P_f^{\mathrm{MC}})$ of the Monte Carlo estimator.

4. Numerical results

This section demonstrates the MFIS method for failure probability estimation on four examples. We first consider a model of a plate in bending in Section 4.1, and a model of an acoustic horn in Section 4.2. We then apply the MFIS method to the time-dependent nonlinear Burgers' equation with uncertain boundary values in Section 4.3, and to a nonlinear model of a combustor in Section 4.4.

In all the following examples, a reference failure probability computed with the Monte Carlo estimator serves as a benchmark. The limit state function in the presented examples is cheap to evaluate and therefore the computational costs are dominated by the high-fidelity and surrogate model evaluations.

The examples were computed with MATLAB. Note that fitting a mixture of normal distributions with EM as needed for Algorithm 1 is available in many other software packages, including Python (sklearn.mixture.GMM) and R (EMCluster). We run the EM algorithm (gmdistribution.fit) with a random initial seeding. If the
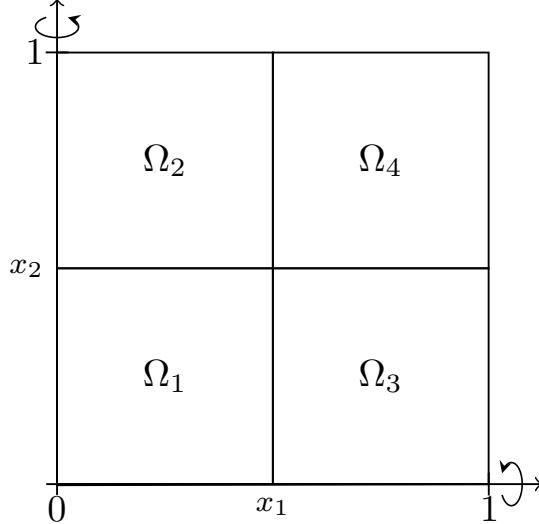
Figure 2: Plate problem: The eight inputs of the plate problem control the thickness and the load in the four subregions $\Omega_1, \ldots, \Omega_4$ of the spatial domain Ω.

algorithm converges to a mixture of normal distributions with singular covariance matrices, we repeat the algorithm with a different random initial seeding. The runtime measurements were performed on an Intel Xeon E5-1620 compute node with 32GB RAM on a single core.

4.1. Deflection of clamped plate in bending

We are interested in the maximum deflection of a clamped plate after a load has been applied [35]. The geometry of the plate is shown in Figure 2. The spatial domain $\Omega = [0,1]^2 \subset \mathbb{R}^2$ is partitioned into four subregions $\Omega_1, \ldots, \Omega_4$. Inputs define the thickness of the plate and the applied load in each subregion. The inputs are the components $\boldsymbol{z} = [z_1, \ldots, z_8]^T \in \mathbb{R}^8$ of a realization of the eight-dimensional input random variable Z, which has a uniform distribution in $[0.05, 0.1]^4 \times [1, 100]^4 \subset \mathbb{R}^8$. The first four inputs $z_1, \ldots, z_4$ define the thickness of the plate and the last four inputs $z_5, \ldots, z_8$ define the load, in the subregions $\Omega_1, \ldots, \Omega_4$, respectively. The output, y, is the maximum deflection of the plate in the spatial domain Ω. The limit state function is $g_{\text{plate}}(y) = -y + 5$, with the threshold 5, which means that a failure occurs if the maximum deflection is higher than 5.

Models. The (steady-state) high-fidelity model $s : \mathbb{R}^8 \to \mathbb{R}$ is derived with the finite element method following [36, Section 12.2, p. 161]. The number of degrees of freedom of the high-fidelity model is 19039. To construct the surrogate model, we sample 1000 inputs at which the high-fidelity model is evaluated to generate the snapshots. There are several sampling schemes available to guide the selection of the snapshots [37, 38, 39]; however, for this example sampling from a uniform distribution in the input domain is sufficient. The proper orthogonal decomposition (POD) is applied to the snapshots to derive the POD basis. The reduced operators are computed with Galerkin projection. This leads to surrogate models s_5 and s_{10} generated from five and ten POD basis vectors, respectively.

Reference failure probability. The reference failure probability $P_f^{\text{MC}} = 9.673 \times 10^{-3}$ is computed with the Monte Carlo estimator with 10^6 samples using the high-fidelity model. The RMSE of the Monte Carlo estimator is 9.787×10^{-5}, where the variance is estimated from the samples.

Illustration of PDFs. For illustration purposes, the estimated PDF of the random variables $s(Z), s_5(Z)$, and $s_{10}(Z)$ are compared in Figure 3. Each of the PDFs is estimated with kernel density estimation (KDE) from

(a) surrogate model, five POD modes

(b) surrogate model, ten POD modes
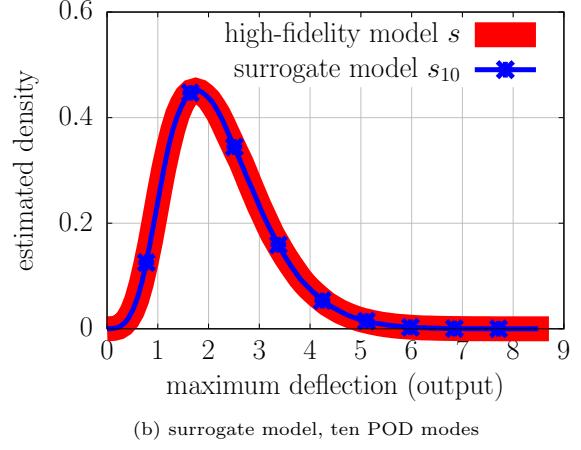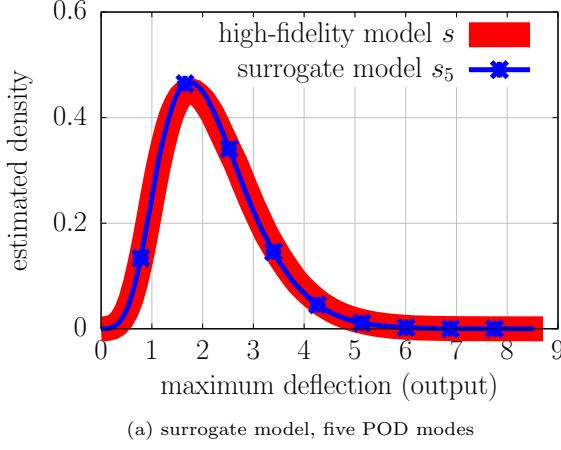
Figure 3: Plate model: The estimated PDFs corresponding to the surrogate model outputs match the PDF of the high-fidelity model outputs well.

(a) biasing distribution with $k = 8$

(b) biasing distribution with $k = 1$
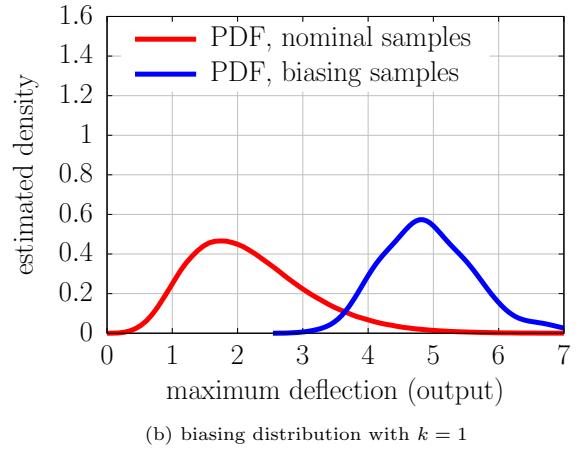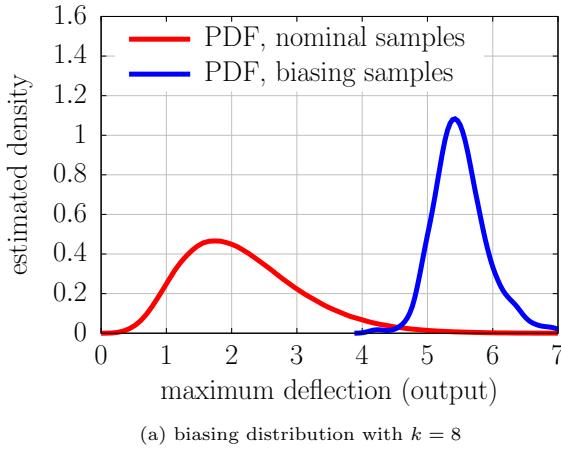
Figure 4: Plate model: The plots compare the estimated PDF of $s(Z)$ (nominal distribution) with $s(Z')$ (biasing distribution). The PDF corresponding to the biasing distribution shows that the samples from the biasing distribution tend to cluster near the threshold 5 of the limit state function.

12

(a) surrogate model s_5 with five POD modes   (b) surrogate model s_{10} with ten POD modes
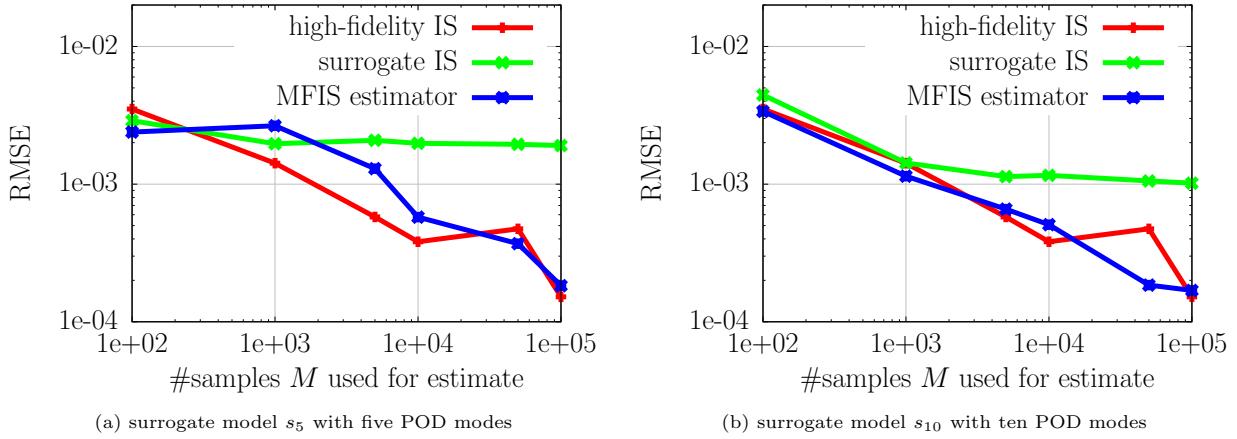
Figure 5: Plate model: Our multifidelity approach (MFIS estimator) shows a similar behavior as using the high-fidelity model only (high-fidelity IS estimator). Using the surrogate model only (surrogate IS estimator) leads to a biased estimator. The number of samples M is the number of samples used to derive the estimate in (18) for the MFIS method and in (9) for the high-fidelity and surrogate IS estimators.

1000 realizations of the corresponding random variable. The surrogate model s_5 with five POD modes is sufficient to derive a PDF that matches the PDF corresponding to the high-fidelity model well.

We generate a biasing distribution by constructing the biasing PDF with Algorithm 1 and the surrogate model s_5. The number of samples is set to $N = 10^6$, and the number of normal distributions in the mixture model is $k = 8$. Surrogate model evaluations are usually computationally cheap and therefore choosing a large N has only a minor impact on the overall costs of the MFIS method, see the discussion on runtime and speedup below. To illustrate the quality of the biasing distribution, we draw 1000 samples $\boldsymbol{z}'_1, \ldots, \boldsymbol{z}'_{1000}$ from the biasing distribution, i.e., realizations of Z', and estimate the PDF of $s(Z')$ from the outputs $s(\boldsymbol{z}'_1), \ldots, s(\boldsymbol{z}'_{1000})$. Figure 4a compares the PDF obtained from the samples from the biasing distribution to the PDF obtained from the samples from the nominal distribution, cf. Figure 3. As desired, the PDF corresponding to the samples from the biasing distribution is larger near the threshold 5 of the limit state function than the PDF corresponding to the nominal distribution.

Parameter selection. Plots such as Figure 4 give guidance on the selection of the parameter k. Biasing distributions are therefore created for several values of k and the corresponding estimated PDFs are plotted as, e.g., in Figure 4a for $k = 8$ and in Figure 4b for $k = 1$. The $N = 10^6$ samples computed with the surrogate model for the construction of the biasing distribution and the 1000 samples from the nominal distribution computed with the high-fidelity model are generated only once, and can be reused for different values of k. We note that the described approach for selecting k is a heuristic. Our MFIS method does not rely on a particular method for the selection of k and therefore other techniques that select the number of normal distributions in a mixture model in a fully automatic way are applicable as well. We refer to, e.g., [40] for an extensive survey of such techniques.

Comparison of estimators w.r.t. RMSE. We now compare three estimators:

- The *MFIS estimator* that uses the surrogate model to construct the biasing distribution with Algorithm 1, with $N = 10^6$ and $k = 8$, and that uses the high-fidelity model to derive the estimate of the failure probability. The MFIS estimator is unbiased.

- The *high-fidelity IS estimator* that uses the high-fidelity model only. The biasing distribution is constructed with Algorithm 1 using the high-fidelity model with the same number $N = 10^6$ of samples and the same number $k = 8$ of normal distributions in the mixture model as the MFIS estimator.

13

The estimate of the failure probability is derived with the high-fidelity model. The high-fidelity IS estimator is unbiased.

- The *surrogate IS estimator* that uses a surrogate model only. The biasing distribution is constructed as in the MFIS estimator. The estimate of the failure probability is computed using the surrogate model. The surrogate IS estimator is biased.

We first consider the RMSE of the three estimators. The RMSE of the MFIS estimator is estimated over $r = 10$ runs as

$$\sqrt{\frac{1}{r} \sum_{i=1}^{r} \left(P_f^{\text{MC}} - P_f^{\text{MFIS}} \right)^2}, \tag{23}$$

with respect to the reference failure probability P_f^{MC}, and analogously for the high-fidelity and surrogate IS estimators. The biasing PDFs of the respective estimators are computed once and then kept fixed over all runs. The biasing PDFs of the three estimators are derived from independent samples. Figure 5 plots the estimated RMSE of the MFIS estimator as a function of the number of samples M used for the estimate of the failure probability, and compares it to the estimated RMSEs of the high-fidelity and surrogate IS estimators. Figure 5a reports the RMSEs for surrogate model s_5 and Figure 5b for surrogate model s_{10}. The results confirm that the our multifidelity approach leads to a similar error decay as the high-fidelity IS estimator that is based on the high-fidelity model. The results also confirm that using the surrogate model only (surrogate IS estimator) leads to a biased estimator of the failure probability P_f.

Comparison of estimators w.r.t. runtime. Figure 6 compares the runtime of the MFIS method to the high-fidelity IS estimator. The runtime of the surrogate IS estimator is not included because it is a biased estimator. The costs of the MFIS method are $c(P_f^{\text{MFIS}}) = b_r N + bM$ because it requires N surrogate model evaluations for the construction of the biasing distribution and M high-fidelity model evaluations to derive the estimate, see (21). The high-fidelity IS estimator has costs $c(P_f^{\text{IS}}) = b(N + M)$, with N high-fidelity model evaluations to generate the biasing distribution and M evaluations to compute the estimate, see (11). The speedup of the MFIS method is $c(P_f^{\text{IS}})/c(P_f^{\text{MFIS}})$. To account for the costs of constructing the surrogate model, we additionally show the speedup

$$\frac{c(P_f^{\text{IS}})}{c(P_f^{\text{MFIS}}) + 1000b}, \tag{24}$$

which includes the costs $1000b$ of the 1000 high-fidelity model evaluations required to construct the surrogate model in this example. Note that the costs to construct the surrogate model are usually dominated by the costs to generate the snapshots; other costs that might arise during the construction of the surrogate model are ignored here. Note also that a surrogate model can be reused for computing failure probabilities that correspond to different limit state functions and thus the costs of building the surrogate model can be further amortized. The speedup $c(P_f^{\text{IS}})/c(P_f^{\text{MFIS}})$ of the MFIS method is shown in Figure 6a, where our multifidelity approach is between 10^2 and 10^4 times faster than the high-fidelity IS estimator. Note that for the same number of samples M the two estimators achieve a similar RMSE, see Figure 5. The MFIS method is between 10^2 and 10^3 times faster if the costs of the construction of the surrogate model (24) are included. The speedup results confirm that the costs of creating the snapshots, i.e., of creating the surrogate model, are quickly compensated here. Figure 6b shows the speedups with respect to the number of degrees of freedom of the high-fidelity model for $M = 100$ samples. As the high-fidelity model becomes more expensive to evaluate, the runtime improvements gained with our multifidelity approach increase.

Effect of number of samples N. Figure 7 demonstrates for s_5 that too few samples N for constructing the biasing distribution can degrade the performance of our MFIS estimator. The results confirm that increasing the number of samples N (in this example to $N = 5 \times 10^5$) leads to an MFIS estimator with a significantly lower RMSE error for the same number M, cf. the discussion on the number of samples N in Section 3.2. In

(b) speedup w.r.t. #samples M

(a) speedup w.r.t. degrees of freedom of the FOM
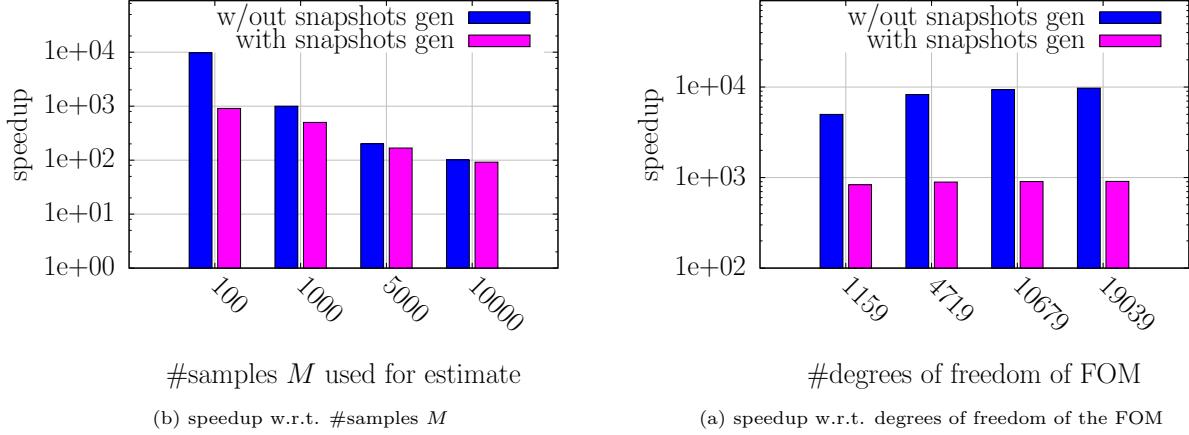
Figure 6: Plate model: If the surrogate model is already available, our MFIS method is between 10^2 and 10^4 times faster than using the high-fidelity model only (high-fidelity IS estimator), while still providing an unbiased estimator ("w/out snapshots gen"). The runtimes include the generation of the biasing distribution and the computation of the estimate but ignore the costs for building the surrogate model. If the runtime for building the surrogate model is included, speedups (24) of the MFIS method are between 10^2 and 10^3 ("with snapshots gen"); thus, the costs of building the surrogate model are compensated quickly here. The number of samples used for the construction of the biasing distribution is set to $N = 10^6$.
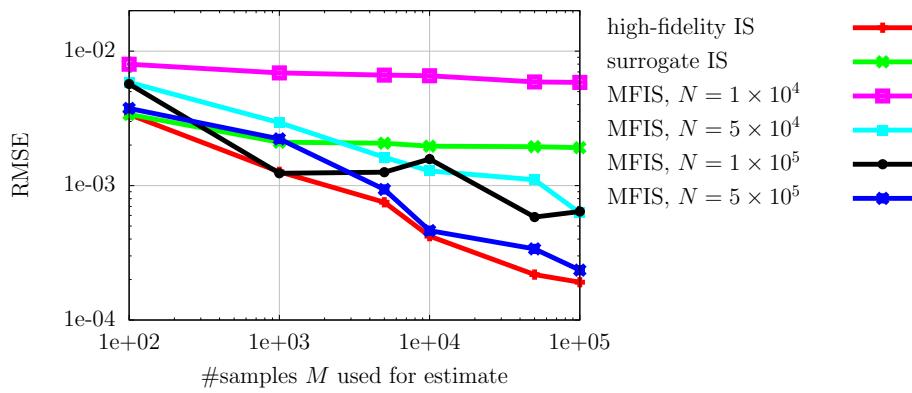
Figure 7: Plate model: If the number of samples N for constructing the biasing distribution is set too low, the constructed biasing PDF cannot capture well the failure boundary, and the RMSE of our MFIS estimator becomes higher than the RMSE of the high-fidelity IS estimator (see MFIS estimator for $N = 10^4 - 10^5$) for the same M; however, such a situation is usually avoided by increasing the number of samples N (in this example to $N = 5 \times 10^5$).

15

(a) surrogate model s_5 with five POD modes

(b) surrogate model s_{10} with ten POD modes
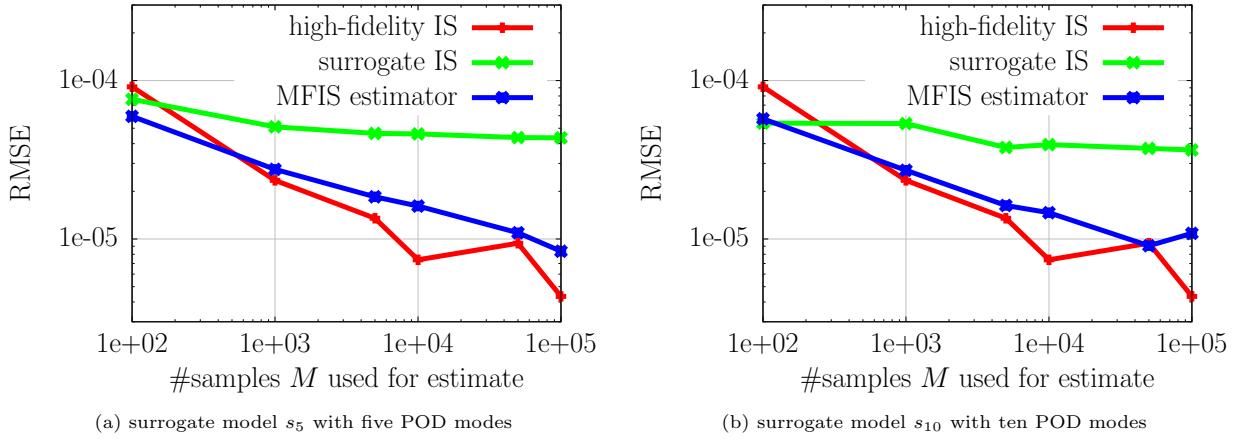
Figure 8: Plate model: Our MFIS method shows a similar behavior as the high-fidelity IS estimator even for limit state functions that lead to a failure probability that is almost two orders of magnitude smaller than in Figure 5.

the case of $N = 10^4$, only about 50 samples are in the failure domain of the surrogate model, whereas about 5000 samples are in the failure domain for $N = 5 \times 10^5$. We set $k = 6$ because in the case $N = 10^4$ there are too few samples in the failure domain for the EM algorithm to converge with $k = 8$ normal distributions in the mixture model. Note that our MFIS estimator is unbiased, independent of the number of samples N.

Smaller failure probability. To demonstrate that our MFIS method is applicable to the estimation of smaller failure probabilities, we consider the limit state function $g_{\text{plate}}^{\text{small}}(y) = -y + 7$. The corresponding reference failure probability is $P_f^{\text{small}} = 1.266 \times 10^{-4}$ with the RMSE 3.5579×10^{-6}. The failure probability P_f^{small} is almost two orders of magnitude smaller than P_f. We increase the number of samples for constructing the biasing PDF from 10^6 to $N = 10^7$ and keep the number of normal distributions $k = 8$ in the mixture model unchanged. Note that an increase in the number of samples N is necessary to ensure that sufficiently many samples are in the failure domain of the surrogate model for the construction of the biasing distribution with the EM algorithm; see previous paragraph for a discussion on the effect of N. Figure 8 compares the RMSE of the MFIS estimator to the high-fidelity and the surrogate IS estimator. Our MFIS estimator shows a similar behavior as the high-fidelity IS estimator.

4.2. Acoustic horn

In this example, the MFIS method is used with a surrogate model that is constructed with the reduced basis method [22].

We consider the acoustic horn model introduced in [41] that is based on the Helmholtz equation in the spatial domain $\Omega \subset \mathbb{R}^2$

$$\Delta u(\boldsymbol{x}) + \kappa^2 u(\boldsymbol{x}) = 0, \qquad \boldsymbol{x} \in \Omega,$$

where $u : \Omega \to \mathbb{C}$ is the solution function, Δ the two-dimensional Laplace operator, and $\kappa \in \mathbb{R}$ the wave number. Following [17], the first six of the $d = 9$ inputs control the geometry of the horn, the seventh input the wave number κ, and the eighth and ninth inputs the upper and lower wall impedance, respectively. The inputs are the components $\boldsymbol{z} = [z_1, \ldots, z_9]^T \in \mathbb{R}^9$ of a realization of the 9-dimensional random variable Z. The first six components of the random variable Z follow a uniform distribution in

$$[0.6785, 1.0357] \times [1.0357, 1.3929] \times [1.3929, 1.7500] \times$$
$$[1.7500, 2.1071] \times [2.1071, 2.4643] \times [2.4643, 2.8214] \subset \mathbb{R}^6 .$$

The seventh component follows a uniform distribution in $[1.3, 1.5] \subset \mathbb{R}$, and the last two components z_8, z_9 each follow a normal distribution with mean 50 and standard deviation 3, see [17]. The output, y, is

16

(a) accuracy (b) speedup
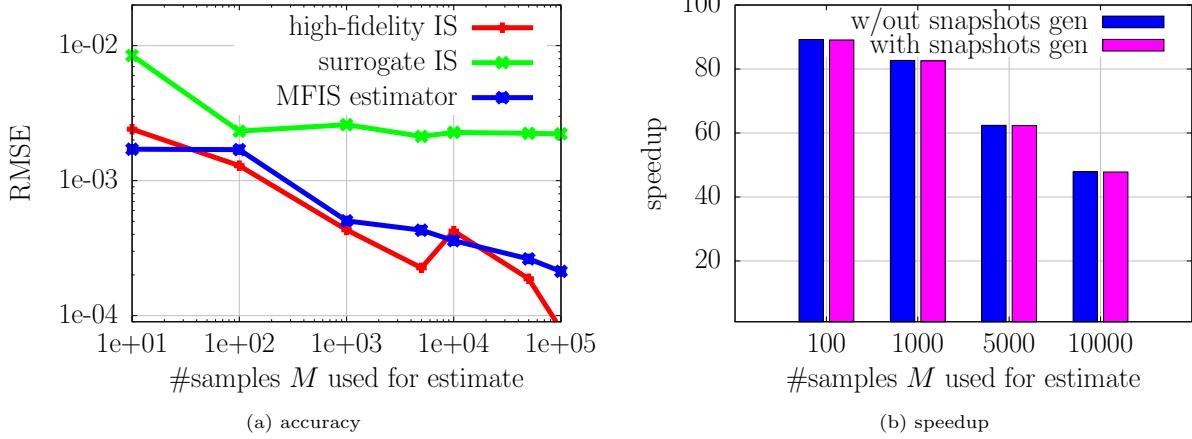
Figure 9: Horn model: The results in (a) show that the MFIS estimator leads to a similar behavior of the RMSE as the high-fidelity IS estimator. The speedups reported in (b) confirm that obtaining an estimate of the failure probability with our multifidelity approach is about 50 to 90 times faster than using the high-fidelity IS estimator.

the reflection coefficient, a measure of the horn's efficiency. We define that the horn fails if the reflection coefficient is above 0.32. The corresponding limit state function is $g_{\mathrm{horn}}(y) = -y + 0.32$.

The (steady-state) high-fidelity model $s : \mathbb{R}^9 \to \mathbb{R}$ is a finite element model with 35895 degrees of freedom [41]. The reduced basis method is used to determine 20 snapshots of the high-fidelity model that form the reduced basis [22]. The surrogate model is derived by projecting the equations onto the space spanned by the reduced basis, which therefore has 20 degrees of freedom. We refer to, e.g., [22] for a general introduction to the reduced basis method and to [41, 17] for details on the acoustic horn problem in the context of the reduced basis method.

We compute a reference failure probability $P_f^{\mathrm{MC}} = 2.858 \times 10^{-3}$ with the Monte Carlo estimator with 10^6 samples. The estimated RMSE of the Monte Carlo estimator is 5.338×10^{-5}.

We generate a biasing distribution with the surrogate model using $N = 10^6$ samples and $k = 8$ normal distributions in the mixture model. The parameter k was selected following the same approach as in Section 4.1. In Figure 9a, the estimated RMSE (23) of the MFIS estimator is compared to the estimated RMSEs of the high-fidelity and surrogate IS estimators, cf. Section 4.1. The estimated RMSE corresponding to the surrogate IS estimator is quickly dominated by the accuracy of the surrogate model. Our multifidelity approach shows a similar behavior as the high-fidelity IS estimator; however, the MFIS method achieves a between 50 and 90 times lower runtime, see Figure 9b. The costs of the high-fidelity model evaluations required for the construction of the surrogate model are low here because the reduced basis method relies on *a posteriori* error estimators [22] to derive the reduced basis, and thus only 20 high-fidelity model evaluations are necessary to obtain the surrogate model in this example [22, 41].

4.3. Burgers' equation with uncertain values at the boundary

This examples demonstrates that a surrogate model can be sufficiently accurate for constructing a biasing distribution even if the surrogate model provides only poor approximations of the high-fidelity model outputs.

Let $\Omega = [-1, 1] \subset \mathbb{R}$ be the spatial domain, $T \in \mathbb{R}$ the end time, and $[0, T] \subset \mathbb{R}$ the time domain. We solve for the solution $u : \Omega \times [0, T] \to \mathbb{R}$ of the time-dependent viscous Burgers' equation

$$\partial_t u(x,t) + u \partial_x u(x,t) = \nu \partial_x^2 u(x,t), \qquad x \in \Omega, t \in [0, T], \tag{25}$$

where ∂_t, ∂_x, and ∂_x^2 are the first derivative in the time coordinate, and the first and second derivative in the spatial coordinate, respectively. The parameter $\nu = 0.05$ is the viscosity parameter. We impose Dirichlet boundary conditions

$$u(-1, t) = 1 + z, \quad u(1, t) = -1, \qquad t \in [0, T],$$

17

|	
(a) nominal distribution	(b) biasing distribution
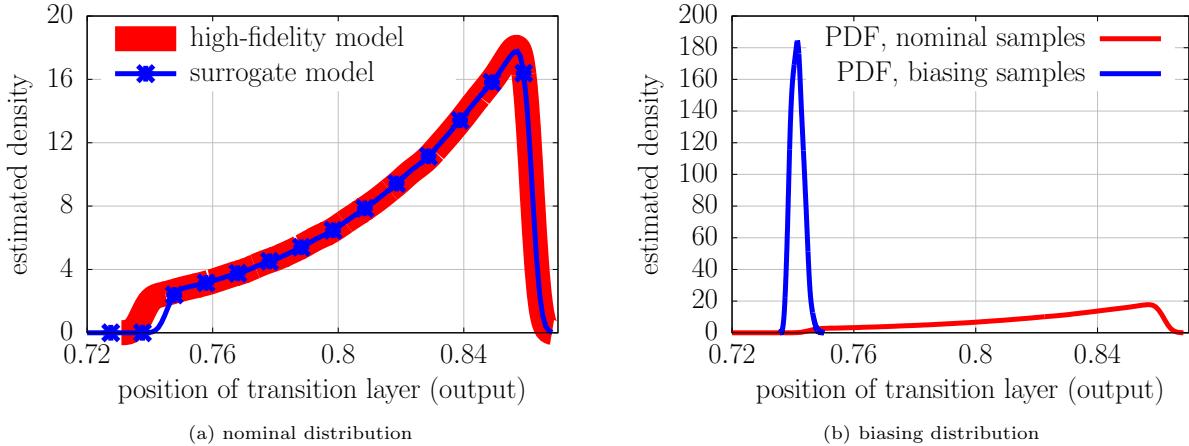
Figure 10: Burgers' equation: In (a), the estimated PDF of the high-fidelity model output $s(Z)$ and the surrogate model output $s_r(Z)$. The PDF corresponding to the surrogate model does not capture the sharp kink near the threshold 0.74; however, the surrogate model is still sufficiently accurate for the construct of a biasing distribution with Algorithm 1, see (b).

where $z \in \mathbb{R}$ is a perturbation. We follow [20] and define the input of our problem as the perturbation z. The input z is a realization of a one-dimensional ($d = 1$) random variable Z. The random variable Z is uniformly distributed in $[10^{-2}, 10^{-1}] \subset \mathbb{R}$. The output is the position $y \in \Omega$ of the transition layer with $u(y, T) = 0$. Note that y is highly sensitive to z [42]. The system fails if the output is below the threshold 0.74, i.e., the limit state function is $g_{\mathrm{Burgers}}(y) = y - 0.74$.

The Burgers' equation (25) is discretized with finite differences on an equidistant grid in Ω with 514 grid points. The time domain is discretized with a semi-implicit Euler method and time step size 10^{-1}. The time stepping is stopped either after 10^5 time steps, i.e., if $T = 10^4$ is reached, or if the L^2 norm of the difference of two successive state vectors is smaller than 10^{-12}. The high-fidelity model is the map $s : \mathbb{R} \to \mathbb{R}$. We evaluate the high-fidelity model at 200 inputs $z_1, \dots, z_{200}$ drawn uniformly in $[10^{-2}, 10^{-1}] \subset \mathbb{R}$, and store the state vector at every 100-th time step. This leads to 2415 snapshots from which we compute a POD basis with 20 basis vectors. The reduced operator corresponding to the diffusion term is derived by Galerkin projection. The corresponding surrogate model is denoted as s_r.

The reference failure probability $P_f^{\mathrm{MC}} = 5.49 \times 10^{-3}$ is computed with the Monte Carlo estimator and 10^6 samples, using the high-fidelity model only. The estimated RMSE of the Monte Carlo estimator is 7.389×10^{-5}.

We now illustrate that the surrogate model is sufficiently accurate to derive a biasing distribution, even though it leads to a poor approximation of the limit state function at the threshold 0.74. The PDFs of $s(Z)$ and $s_r(Z)$ are estimated from 1000 samples each and are shown in Figure 10a. The surrogate model leads to a similar PDF as the high-fidelity model; however, near the threshold 0.74 the PDF corresponding to the surrogate model does not have the sharp kink of the PDF corresponding to the high-fidelity model. Still, Figure 10b confirms that Algorithm 1 with the surrogate model leads to a biasing distribution that puts many samples near the failure boundary. The number of normal distributions in the mixture model is set to $k = 1$ and the number of samples to construct the biasing distribution is $N = 10^6$.

Figure 11a reports the estimated RMSE (23) corresponding to the MFIS, the high-fidelity IS, and the surrogate IS estimator. The estimated RMSE corresponding to the surrogate IS estimator stays high even if the number of samples M is increased. The estimated RMSE of the MFIS estimator is slightly offset from the estimated RMSE of the high-fidelity IS estimator but decays at a similar rate. The speedup of the MFIS method is reported in Figure 11b. The MFIS method achieves speedups of about 14 in this example. Note that only 200 high-fidelity model evaluations are used during the construction of the surrogate model, and therefore the costs of the snapshot generation, i.e., the costs of building the surrogate model, are compensated.

(a) accuracy

(b) speedup
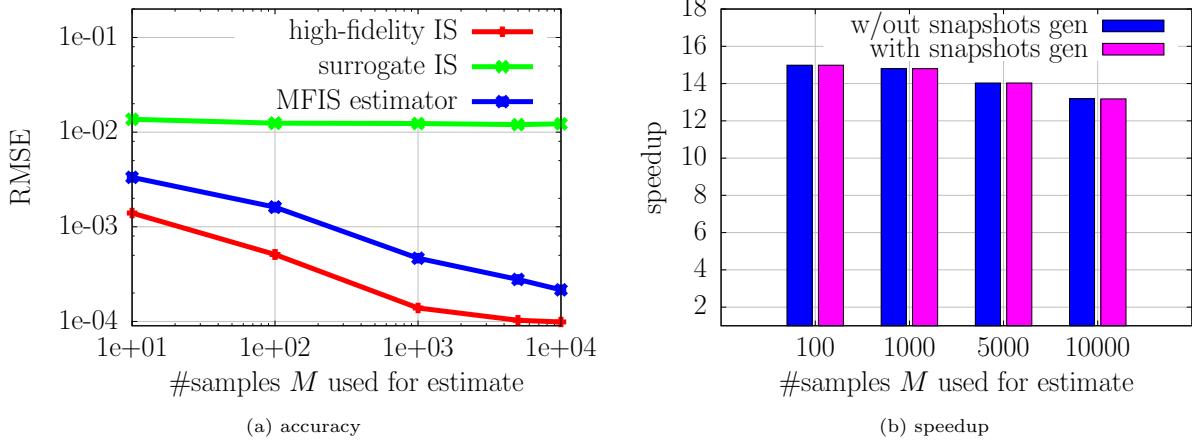
Figure 11: Burgers' equation: The RMSE of the surrogate IS estimator is limited by the poor approximation quality of the surrogate model; however, the surrogate model is sufficiently accurate to construct a biasing distribution so that the MFIS estimator shows a similar behavior as the computationally more expensive high-fidelity IS estimator.

4.4. Combustor model

In this example, a surrogate model with a separate approximation of the nonlinear term with the discrete empirical interpolation method (DEIM) [43] is used in the MFIS method.

We consider the simplified combustor model introduced in [44]. The reaction underlying the model is

$$2H_2 + O_2 \rightarrow 2H_2O,$$
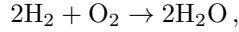
where H_2 is the fuel, O_2 the oxidizer, and H_2O the product. The governing equation is a nonlinear advection-diffusion-reaction equation with $d = 2$ inputs, which are the activation energy and the pre-exponential factor in an Arrhenius-type nonlinear reaction source term. The inputs are the components $\boldsymbol{z} = [z_1, z_2]^T \in \mathbb{R}^2$ of a realization of a two-dimensional random variable Z, with a uniform distribution in

$$[5.5 \times 10^{11}, 1.5 \times 10^{13}] \times [1.5 \times 10^3, 9.5 \times 10^3] \subset \mathbb{R}^2. \tag{26}$$

The output, y, is the maximum temperature. The combustor fails if the maximum temperature exceeds 2420K. The corresponding limit state function is $g_{\text{combustor}}(y) = -y + 2240$. We refer to [44] for details of the model.

The high-fidelity model $s : \mathbb{R}^2 \rightarrow \mathbb{R}$ and the surrogate model s_r are derived following [44]. The high-fidelity model is based on a finite difference discretization of the governing equation that leads to 10804 degrees of freedom. The discrete system of nonlinear equations is solved with the Newton method. POD is applied to 2500 snapshots corresponding to an equidistant 50×50 grid in the domain (26). The surrogate model s_r is derived from the POD basis with four basis vectors and Galerkin projection. The nonlinear term is approximated using DEIM [43, 45], where the DEIM basis is generated by applying POD to the nonlinear reaction source term evaluated at the 2500 snapshots. The DEIM approximation uses four DEIM basis vectors and four DEIM interpolation points. We refer to [44] for details on the derivation of the surrogate model.

The reference failure probability $P_f^{\text{MC}} = 4.560 \times 10^{-3}$ is computed with the Monte Carlo estimator and 10^6 samples. The estimated RMSE of the failure probability estimator is 6.737×10^{-5}.

The surrogate model is used to construct a biasing distribution with Algorithm 1, with $k = 1$ and $N = 10^6$, cf. Section 4.1. In Figure 12a the estimated RMSEs (23) of the MFIS estimator and the high-fidelity and surrogate IS estimators are reported. The results confirm that the surrogate IS estimator leads to a high estimated RMSE that levels off after 100 samples; however, the surrogate model is sufficiently accurate to construct a suitable biasing distribution. The speedup of the MFIS method compared to high-fidelity IS

19

|(a) accuracy|(b) speedup|
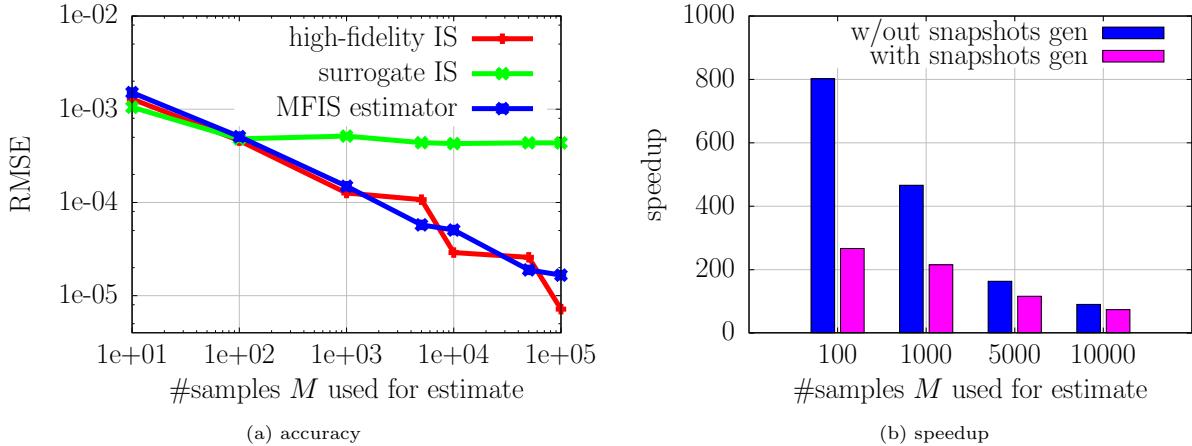
Figure 12: Combustor model: The surrogate model with four POD and four DEIM modes is a poor approximation of the high-fidelity model but it is sufficiently accurate to construct a biasing distribution such that the MFIS estimator achieves a similar estimated RMSE as the high-fidelity IS estimator. The multifidelity approach achieves speedups between 100 and 800 compared to using the high-fidelity model only.

estimator is shown in Figure 12b. Our multifidelity approach is between 100 and 800 times faster than using the high-fidelity model only. We note that the costs of building the surrogate model are compensated here again.

5. Conclusions and future work

This paper introduced a multifidelity approach for failure probability computation with the Monte Carlo method and importance sampling. Our MFIS method employs a surrogate model to construct the biasing distribution and the high-fidelity model to compute the estimate. We have shown that this leads to an unbiased estimator of the failure probability. Numerical experiments on a wide range of linear and nonlinear problems demonstrated the significant runtime improvements gained with our MFIS method compared to the Monte Carlo method with importance sampling that invokes a single model only.

The key property of our MFIS method is that it speeds up the runtime of the Monte Carlo method by using a surrogate model, but it leads to an unbiased estimator even in the absence of accuracy guarantees on the surrogate model. Our multifidelity method shows that, in some situations, recourse to the high-fidelity model can compensate for the lack of *a priori* error bounds or *a posteriori* error estimators for the surrogate model. Furthermore, it demonstrates that accuracy guarantees on the surrogate model are not a prerequisite to derive guarantees on the Monte Carlo estimate. This is particularly useful in complex problem settings, for example in computational fluid dynamics and fluid-structure interaction, where surrogate models can often be derived but error bounds or error estimators for these surrogate models are unavailable due to a lack of mathematical theory.

For smaller failure probabilities (i.e., in the range of 10^{-9}), the number of samples in the failure domain of the surrogate model typically decreases significantly such that fitting a mixture model with the EM algorithm fails. Future work therefore includes an iterative approach, similar to the cross entropy method [7, 8], where the biasing distribution is repeatedly refined using samples of the biasing distribution of the previous iteration.

Acknowledgment

as part of the DiaMonD Multifaceted Mathematics Integrated Capability Center. Some of the numerical examples were computed on the computer cluster of the Munich Centre of Advanced Computing.

References

[1] J. Li, D. Xiu, Evaluation of failure probability via surrogate models, Journal of Computational Physics 229 (23) (2010) 8966–8980.

[2] S. Engelund, R. Rackwitz, A benchmark study on importance sampling techniques in structural reliability, Structural Safety 12 (4) (1993) 255–276.

[3] R. Rackwitz, Reliability analysis—a review and some perspectives, Structural Safety 23 (4) (2001) 365–395.

[4] C. Robert, G. Casella, Monte Carlo Statistical Methods, Springer, 2004.

[5] G. Fishman, Monte Carlo, Springer, 1996.

[6] R. Srinivasan, Importance Sampling, Springer, 2002.

[7] R. Rubinstein, Optimization of computer simulation models with rare events, European Journal of Operational Research 99 (1) (1997) 89–112.

[8] P.-T. d. Boer, D. Kroese, S. Mannor, R. Rubinstein, A tutorial on the cross-entropy method, Annals of Operations Research 134 (1) (2005) 19–67.

[9] O. Cappé, A. Guillin, J. M. Marin, C. Robert, Population Monte Carlo, Journal of Computational and Graphical Statistics 13 (4) (2004) 907–929.

[10] R. Douc, A. Guillin, J.-M. Marin, C. Robert, Convergence of adaptive mixtures of importance sampling schemes, The Annals of Statistics 35 (1) (2007) 420–448.

[11] R. Douc, A. Guillin, J.-M. Marin, C. Robert, Minimum variance importance sampling via population Monte Carlo, ESAIM: Probability and Statistics 11 (2007) 427–447.

[12] O. Cappé, R. Douc, A. Guillin, J.-M. Marin, C. Robert, Adaptive importance sampling in general mixture classes, Statistics and Computing 18 (4) (2008) 447–459.

[13] J.-M. Cornuet, J.-M. Marin, A. Mira, C. Robert, Adaptive multiple importance sampling, Scandinavian Journal of Statistics 39 (4) (2012) 798–812.

[14] B. Nelson, On control variate estimators, Computers & Operations Research 14 (3) (1987) 219–225.

[15] M. Giles, Multi-level Monte Carlo path simulation, Operations Research 56 (3) (2008) 607–617.

[16] K. Cliffe, M. Giles, R. Scheichl, A. Teckentrup, Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients, Computing and Visualization in Science 14 (1) (2011) 3–15.

[17] L. Ng, K. Willcox, Multifidelity approaches for optimization under uncertainty, International Journal for Numerical Methods in Engineering 100 (10) (2014) 746–772.

[18] L. Ng, K. Willcox, Monte-Carlo information-reuse approach to aircraft conceptual design optimization under uncertainty, Journal of Aircraft (2015) 1–12.

[19] A. Litvinenko, H. Matthies, T. El-Moselhy, Sampling and low-rank tensor approximation of the response surface, in: J. Dick, F. Y. Kuo, G. W. Peters, I. H. Sloan (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2012, Vol. 65 of Springer Proceedings in Mathematics & Statistics, Springer, 2013, pp. 535–551.

[20] J. Li, J. Li, D. Xiu, An efficient surrogate-based method for computing rare failure probability, Journal of Computational Physics 230 (24) (2011) 8683–8697.

[21] P. Chen, A. Quarteroni, Accurate and efficient evaluation of failure probability for partial different equations with random input data, Computer Methods in Applied Mechanics and Engineering 267 (2013) 233–260.

[22] G. Rozza, D. Huynh, A. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations, Archives of Computational Methods in Engineering 15 (3) (2008) 229–275.

[23] M. Grepl, Y. Maday, N. C. Nguyen, A. Patera, Efficient reduced-basis treatment of nonaffine and nonlinear partial differential equations, ESAIM: Mathematical Modelling and Numerical Analysis 41 (03) (2007) 575–605.

[24] J. Bucklew, Introduction to Rare Event Simulation, Springer, 2003.

[25] C. Rasmussen, C. Williams, Gaussian Processes for Machine Learning, MIT Press, 2006.

[26] A. Ionita, A. Antoulas, Data-driven parametrized model reduction in the Loewner framework, SIAM Journal on Scientific Computing 36 (3) (2014) A984–A1007.

[27] B. Peherstorfer, S. Zimmer, H.-J. Bungartz, Model reduction with the reduced basis method and sparse grids, in: J. Garcke, M. Griebel (Eds.), Sparse Grids and Applications, Vol. 88 of Lecture Notes in Computational Science and Engineering, Springer, 2013, pp. 223–242.

[28] S. Gugercin, A. Antoulas, A survey of model reduction by balanced truncation and some new results, International Journal of Control 77 (8) (2004) 748–766.

[29] P. Mills, Efficient statistical classification of satellite measurements, International Journal of Remote Sensing 32 (21) (2011) 6109–6132.

[30] B. Peherstorfer, D. Pflüger, H.-J. Bungartz, Density Estimation with Adaptive Sparse Grids for Large Data Sets, SIAM, 2014, Ch. 50, pp. 443–451.

[31] B. W. Silverman, Density Estimation for Statistics and Data Analysis, Chapman and Hall, 1986.

[32] C. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.

[33] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, Springer, 2009.

[34] K. Murphy, Machine Learning, MIT Press, 2012.

[35] B. Peherstorfer, K. Willcox, Dynamic data-driven reduced-order models, Computer Methods in Applied Mechanics and Engineering 291 (2015) 21–41.

[36] A. Ferreira, MATLAB Codes for Finite Element Analysis, Springer, 2008.

[37] C. Prud'homme, Y. Maday, A. T. Patera, G. Turinici, D. V. Rovas, K. Veroy, L. Machiels, Reliable Real-Time Solution of Parametrized Partial Differential Equations: Reduced-Basis Output Bound Methods, Journal of Fluids Engineering 124 (1) (2001) 70–80.

[38] T. Bui-Thanh, K. Willcox, O. Ghattas, Model reduction for large-scale systems with high-dimensional parametric input space, SIAM Journal on Scientific Computing 30 (6) (2008) 3270–3288.

[39] B. Haasdonk, Convergence rates of the POD-Greedy method, ESAIM: Mathematical Modelling and Numerical Analysis 47 (2013) 859–873.

[40] S. Salvador, P. Chan, Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms, in: Tools with Artificial Intelligence, 2004. ICTAI 2004. 16th IEEE International Conference on, 2004, pp. 576–584.

[41] J. Eftang, D. Huynh, D. Knezevic, A. Patera, A two-step certified reduced basis method, Journal of Scientific Computing 51 (1) (2012) 28–58.

[42] D. Xiu, G. Karniadakis, Supersensitivity due to uncertain boundary conditions, International Journal for Numerical Methods in Engineering 61 (12) (2004) 2114–2138.

[43] S. Chaturantabut, D. Sorensen, Nonlinear model reduction via discrete empirical interpolation, SIAM Journal on Scientific Computing 32 (5) (2010) 2737–2764.

[44] M. Buffoni, K. Willcox, Projection-based model reduction for reacting flows, in: 40th Fluid Dynamics Conference and Exhibit, Fluid Dynamics and Co-located Conferences, American Institute of Aeronautics and Astronautics, 2010, pp. 1–14.

[45] B. Peherstorfer, D. Butnaru, K. Willcox, H. Bungartz, Localized discrete empirical interpolation method, SIAM Journal on Scientific Computing 36 (1) (2014) A168–A192.