Multifidelity Cross-Entropy Estimation of Conditional Value-at-Risk for Risk-Averse Design Optimization

Anirban Chaudhuri*
Massachusetts Institute of Technology, Cambridge, MA 02139

Benjamin Peherstorfer[†]
Courant Institute of Mathematical Sciences, New York University, New York, NY 10012

Karen Willcox[‡]
University of Texas at Austin, Austin, TX 78712

There is an increasing demand for designs in aerospace engineering that guarantee baseline performance even in limit states, i.e., outside of nominal operating conditions of vehicles. Typically, the numerical optimization for such risk-averse designs is computationally challenging because in each iteration of the optimization loop the performances of the designs are estimated for the rare events corresponding to the limit states. This work proposes a multifidelity approach to make tractable the optimization of large-scale risk-averse designs that are based on the conditional value-at-risk (CVaR) as the risk measure. The multifidelity method leverages low-cost, low-fidelity models to speed up the CVaR estimation in each iteration of the risk-averse optimization to reduce the runtime compared to traditional Monte Carlo estimators that rely on the high-fidelity models alone. At the same time, the proposed approach makes occasional recourse to the expensive high-fidelity model to guarantee convergence to design points that satisfy the high-fidelity optimality conditions. In numerical experiments with an aerostructural design problem, the multifidelity approach achieves speedups of almost one order of magnitude compared to a traditional single-fidelity method.

I. Introduction

To meet the ever increasing demands on robustness and reliability in aerospace engineering, it becomes necessary to optimize for risk-averse designs that guarantee baseline performance even in limit states, i.e., outside of nominal operating conditions. The objective functions in such risk-averse optimization problems incorporate uncertainties via random variables and then build on risk measures to quantify the value of the objective in the tails of the distributions corresponding to limit states. However, estimating risk measures with, e.g., standard Monte Carlo methods typically requires large numbers of realizations of the random variables per objective function evaluation, at which then a model of the system has to be evaluated. Thus, if the model is expensive-to-evaluate, then risk-averse optimization quickly becomes intractable. We present a multifidelity approach [1] that leverages low-cost, low-fidelity models to speed up risk-averse optimization and that uses occasional recourse to the computationally expensive, high-fidelity model to establish accuracy guarantees.

Low-fidelity models provide computationally cheap approximations of high-fidelity model outputs. In many situations, the runtime of low-fidelity models is orders of magnitudes lower than the runtime of high-fidelity models. Following [1], one distinguishes between three types of low-fidelity models. The first type are data-fit surrogate models [2, 3], e.g., kriging models [4–8], that are learned from input-output data of the high-fidelity models. Second, there are projection-based reduced models that solve for the states of the high-fidelity models in low-dimensional solutions spaces [9–13]. Third, there are simplified-physics models that are derived from high-fidelity models by exploiting domain knowledge [14]. A typical approach to speed up optimization with low-fidelity models is to construct a low-fidelity model with one-time high costs that satisfies the accuracy requirements of the problem at hand and then to use the low-fidelity model to replace the high-fidelity model in the optimization loop. However, just replacing the high-fidelity model with the low-fidelity model introduces the error of the low-fidelity model into the optimization result. In particular,

---

*Postdoctoral Associate, Department of Aeronautics and Astronautics. AIAA member.
[†]Assistant Professor, Courant Institute of Mathematical Sciences.
[‡]Professor, Aerospace Engineering and Engineering Mechanics, Oden Institute for Computational Engineering and Sciences. Fellow AIAA.

this means that the optimizer converges to a design point that satisfies the optimality conditions with respect to the low-fidelity model, instead of the high-fidelity model.

Multifidelity methods combine, instead of replace, the high-fidelity model with one or multiple low-fidelity models [1]. The low-fidelity models are leveraged for speedup and occasional recourse to the high-fidelity model establishes convergence to an optimum that satisfies the optimality conditions with respect to the high-fidelity model. Thus, multifidelity methods rely on multiple models and use them in concert to speed up computations while establishing the same accuracy guarantees as methods that use the high-fidelity model alone. Multifidelity methods have a long tradition in deterministic optimization. The works [15, 16] establish a multifidelity trust region framework and formulate a first-order consistency requirement on the low-fidelity model. Another multifidelity method for optimization is efficient global optimization (EGO) [17, 18], which adapts a low-fidelity Gaussian process model to the high-fidelity model. The work [3] gives a survey on multifidelity methods in optimization. The survey [1] gives an overview about multifidelity methods in optimization, uncertainty propagation, and statistical inverse problems.

We propose a multifidelity method for risk-averse optimization. We build on the conditional value-at-risk (CVaR), which is a common risk measures in financial mathematics [19–22] and engineering applications [23–25]. Estimating CVaR is computationally demanding for two reasons. First, estimating CVaR typically requires a large number of model evaluations, which quickly becomes computationally intractable if each model evaluation is expensive. Second, since CVaR targets rare and low-probability events, standard Monte Carlo estimators of CVaR are often inefficient because a large number of samples is required to accurately estimate the tail behavior. There are several multifidelity methods for estimating rare event probabilities, e.g., the work [26, 27] uses general polynomial chaos low-fidelity models, the work [28] uses reduced-basis models and exploits the availability of error estimators, and [29–31] builds on a multilevel hierarchy of coarse-grid approximations. In [32], reduced models are used together with importance sampling to efficiently estimate CVaR. In contrast, the works [33, 34] are applicable to general types of low-fidelity models. In the context of reliability-based design optimization, there are multifidelity methods [35, 36] that reuse information from previous iterations to obtain speedups compared to single-fidelity methods [37]. We build on the multifidelity pre-conditioned cross-entropy (MFCE) method [38], which is based on importance sampling and the cross-entropy method [39–41] and was developed for estimating rare event probabilities. MFCE preconditions the underlying estimation problem with low-fidelity models such that only few evaluations of the high-fidelity models are required at the final step. In this work, we show that MFCE can be used to efficiently estimate the CVaR from model outputs to reduce the costs of risk-averse optimization. Our numerical results demonstrate speedups of up to two orders of magnitude compared to single-fidelity approaches that use the high-fidelity model alone.

This paper is organized as follows. Notation to formulate the optimization problems of interest is introduced in Section II. Section III formulates risk-averse optimization with CVaR and summarizes that risk-averse optimization based on standard Monte Carlo methods quickly becomes intractable if model evaluations are computationally expensive. The proposed multifidelity method that leverages low-cost, low-fidelity models to address these computational challenges is presented in Section IV. Numerical experiments in Section V demonstrate that the proposed approach achieves speedups between one and two orders of magnitude compared to a standard Monte Carlo estimator that uses the high-fidelity model alone. Conclusions are drawn in Section VI.

II. Deterministic optimization

In a typical situation, a potentially computationally expensive high-fidelity model of the system of interest is available, which describes the system behavior for a given design point. Let $x \in \mathcal{X}$ be a design point in the design domain $\mathcal{X}$. The high-fidelity model that takes design variables as the input is denoted by $f : \mathcal{X} \to \mathbb{R}$. The function value $f(x)$ at a design point $x \in \mathcal{X}$ is the quantity of interest. The performance of a design x is assessed by the objective $J : \mathcal{X} \to \mathbb{R}$. Evaluating J at a design point typically entails at least one evaluation of the model f. The goal of deterministic optimization is to find the optimal design $x^* \in \mathcal{X}$ that solves

$$\min_{x \in \mathcal{X}} \quad J(x) \, . \tag{1}$$

Numerical optimization methods are typically of iterative nature, which means that the optimization can be imagined as a loop. In each iteration of the loop, at least one evaluation of the objective J, and thus of the high-fidelity model f, is necessary. Figure 1a depicts the optimization loop. If one evaluation of f incurs costs $0 < c \in \mathbb{R}$ and if $K \in \mathbb{N}$ iterations are performed by the optimizer, then the total costs of the optimization are at least $c \times K$.

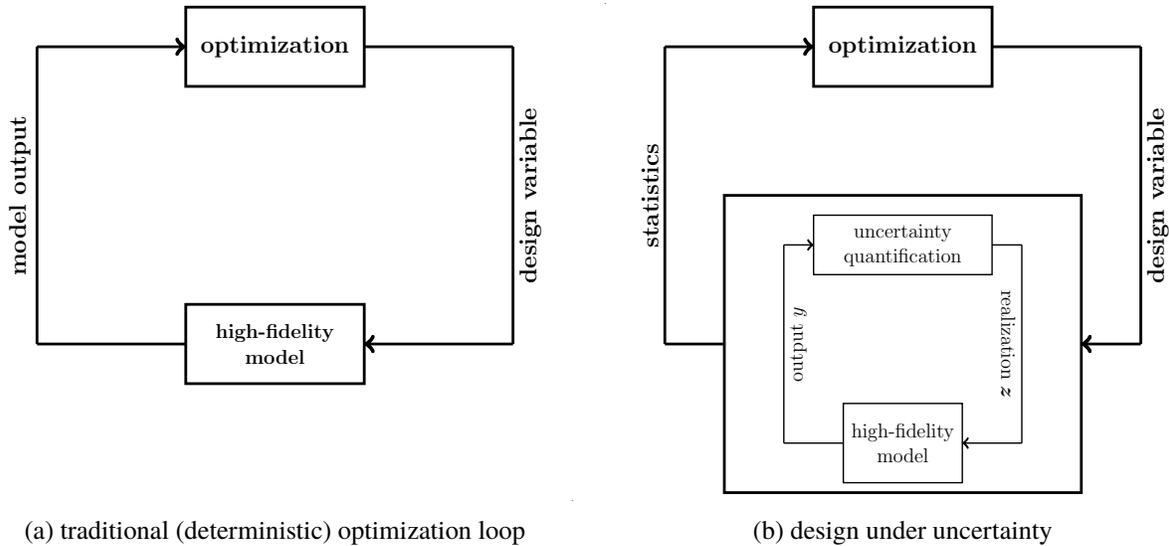(a) traditional (deterministic) optimization loop   (b) design under uncertainty

Fig. 1   Optimization methods are typically of iterative nature and therefore can be depicted as loops. In each iteration, at least one evaluation of the objective, and thus of the high-fidelity model, is necessary. In case of risk-averse design, an uncertainty quantification loop is embedded within each optimization iteration. Thus, the computational costs of risk-averse designs can be orders of magnitude higher than the costs of traditional (deterministic) design.

III. Risk-Averse Optimization based on the Conditional Value-at-Risk (CVaR)

Consider now the situation where the high-fidelity model f depends on the design variable $\boldsymbol{x} \in \mathcal{X}$ and, additionally, on a random variable $\Xi : \Omega \to \mathcal{Z}$, with sample space Ω. Thus, the model is a function $f : \mathcal{X} \times \mathcal{Z} \to \mathbb{R}$ that maps $\boldsymbol{x}$ and a realization $\boldsymbol{\xi} \in \mathcal{Z}$ of the random variable Ξ onto $\mathbb{R}$. This means that the quantity of interest of the high-fidelity model $f(\boldsymbol{x}, \Xi)$ becomes a random variable. The random variable Ξ models uncertainties in the system, e.g., model measurement errors, unknown environments conditions, and manufacturing variations. We now consider risk-averse optimization, where the objective depends on a risk measure that quantifies statistics corresponding to the tails of the distribution of the random variable $f(\boldsymbol{x}, \Xi)$.

A. CVaR risk measure

The risk measure we will focus on in the following is the CVaR. To define CVaR, let us start with the definition of the value-at-risk (VaR) [22, 42] at level $\beta \in (0, 1)$, VaR_β, which is equal to the β-quantile of the distribution of $f(\boldsymbol{x}, \Xi)$

$$\mathrm{VaR}_\beta[f(\boldsymbol{x}, \Xi)] = \min \{y \in \mathbb{R} \mid \mathbb{P}[f(\boldsymbol{x}, \Xi) \leq y] \geq \beta\}.$$

In case of our model output random variable $f(\boldsymbol{x}, \Xi)$, the $\mathrm{VaR}_\beta[f(\boldsymbol{x}, \Xi)] \in \mathbb{R}$ is the value such that $f(\boldsymbol{x}, \Xi)$ is below $\mathrm{VaR}_\beta[f(\boldsymbol{x}, \Xi)]$ with probability β which is visualized in Figure 2b.

The $\mathrm{CVaR}_\beta[f(\boldsymbol{x}, \Xi)]$ (also known as β-superquantile) is the expectation of $f(\boldsymbol{x}, \Xi)$ conditioned on $\mathrm{VaR}_\beta[f(\boldsymbol{x}, \Xi)]$

$$\mathrm{CVaR}_\beta[f(\boldsymbol{x}, \Xi)] = \mathbb{E}\left[f(\boldsymbol{x}, \Xi) \mid f(\boldsymbol{x}, \Xi) \leq \mathrm{VaR}_\beta[f(\boldsymbol{x}, \Xi)]\right] = \int_{\mathcal{Z}} \mathbb{I}_\beta(\boldsymbol{\xi}) f(\boldsymbol{x}, \boldsymbol{\xi}) p(\boldsymbol{\xi}) \mathrm{d}\boldsymbol{\xi}, \tag{2}$$

with the indicator function

$$\mathbb{I}_\beta(\boldsymbol{\xi}) = \begin{cases} 1, & f(\boldsymbol{x}, \boldsymbol{\xi}) \leq \mathrm{VaR}_\beta[f(\boldsymbol{x}, \boldsymbol{\xi})] \\ 0, & \text{else} \end{cases} \tag{3}$$

and density p of Ξ; see [20]. Note that $\mathrm{VaR}_\beta[f(\boldsymbol{x}, \Xi)]$ and $\mathrm{CVaR}_\beta[f(\boldsymbol{x}, \Xi)]$ are quantities that depend on the design point $\boldsymbol{x}$.

B. CVaR-based optimization

The risk-averse optimization formulation defines an objective function

$$J_\beta(\boldsymbol{x}) = \mathrm{CVaR}_\beta[f(\boldsymbol{x}, \Xi)] \tag{4}$$

3

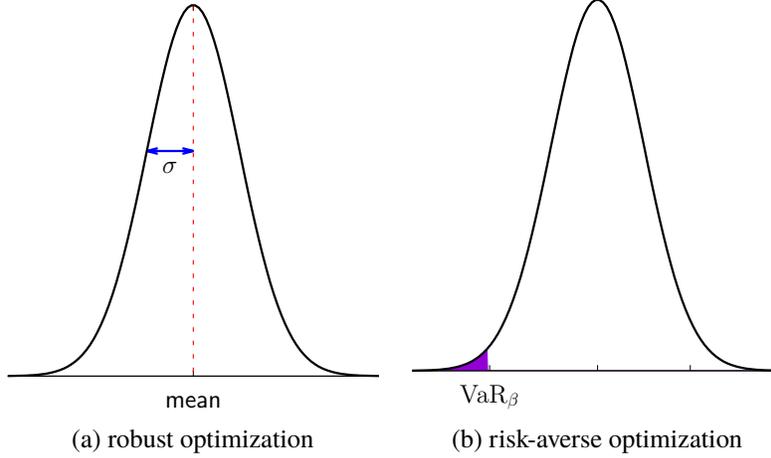(a) robust optimization   (b) risk-averse optimization

Fig. 2   Robust optimization penalizes deviations from the mean behavior, indicated by the standard deviation σ in plot (a). In contrast, risk-averse optimization emphasizes rare and low-probability events that are in the tails of the distribution, see plot (b). The VaR_β is the β-quantile of the distribution and CVaR_β is the expectation in the tail conditioned on the VaR_β, which is indicated by the shaded area in (b).

with the probability level $\beta \in (0, 1)$ to solve

$$\min_{x \in \mathcal{X}} J_\beta(x) . \tag{5}$$

Since, CVaR_β is the expectation of $f(x, \Xi)$ conditioned on the tail $f(x, \Xi) \leq \text{VaR}_\beta[f(x, \Xi)]$, minimizing the CVaR of $f(x, \Xi)$ minimizes the expectation of $f(x, \Xi)$ in the β-quantile. This means, even in limit states corresponding to $f(x, \Xi) \leq \text{VaR}_\beta[f(x, \Xi)]$, the expectation of $f(x, \Xi)$ is enforced to remain small. The readers are referred to Refs. [19–21, 23–25] for further discussion on the properties of CVaR and CVaR-based optimization.

C. Monte Carlo estimators of VaR and CVaR

Evaluating the objective J_β defined in (4) requires estimating $\text{CVaR}_\beta[f(x, \Xi)]$. The work [22] proposes a two-step Monte Carlo estimator: First, a Monte Carlo estimate $\widehat{\text{VaR}}_\beta[f(x, \Xi)]$ of $\text{VaR}_\beta[f(x, \Xi)]$ is obtained. Then, an estimate $\widehat{\text{CVaR}}_\beta[f(x, \Xi)]$ of $\text{CVaR}_\beta[f(x, \Xi)]$ is derived that builds on the estimate $\widehat{\text{VaR}}_\beta[f(x, \Xi)]$.

Consider the estimation of $\text{VaR}_\beta[f(x, \Xi)]$ first. Let $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_m$ be m realizations of Ξ and evaluate the high-fidelity model at the realizations, i.e.,

$$f(x, \boldsymbol{\xi}_1), \ldots, f(x, \boldsymbol{\xi}_m) . \tag{6}$$

Consider now the sorted outputs

$$f(x, \boldsymbol{\xi}_{i_1}) \leq f(x, \boldsymbol{\xi}_{i_2}) \leq \cdots \leq f(x, \boldsymbol{\xi}_{i_m}),$$

with indices $i_1, \ldots, i_m \in \{1, \ldots, m\}$. Note that such an ordering exists because the range of f is $\mathbb{R}$. Then, following [22], an estimate of $\text{VaR}_\beta[f(x, \Xi)]$ is $\widehat{\text{VaR}}_\beta[f(x, \Xi)] = f(x, \boldsymbol{\xi}_{i_r})$, where $r \in \{1, \ldots, m\}$ such that

$$\frac{r-1}{m} \leq \beta \leq \frac{r}{m} .$$

The estimator $\widehat{\text{VaR}}_\beta[f(x, \Xi)]$ is a biased estimator of $\text{VaR}_\beta[f(x, \Xi)]$. The bias converges to 0 for $m \to \infty$, see, e.g., [42–44].

To estimate CVaR_β, define the indicator function

$$\hat{\mathbb{1}}_\beta(\boldsymbol{\xi}) = \begin{cases} 1, & f(x, \boldsymbol{\xi}) \leq \widehat{\text{VaR}}_\beta[f(x, \Xi)], \\ 0, & \text{else}, \end{cases}$$

which is based on the estimate $\widehat{\text{VaR}}_\beta$, instead of VaR_β as in (3). Consider now again the realizations $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_m$ that

have been used in (6) to evaluate the high-fidelity model and estimate

$$\widehat{\text{CVaR}}_\beta[f(\boldsymbol{x}, \Xi)] = \frac{1}{m} \sum_{i=1}^m \hat{\mathbb{1}}_\beta(\boldsymbol{\xi}_i) f(\boldsymbol{x}, \boldsymbol{\xi}_i) p(\boldsymbol{\xi}_i).$$

The consistency and asymptotic normality of the estimators $\widehat{\text{VaR}}_\beta$ and $\widehat{\text{CVaR}}_\beta$ are studied in the literature [22, 42].

D. Computational challenges of risk-averse optimization with CVaR

Monte Carlo methods provide one versatile way of estimating the CVaR_β; however, if $\beta \ll 1$, then the $\widehat{\text{VaR}}_\beta$ estimator and the $\widehat{\text{CVaR}}_\beta$ estimator derived in Section III.C are inefficient. In case of $\widehat{\text{VaR}}_\beta$, a large number of realizations of Ξ, and consequentially a large number of model evaluations, are necessary to obtain a reasonable number of realizations in the β-quantile to obtain an acceptable estimate of VaR_β. Similarly, for $\beta \ll 1$, a large number of realizations is necessary for the indicator function $\hat{\mathbb{1}}_\beta$ in $\widehat{\text{CVaR}}_\beta$ to evaluate to 1 sufficiently many times so that an accurate estimate of CVaR_β is obtained.

IV. Risk-averse optimization with the multifidelity cross-entropy method

We develop a multifidelity approach for risk-averse optimization with CVaR to speed up the estimation of VaR and CVaR and so make large-scale risk-averse optimization tractable. We build on the MFCE method that was introduced in [38]. The MFCE method is based on importance sampling, which is a variance reduction approach for Monte Carlo estimators. In Section IV.A, importance-sampling estimators of VaR and CVaR are briefly reviewed. Section IV.B describes the multifidelity setting and MFCE to construct the importance sampling estimators of VaR and CVaR. Section IV.C gives details on the overall procedure for risk-averse optimization with the MFCE estimators of CVaR.

A. Variance reduction with importance sampling for estimating risk measures

Importance sampling is a variance reduction strategy for Monte Carlo estimation. Consider the model output random variable $f(\boldsymbol{x}, \Xi)$ with density p and define the support of a density p to be $\text{supp}(p) = \{\boldsymbol{\xi} \in \mathcal{Z} \mid p(\boldsymbol{\xi}) > 0\}$. Let $\tilde{p}$ be a biasing density with $\text{supp}(p) \subseteq \text{supp}(\tilde{p})$. An importance sampling estimator of $\widehat{\text{VaR}}_\beta$ is then $\widehat{\text{VaR}}_\beta^{(\text{IS})}[f(\boldsymbol{x}, \Xi)] = f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}_{i_r})$ with $r \in \{1, \ldots, m\}$ such that

$$\frac{1}{m} \sum_{j=1}^{r-1} \frac{p(\tilde{\boldsymbol{\xi}}_{i_j})}{\tilde{p}(\tilde{\boldsymbol{\xi}}_{i_j})} \leq \beta \leq \frac{1}{m} \sum_{j=1}^r \frac{p(\tilde{\boldsymbol{\xi}}_{i_j})}{\tilde{p}(\tilde{\boldsymbol{\xi}}_{i_j})},$$

where $\tilde{\boldsymbol{\xi}}_1, \ldots, \tilde{\boldsymbol{\xi}}_m$ are m realizations of the biasing random variable $\tilde{\Xi}$ with density $\tilde{p}$ and $f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}_{i_1}) \leq f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}_{i_2}) \leq \cdots \leq f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}_{i_m})$.

The importance sampling estimator $\widehat{\text{CVaR}}_\beta^{(\text{IS})}$ of CVaR_β is

$$\widehat{\text{CVaR}}_\beta^{(\text{IS})}[f(\boldsymbol{x}, \Xi)] = \frac{1}{m} \sum_{i=1}^m \hat{\mathbb{1}}_\beta(\tilde{\boldsymbol{\xi}}_i) f(\boldsymbol{x}, \tilde{\boldsymbol{\xi}}_i) \frac{p(\tilde{\boldsymbol{\xi}}_i)}{\tilde{p}(\tilde{\boldsymbol{\xi}}_i)}.$$

We refer to the survey in [22] for details and analyses.

A critical step in importance sampling is the construction of the biasing density such that the importance sampling estimators have a lower variance than the standard Monte Carlo estimators

$$\mathbb{V}\left[\widehat{\text{VaR}}_\beta^{(\text{IS})}\right] \ll \mathbb{V}\left[\widehat{\text{VaR}}_\beta\right]$$

and

$$\mathbb{V}\left[\widehat{\text{CVaR}}_\beta^{(\text{IS})}\right] \ll \mathbb{V}\left[\widehat{\text{CVaR}}_\beta\right].$$

B. Constructing biasing densities with the cross-entropy method and MFCE

The cross-entropy method is an iterative process to construct biasing densities [39–41]. Let $Q = \{\tilde{p}_{\boldsymbol{v}} \mid \boldsymbol{v} \in \mathcal{D}\}$ be a set of parametrized densities $\tilde{p}_{\boldsymbol{v}}$ with parameters $\boldsymbol{v} \in \mathcal{D}$ in the parameter domain $\mathcal{D}$. A common choice for Q is the set
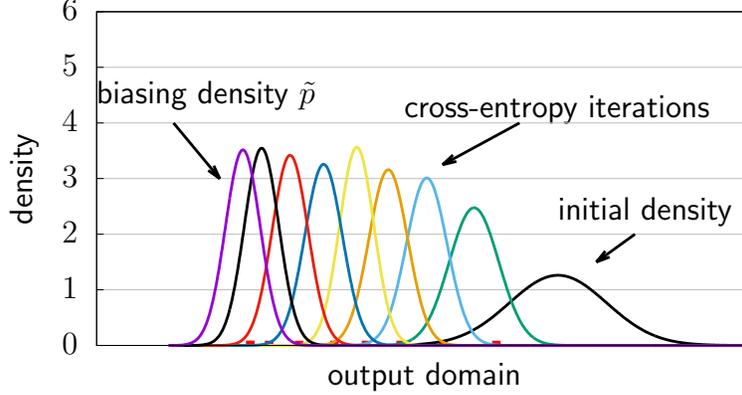
5

Fig. 3 The cross-entropy method iteratively refines an initial density until a biasing density $\tilde{q}$ is found with minimal Kullback-Leibler distance to the zero-variance density $\tilde{q}_0$.

of multivariate Gaussian random variables, where the parameters $\boldsymbol{v}$ are the mean and the covariance matrix. Let $\tilde{p}_0$ be the zero-variance biasing density such that

$$\mathbb{V}[\widehat{\text{CVaR}}_\beta^{(\text{IS})}] = 0,$$

which is known up to a normalizing constant [41]. The cross-entropy method derives an approximation $\tilde{p} \in Q$ in the set Q of the zero-variance biasing density $\tilde{p}_0$. The cross-entropy method solves for the biasing density $\tilde{p} \in Q$ that has minimal distance with respect to the Kullback-Leibler divergence from the zero-variance density $\tilde{p}_0$. The corresponding optimization problem is solved iteratively, as shown in Figure 3. The optimization problem in each step of the cross-entropy method can be solved analytically for a wide selection of sets Q of parametrized densities; see, e.g., [39–41].

Let $\tilde{f}^{(1)}, \ldots, \tilde{f}^{(k)} : \mathcal{X} \times \mathcal{Z} \to \mathbb{R}$ be $k \in \mathbb{N}$ low-fidelity models with evaluation costs $\tilde{w}_1, \ldots, \tilde{w}_k$ that are lower than the evaluation costs w of the high-fidelity model f. The MFCE method [38] leverages the low-fidelity models $\tilde{f}^{(1)}, \ldots, \tilde{f}^{(k)}$ to speed up the construction of biasing densities compared to the cross-entropy method that relies on the high-fidelity model f alone. The MFCE method is an extension of the cross-entropy method that solves for a biasing density $\tilde{p}^{(i)}$ with the low-fidelity models $\tilde{f}^{(i)}$ for $i = 1, \ldots, k$ first, and then uses these biasing densities to precondition the optimization problem in the final step when using the high-fidelity model to obtain $\tilde{p}$. Once the biasing density $\tilde{p}$ has been found, the high-fidelity model is used to derive an estimate of VaR and CVaR. We denote the MFCE estimators of VaR_β and CVaR_β as $\widehat{\text{VaR}}_\beta^{(\text{MF})}$ and $\widehat{\text{CVaR}}_\beta^{(\text{MF})}$, respectively. We refer to [38] for details on the MFCE method

C. Multifidelity risk-averse optimization with MFCE

MFCE provides an efficient way of estimating VaR and CVaR for a given β, if low-fidelity models are available. We combine the multifidelity estimation of VaR and CVaR into an optimization loop to speed up risk-averse optimization under uncertain. Thus, instead of the Monte Carlo estimator discussed in Section III.C, MFCE is run to estimate VaR and subsequently CVaR. This means that the objective function is

$$J_\beta^{(\text{MF})}(\boldsymbol{x}) = \widehat{\text{CVaR}}_\beta^{(\text{MF})}[f(\boldsymbol{x}, \Xi)],$$

with the MFCE estimator $\widehat{\text{CVaR}}_\beta^{(\text{MF})}$ of CVaR_β. We then plug $J_\beta^{(\text{MF})}$ in a standard optimization routine to compute $\hat{\boldsymbol{x}}^* \in \mathcal{X}$ that solves

$$\min_{\boldsymbol{x} \in \mathcal{X}} J_\beta^{(\text{MF})}(\boldsymbol{x}). \tag{7}$$

In the following, we will use derivative-free optimizers to solve (7), which typically are robust with respect to small magnitude noise. This robustness is important for us because of the variance of the estimators of CVaR_β when evaluating the objective function $J_\beta^{(\text{MF})}$. We use implicit filtering [45], for which a MATLAB implementation exists.

V. Numerical results

The numerical examples in this section demonstrate our multifidelity approach for risk-averse optimization. All runtime measurements are performed with MATLAB 2017b on compute nodes with Intel Xeon E5-1660v4 and 64GB RAM. The parameters of MFCE are chosen as described in [38, Section 4]. The implementation [45] `imfil` is used with its default configuration. Section V.A shows speedups of up to two orders of magnitude for a benchmark example from heat transfer. Section V.B presents a coupled aero-structural optimization problem. The objective is to minimize fuel burn under uncertain flight conditions. We report speedups of up to one order of magnitude compared to risk-averse optimization with standard Monte Carlo estimation that uses the high-fidelity model alone.

A. Elliptic problem

We first consider a one-dimensional heat problem with two design variables and two uncertain parameters.

1. Problem setup

Consider the PDE with random coefficients

$$-\nabla \cdot (a(\Xi, \boldsymbol{x}, v)\nabla u(\Xi, \boldsymbol{x}, v)) = 1\,, \qquad\qquad v \in \mathcal{V}\,, \tag{8}$$

$$u(\Xi, \boldsymbol{x}, 0) = 0\,, \tag{9}$$

$$\partial_n u(\Xi, \boldsymbol{x}, 1) = 0\,, \tag{10}$$

where $\mathcal{V} = (0, 1) \subset \mathbb{R}$ is the spatial domain and $\partial \mathcal{V} = \{0, 1\}$ is the boundary of $\mathcal{V}$. The solution function $u : \mathcal{Z} \times \mathcal{X} \times \mathcal{V} \to \mathbb{R}$ is defined on the Cartesian product of the set $\mathcal{Z}$, the design domain $\mathcal{X}$, and the closure $\bar{\mathcal{V}}$ of the spatial domain $\mathcal{V}$. Equation (9) imposes homogeneous Dirichlet boundary conditions on the left boundary $v = 0$ and equation (10) imposes homogeneous Neumann boundary conditions on the right boundary $v = 1$. The coefficient is

$$a(\Xi, \boldsymbol{x}, v) = \sum_{i=1}^{2} \exp(\xi_i) \exp\left(-0.5\frac{|v - x_i|}{0.0225}\right)\,,$$

where $\boldsymbol{x} = [x_1, x_2]^T \in \mathcal{X}$ is the design variable and $\Xi = [\xi_1, \xi_2]^T$ is a random vector. The design domain is

$$\mathcal{X} = [0.1, 0.4] \times [0.6, 0.9] \subset \mathbb{R}^2\,,$$

and the components of the random vector Ξ are independent and distributed normally with mean 1 and variance 0.1.

2. High- and low-fidelity models

Problem (8)–(10) is discretized with linear finite elements on an equidistant grid in $\bar{\mathcal{V}}$ with mesh width $h^{(\ell)} = 2^{-\ell}$ with $\ell \in \mathbb{N}$. The high-fidelity model f corresponds to level $\ell = 8$. The quantity of interest is the value of u at the right boundary, i.e.,

$$f(\boldsymbol{x}, \Xi) = u^{(8)}(\Xi, \boldsymbol{x}, 1)\,,$$

where $u^{(8)}$ is the finite-element approximation of u on a grid with mesh width $h^{(8)}$. Additionally, we build four low-fidelity models $\tilde{f}^{(1)}, \ldots, \tilde{f}^{(4)}$ corresponding to levels $\ell = 4, \ldots, 7$.

3. Multifidelity risk-averse optimization

Consider now the optimization problem

$$\min_{\boldsymbol{x} \in \mathcal{X}} \text{CVaR}_\beta[f(\boldsymbol{x}, \Xi)]\,,$$

where f is the high-fidelity model defined in Section V.A.2. We consider two estimators of CVaR_β. We first have the single-fidelity estimator $\widehat{\text{CVaR}}_\beta^{(\text{HF})}$ that uses the high-fidelity model f alone. The estimator $\widehat{\text{CVaR}}_\beta^{(\text{HF})}$ is obtained by running MFCE with the high-fidelity model alone, which means that the biasing density is constructed from the high-fidelity model using the cross-entropy method. Second, we have the multifidelity estimator $\widehat{\text{CVaR}}_\beta^{(\text{MF})}$ that uses the high-fidelity model f and the four low-fidelity models $\tilde{f}^{(1)}, \ldots, \tilde{f}^{(4)}$ for constructing the biasing density. Thus, MFCE is run with the models $f, \tilde{f}^{(1)}, \ldots, \tilde{f}^{(4)}$.

(a) objective with $m = 10^3$ samples

(b) objective with $m = 10^4$ samples

(c) objective with $m = 10^5$ samples
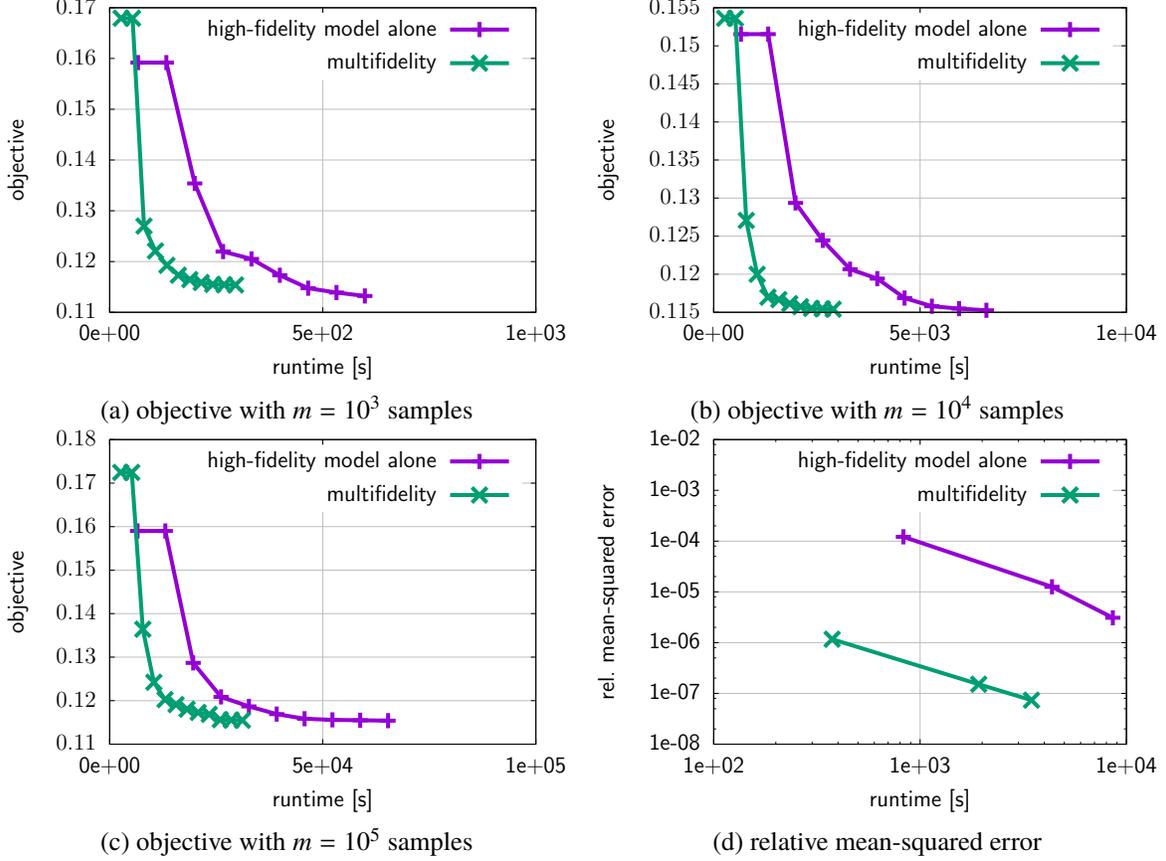
(d) relative mean-squared error

Fig. 4 Heat example: The probability level is set to $\beta = 10^{-7}$. Plots (a)-(c) show the value of the objective function for the single-fidelity and the multifidelity approach. The costs of the multifidelity approach seem lower than the costs of the single-fidelity approach that uses the high-fidelity model alone. The relative mean-squared error plotted in (d) shows that the multifidelity approach achieves a speedup of up to two orders of magnitude in this example.

Figure 4 reports result for risk-averse optimization with $\widehat{\mathrm{CVaR}}_\beta^{(\mathrm{HF})}$ and $\widehat{\mathrm{CVaR}}_\beta^{(\mathrm{MF})}$ for probability level $\beta = 10^{-7}$. Figure 4a shows the value of the objective function, i.e., either $\widehat{\mathrm{CVaR}}_\beta^{(\mathrm{HF})}$ or $\widehat{\mathrm{CVaR}}_\beta^{(\mathrm{MF})}$, during iterations of the optimization process. The number of samples is set to $m = 10^3$. The reported results are the average over 10 runs. The results indicate that the multifidelity approach locates the optimum with a lower runtime than the single-fidelity approach that uses the high-fidelity model alone. Similar observations are made in Figure 4b and Figure 4c where $m = 10^4$ and $m = 10^5$ samples, respectively, are used.

Next, we compare the value of objective function at the final iteration for the single- and the multifidelity approach. We optimize with the estimator $\widehat{\mathrm{CVaR}}_\beta^{(\mathrm{HF})}$ that uses the high-fidelity model and take $m = 5 \times 10^5$ samples. We obtain a design point $\hat{x}^*$ at which we evaluate the objective $\widehat{\mathrm{CVaR}}_\beta^{(\mathrm{HF})}[f(\hat{x}^*, \Xi)]$ using $m = 5 \times 10^5$ samples and store the value. We repeat this process ten times and take the average of the corresponding objective values as our reference value J^{Ref}. We then optimize with m samples with the high-fidelity model alone and with the multifidelity approach. Again, this process is repeated 10 times to obtain the objective values J_i^{MF} and J_i^{HF} for the runs $i = 1, \ldots, 10$. With

$$e^{(\mathrm{MF})} = \frac{1}{10} \sum_{i=1}^{10} \left(\frac{J^{(\mathrm{Ref})} - J_i^{(\mathrm{MF})}}{J^{(\mathrm{Ref})}} \right)^2, \ e^{(\mathrm{HF})} = \frac{1}{10} \sum_{i=1}^{10} \left(\frac{J^{(\mathrm{Ref})} - J_i^{(\mathrm{HF})}}{J^{(\mathrm{Ref})}} \right)^2, \tag{11}$$

we denote the relative mean-squared errors, which are reported in Figure 4d for $m \in \{10^3, 5 \times 10^3, 10^4\}$. The multifidelity approach achieves a speedup of about two orders of magnitude compared to the single-fidelity method that uses the high-fidelity model alone. Figure 5 shows the analogous results for the probability level $\beta = 10^{-8}$, which indicates that
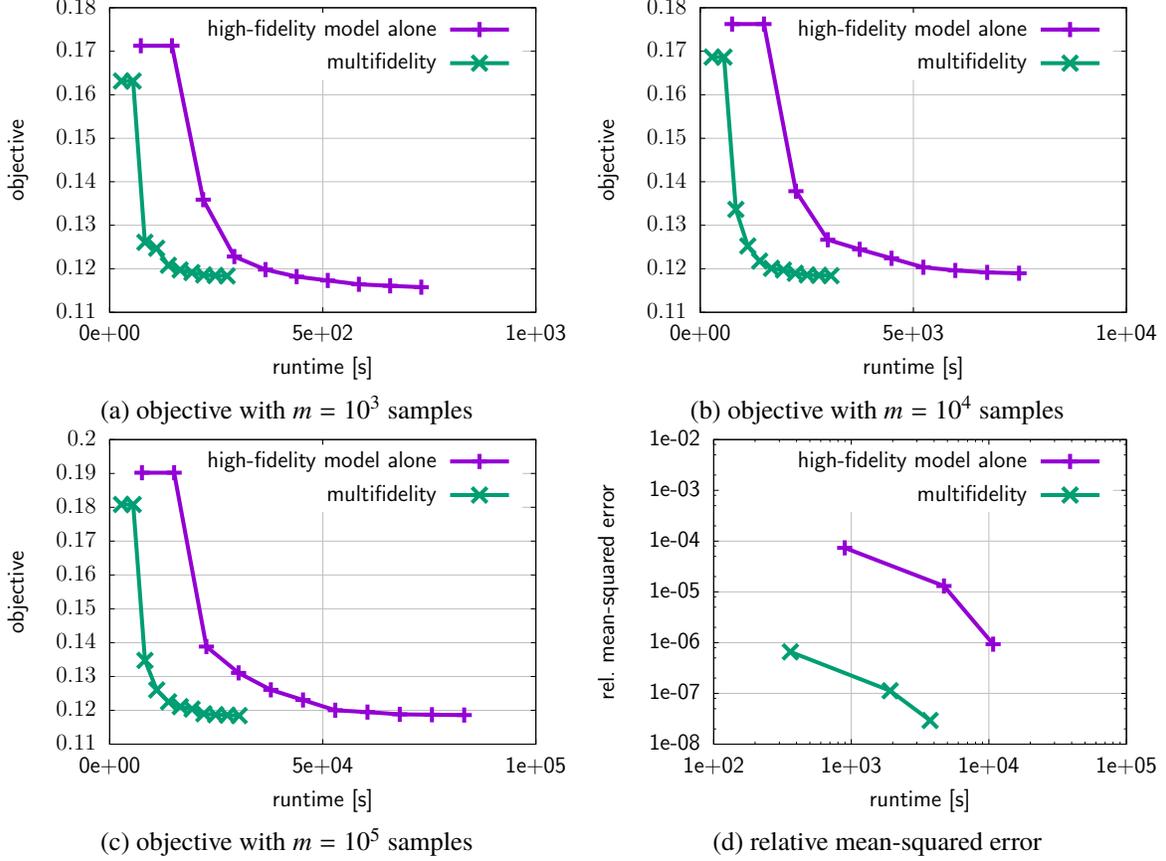
(a) objective with $m = 10^3$ samples

(b) objective with $m = 10^4$ samples

(c) objective with $m = 10^5$ samples

(d) relative mean-squared error

Fig. 5 Heat example: The probability level is set to $\beta = 10^{-8}$. The plots (a)-(c) show the convergence history of the multifidelity and the single-fidelity approach. A speedup of up to two orders of magnitude is obtained with the multifidelity approach with respect to the relative mean-squared error in the objective value compared to the single-fidelity approach, see plot (d).

our approach achieves a comparable performance as in Figure 4 for probability level $\beta = 10^{-7}$.

B. Risk-averse aero-structural design problem

In this section, we consider a coupled aero-structural design problem to determine the geometry of a wing with the objective to minimize fuel burn. We apply our multifidelity risk-averse optimization approach to minimize fuel burn at limit states.

1. Problem setup

Our high-fidelity model corresponds to a coupled aero-structural analysis* [46] that utilizes a vortex-lattice method and a 6 degrees of freedom 3-dimensional spatial beam model. The code simulates aerodynamic and structure analyses using lifting surfaces. The code is built in the framework of OpenMDAO [47], which is a platform for systems analysis and multidisciplinary optimization. An illustration of a wing is given in Figure 6. We consider wing meshes with 5 evenly spaced spanwise and 2 chordwise points, which is the standard configuration in the code. The design is parametrized by three design variables. The first two design variables control the thickness of the structural spar and the third design variable controls the twist variation along the span. The ranges of the design variables are given in Table 1. The random variable Ξ describes uncertain flight conditions as follows. The random variable $\Xi = [\xi_1, \xi_2, \xi_3]$ has three components, where ξ_1 describes the angle of attack, ξ_2 Mach number, and ξ_3 air density. The distribution of Ξ is a

---

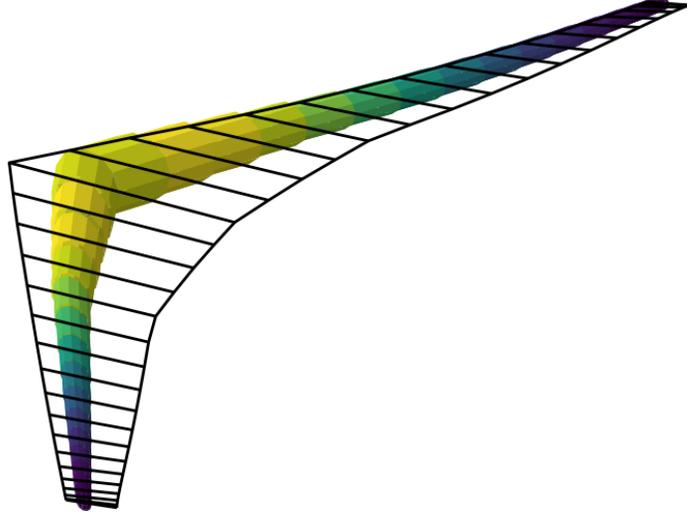*Code available at https://github.com/johnjasa/OpenAeroStruct/

Fig. 6 Illustration of a wing that can be analyzed with the OpenAeroStruct code. Figure from OpenAeroStruct documentation[*].

Table 1 Aero-structural design problem: The three-dimensional design variable $x = [x_1, x_2, x_3]^T$ controls the thickness and the twist of the wing configuration.

notation	design variable	lower bound	upper bound
x_1	thickness at location 1	0.01	0.25
x_2	thickness at location 2	0.01	0.25
x_3	twist	-2	10

mixture distribution of 2 normal distributions with mean $\boldsymbol{\mu}_1 = [5, 0.38, 0.80]^T$ and covariance

$$\Sigma_1 = \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.001 & 0 \\ 0 & 0 & 0.001 \end{bmatrix}$$

and mean $\boldsymbol{\mu}_2 = [5, 0.38, 0.84]^T$ and covariance

$$\Sigma_2 = \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.001 & 0 \\ 0 & 0 & 0.006 \end{bmatrix},$$

respectively. The rest of the parameters of OpenAeroStruct are set to their default values. The high-fidelity model $f : \mathcal{X} \times \mathcal{Z} \to \mathbb{R}$ returns the fuel burn for a design variable x and a realization $\xi \in \mathcal{Z}$ of Ξ.

2. Low-fidelity model

A spline interpolant of the function f with respect to the uncertain parameters serves as a low-fidelity model. Consider the high-fidelity model restricted to a design point $x \in \mathcal{X}$, i.e., $f_x(\xi) = f(x, \xi)$. Given $x \in \mathcal{X}$, we construct a spline interpolant of f_x. Let $n \in \mathbb{N}$ and consider an equidistant grid with n grid points in each direction in the domain

$$[4.5, 5.5] \times [0.2, 0.6] \times [0.3, 1.25] \subset \mathbb{R}^3.$$

To construct the spline interpolant $\tilde{f}_x$, the high-fidelity model f_x is evaluated at all grid points and a spline interpolant is constructed with the MATLAB function `griddedInterpolant` and the option `'spline'`. Thus, the costs of constructing the low-fidelity model are $n^3 \times w$, if one evaluation of the high-fidelity model has costs w. The low-fidelity model $\tilde{f}_x$ is about five orders of magnitude cheaper to evaluate than the high-fidelity model f.

(a) objective with $m = 10^3$ samples

(b) objective with $m = 10^4$ samples

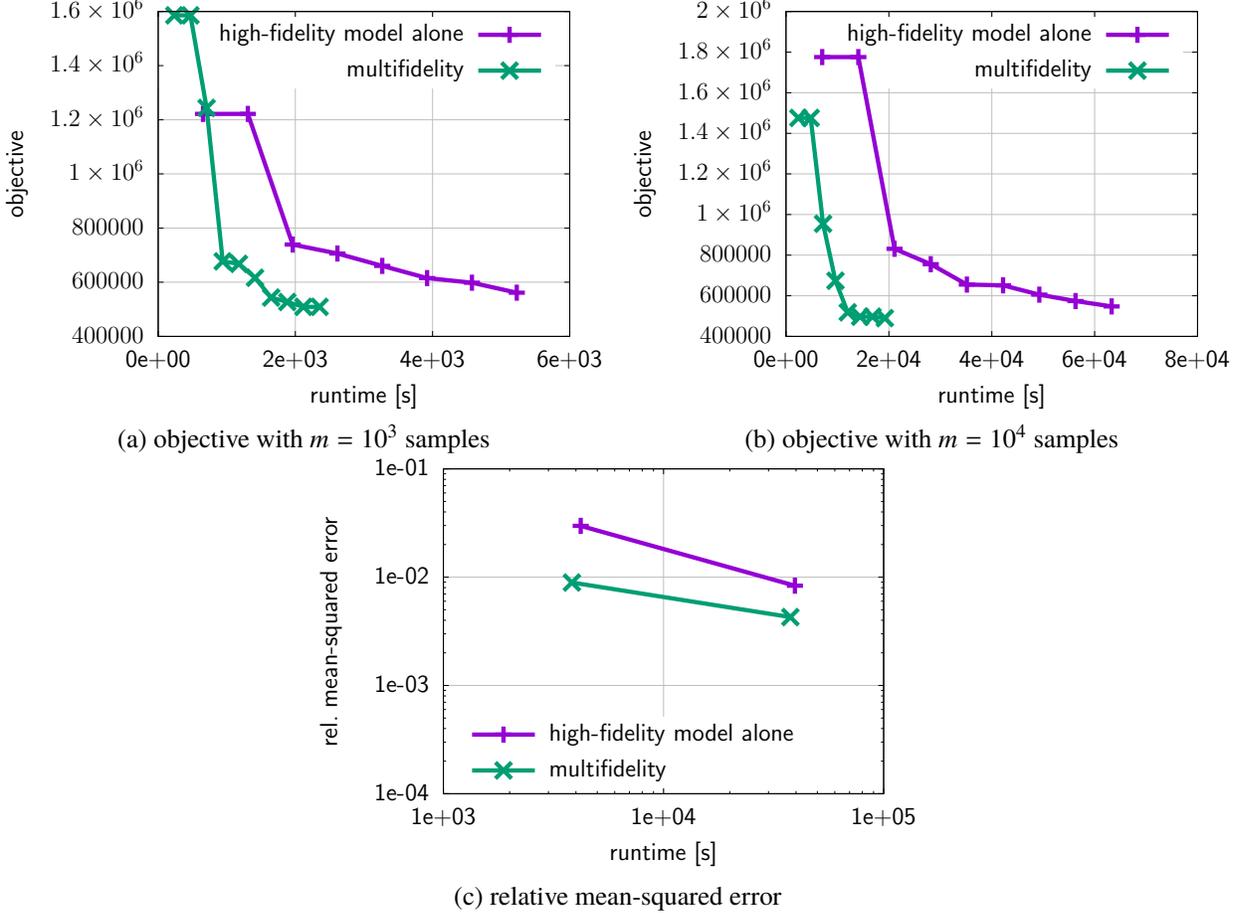(c) relative mean-squared error

Fig. 7   Coupled aero-structural optimization: The probability level is set to $\beta = 10^{-5}$. Plots (a) and (b) show the convergence history of the single-fidelity approach that uses the high-fidelity model alone and the multifidelity approach. Plot (c) shows the relative mean-squared error of the objective value achieved after a fixed computational budget. For the same budget, the multifidelity approach achieves an objective value with a lower relative mean-squared error than the single-fidelity approach that uses the high-fidelity model alone.

3. Multifidelity risk-averse optimization in aero-structural design problem

Our goal is to minimize the CVaR_β of the fuel burn for different probability levels β. We estimate CVaR_β with the single-fidelity estimator $\widehat{\text{CVaR}}_\beta^{(\text{HF})}$ that uses the high-fidelity model alone and the multifidelity estimator $\widehat{\text{CVaR}}_\beta^{(\text{MF})}$ as in Section V.A. In case of the multifidelity estimator, in each evaluation of the objective function, the low-fidelity model needs to be constructed for the current design point x before it can be evaluated. We use $n = 5$ grid points in each dimension to derive the spline interpolant. Thus, in each evaluation of the objective function, the low-fidelity model is first built from $n^5 = 125$ high-fidelity model evaluations.

Figure 7 reports results for probability level $\beta = 10^{-5}$. Figure 7a reports the average convergence history over five runs of the single-fidelity and the multifidelity approach for $m = 10^3$ samples. The results indicate that the multifidelity approach has a lower runtime than the single-fidelity approach. Similar results are obtained for $m = 10^4$ as reported in Figure 7b.

To compute the relative mean-squared error, we first derive a reference value using the high-fidelity model alone with $m = 10^3$ samples and by iterating the optimization until `imfil` stops with default configuration. The result of the optimization is a design point $\hat{x}^*$ at which we evaluate CVaR with $\widehat{\text{CVaR}}_\beta^{(\text{HF})}$ using the high-fidelity model alone and $m = 10^3$ samples. We repeat this process five times and the average of the CVaR estimates is our reference value $J^{(\text{Ref})}$. We then optimize with the multifidelity approach until `imfil` stops and obtain the objective value $J_i^{(\text{MF})}$ for $i = 1, \ldots, 5$ runs. We measure the average runtime of computing $J_i^{(\text{MF})}$. We take these average costs as the cost budget for the

11

(a) objective with $m = 10^3$ samples

(b) objective with $m = 10^4$ samples
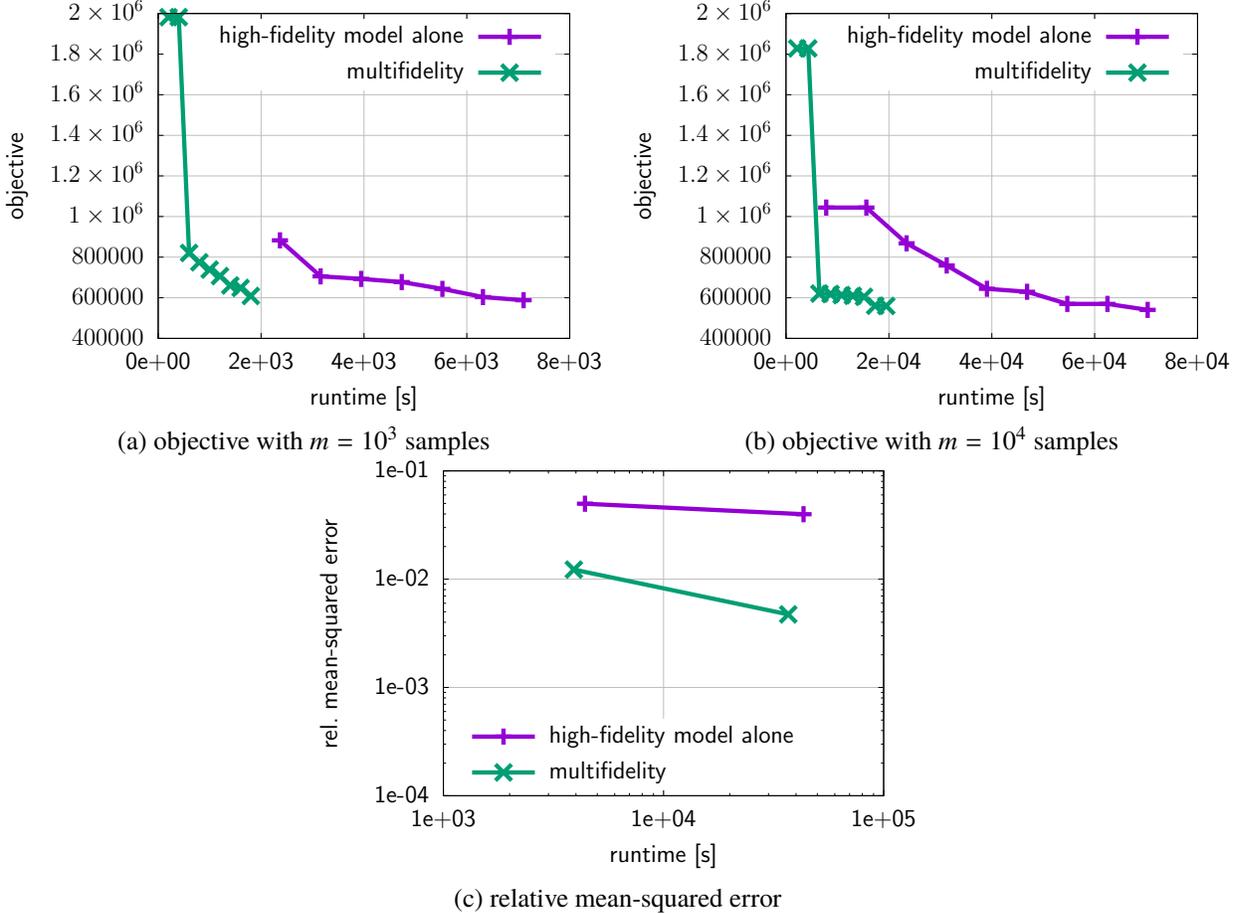
(c) relative mean-squared error

Fig. 8 Coupled aero-structural optimization: The plots in (a) and (b) show the objective value versus the costs for probability level $\beta = 10^{-6}$. Note that in plot (a) the value of the objective corresponding to the single-fidelity approach with the high-fidelity model alone is approximately 10^{12} in the first two iterations and therefore truncated from the plot. Plot (c) shows the runtime of the multifidelity risk-averse optimization versus the relative mean-squared error.

optimization with the high-fidelity model alone, i.e., we stop the optimization when the costs of evaluating the objective functions exceeds this cost budget and obtain objective values $J_i^{(\mathrm{HF})}$ for $i = 1, \ldots, 5$ runs. The relative mean-squared error is then computed as in (11) over the five runs. The results are reported in Figure 7c for $m = 10^3$ and $m = 10^4$ samples. The plot shows that our approach achieves speedups of a factor 2-3 compared to using the high-fidelity model alone. Figure 8 reports results for probability level $\beta = 10^{-6}$, where our multifidelity approach achieves speedups of almost one order of magnitude.

VI. Conclusions

This work presented a multifidelity approach for risk-averse optimization. The proposed approach leverages low-cost, low-fidelity models to obtain speedups compared to single-fidelity methods that use a high-fidelity model alone. At the same time, the proposed approach makes occasional recourse to the high-fidelity model to guarantee that the optimality conditions with respect to the high-fidelity model are satisfied. In the presented numerical examples, speedups of up to two orders of magnitude are obtained. We conclude with two remarks on the wide scope of the proposed approach. First, the numerical results demonstrate that the proposed approach is applicable with off-the-shelf low-fidelity models such as spline interpolants. Many standard scientific-computing software packages, e.g., MATLAB and scipy, implement routines for deriving spline interpolants, which reduces the implementation effort of the proposed approach compared to other approaches that require specific types of low-fidelity models. Second, the optimizer is used as a black box, which means that the proposed approach can be integrated into existing code frameworks in a non-intrusive way without

12

entailing in-depth changes to the optimization code.

Acknowledgments

References

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

[2] Forrester, A., Sóbester, A., and Keane, A., *Engineering design via surrogate modelling: a practical guide*, Wiley, 2008.

[3] Forrester, A. I. J., and Keane, A. J., "Recent advances in surrogate-based optimization," *Progress in Aerospace Sciences*, Vol. 45, No. 1–3, 2009, pp. 50–79.

[4] Krige, D. G., "A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand," *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, Vol. 52, No. 6, 1951, pp. 119–139.

[5] Cressie, N., "The origins of kriging," *Mathematical Geology*, Vol. 22, No. 3, 1990, pp. 239–252.

[6] Rasmussen, C., and Williams, C., *Gaussian Processes for Machine Learning*, MIT Press, 2006.

[7] Rumpfkeil, M. P., Hanazaki, K., and Beran, P. S., "Construction of Multi-Fidelity Locally Optimized Surrogate Models for Uncertainty Quantification," *19th AIAA Non-Deterministic Approaches Conference*, AIAA SciTech Forum, American Institute of Aeronautics and Astronautics, 2017, pp. 1–15.

[8] Rumpfkeil, M. P., and Beran, P., "Construction of Dynamic Multifidelity Locally Optimized Surrogate Models," *AIAA Journal*, Vol. 55, No. 9, 2017, pp. 3169–3179.

[9] Sirovich, L., "Turbulence and the dynamics of coherent structures," *Quarterly of Applied Mathematics*, Vol. 45, 1987, pp. 561–571.

[10] Rozza, G., Huynh, D., and Patera, A., "Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations," *Archives of Computational Methods in Engineering*, Vol. 15, No. 3, 2007, pp. 1–47.

[11] Gugercin, S., Antoulas, A., and Beattie, C., "$\mathcal{H}_2$ Model Reduction for Large-Scale Linear Dynamical Systems," *SIAM Journal on Matrix Analysis and Applications*, Vol. 30, No. 2, 2008, pp. 609–638.

[12] Anttonen, J., King, P., and Beran, P., "POD-Based reduced-order models with deforming grids," *Mathematical and Computer Modelling*, Vol. 38, No. 1, 2003, pp. 41 – 62.

[13] Lucia, D. J., Beran, P. S., and Silva, W. A., "Reduced-order modeling: new approaches for computational physics," *Progress in Aerospace Sciences*, Vol. 40, No. 1–2, 2004, pp. 51 – 117.

[14] Majda, A. J., and Gershgorin, B., "Quantifying uncertainty in climate change science through empirical information theory," *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 107, No. 34, 2010, pp. 14958–14963.

[15] Alexandrov, N. M., Jr, J. E. D., Lewis, R. M., and Torczon, V., "A trust-region framework for managing the use of approximation models in optimization," *Structural optimization*, Vol. 15, No. 1, 1998, pp. 16–23.

[16] Alexandrov, N. M., Lewis, R. M., Gumbert, C. R., Green, L. L., and Newman, P. A., "Approximation and Model Management in Aerodynamic Optimization with Variable-Fidelity Models," *Journal of Aircraft*, Vol. 38, No. 6, 2001, pp. 1093–1101.

[17] Jones, D. R., Schonlau, M., and Welch, W. J., "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, Vol. 13, No. 4, 1998, pp. 455–492.

[18] Jones, D. R., "A Taxonomy of Global Optimization Methods Based on Response Surfaces," *Journal of Global Optimization*, Vol. 21, No. 4, 2001, pp. 345–383.

[19] Rockafellar, R. T., and Uryasev, S., "Conditional value-at-risk for general loss distributions," *Journal of Banking & Finance*, Vol. 26, No. 7, 2002, pp. 1443 – 1471.

[20] Rockafellar, R. T., and Uryasev, S., "Optimization of conditional value-at-risk," *Journal of Risk*, Vol. 2, No. 3, 2000, pp. 21–41.

[21] Rockafellar, R. T., and Royset, J. O., *Superquantiles and Their Applications to Risk, Random Variables, and Regression*, Informs, 2013, Chap. Chapter 8, pp. 151–167.

[22] Hong, L. J., Hu, Z., and Liu, G., "Monte Carlo Methods for Value-at-Risk and Conditional Value-at-Risk: A Review," *ACM Trans. Model. Comput. Simul.*, Vol. 24, No. 4, 2014, pp. 22:1–22:37.

[23] Kouri, D. P., and Surowiec, T. M., "Risk-Averse PDE-Constrained Optimization Using the Conditional Value-At-Risk," *SIAM Journal on Optimization*, Vol. 26, No. 1, 2016, pp. 365–396.

[24] Yang, H., and Gunzburger, M., "Algorithms and analyses for stochastic optimization for turbofan noise reduction using parallel reduced-order modeling," *Computer Methods in Applied Mechanics and Engineering*, Vol. 319, 2017, pp. 217 – 239.

[25] Zou, Z., Kouri, D. P., and Aquino, W., "An Adaptive Sampling Approach for Solving PDEs with Uncertain Inputs and Evaluating Risk," *19th AIAA Non-Deterministic Approaches Conference*, AIAA SciTech Forum, American Institute of Aeronautics and Astronautics, 2017, pp. 1–13.

[26] Li, J., and Xiu, D., "Evaluation of failure probability via surrogate models," *Journal of Computational Physics*, Vol. 229, No. 23, 2010, pp. 8966–8980.

[27] Li, J., Li, J., and Xiu, D., "An efficient surrogate-based method for computing rare failure probability," *Journal of Computational Physics*, Vol. 230, No. 24, 2011, pp. 8683–8697.

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

[29] Fagerlund, F., Hellman, F., Målqvist, A., and Niemi, A., "Multilevel Monte Carlo methods for computing failure probability of porous media flow systems," *Advances in Water Resources*, Vol. 94, 2016, pp. 498 – 509.

[30] Elfverson, D., Hellman, F., and Målqvist, A., "A Multilevel Monte Carlo Method for Computing Failure Probabilities," *SIAM/ASA Journal on Uncertainty Quantification*, Vol. 4, No. 1, 2016, pp. 312–330.

[31] Ullmann, E., and Papaioannou, I., "Multilevel Estimation of Rare Events," *SIAM/ASA Journal on Uncertainty Quantification*, Vol. 3, No. 1, 2015, pp. 922–953.

[32] Heinkenschloss, M., Kramer, B., Takhtaganov, T., and Willcox, K., "Conditional-Value-at-Risk Estimation via Reduced-Order Models," *SIAM/ASA Journal on Uncertainty Quantification*, Vol. 6, No. 4, 2018, pp. 1395–1423.

[33] Peherstorfer, B., Cui, T., Marzouk, Y., and Willcox, K., "Multifidelity Importance Sampling," *Computer Methods in Applied Mechanics and Engineering*, Vol. 300, 2016, pp. 490–509.

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

[35] Chaudhuri, A., Kramer, B., and Willcox, K., "Information Reuse for Importance Sampling in Reliability-Based Design Optimization," *ACDL Technical Report TR-2019-01*, 2019.

[36] Chaudhuri, A., Marques, A. N., Lam, R., and Willcox, K., "Reusing Information for Multifidelity Active Learning in Reliability-Based Design Optimization," *AIAA Scitech 2019 Forum*, 2019, p. 1222.

[37] Ng, L. W. T., and Willcox, K., "Monte Carlo Information-Reuse Approach to Aircraft Conceptual Design Optimization Under Uncertainty," *Journal of Aircraft*, 2015, pp. 1–12.

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

[39] Rubinstein, R. Y., "Optimization of computer simulation models with rare events," *European Journal of Operational Research*, Vol. 99, No. 1, 1997, pp. 89 – 112.

[40] Rubinstein, R., "The Cross-Entropy Method for Combinatorial and Continuous Optimization," *Methodology And Computing In Applied Probability*, Vol. 1, No. 2, 1999, pp. 127–190.

[41] de Boer, P.-T., Kroese, D. P., Mannor, S., and Rubinstein, R. Y., "A Tutorial on the Cross-Entropy Method," *Annals of Operations Research*, Vol. 134, No. 1, 2005, pp. 19–67.

[42] Trindade, A. A., Uryasev, S., Shapiro, A., and Zrazhevsky, G., "Financial prediction with constrained tail risk," *Journal of Banking & Finance*, Vol. 31, No. 11, 2007, pp. 3524 – 3538. Risk Management and Quantitative Approaches in Finance.

[43] Hyndman, R. J., and Fan, Y., "Sample Quantiles in Statistical Packages," *The American Statistician*, Vol. 50, No. 4, 1996, pp. 361–365.

[44] Okolewski, A., and Rychlik, T., "Sharp distribution-free bounds on the bias in estimating quantiles via order statistics," *Statistics & Probability Letters*, Vol. 52, No. 2, 2001, pp. 207 – 213.

[45] Kelley, C., *Implicit Filtering*, Society for Industrial and Applied Mathematics, 2011.

[46] Jasa, J. P., Hwang, J. T., and Martins, J. R. R. A., "Open-source coupled aerostructural optimization using Python," *Structural and Multidisciplinary Optimization*, Vol. 57, 2018, pp. 1815–1827.

[47] Gray, J., Moore, K., and Naylor, B., "OpenMDAO: An Open Source Framework for Multidisciplinary Analysis and Optimization," *13th AIAA/ISSMO Multidisciplinary Analysis Optimization Conference*, Multidisciplinary Analysis Optimization Conferences, American Institute of Aeronautics and Astronautics, 2010, pp. 1–12.