Optimization and Engineering
https://doi.org/10.1007/s11081-026-10078-5

RESEARCH ARTICLE

Consequential improvement acquisition function for efficient multi-fidelity Bayesian optimization

Ibrahim Aydogdu[1,2] · Michaela Kempner[1] · Jesse M. Sestito[3] · Eva Zarkadoula[4] · Hongyuan Zha[5] · Yan Wang[1]

Received: 18 October 2024 / Revised: 19 December 2025 / Accepted: 21 January 2026
© The Author(s) 2026

Abstract

Surrogate-based Bayesian optimization has been widely applied in design optimization to increase sampling efficiency. However, the cost for each evaluation of the objective function can still be very high when physical experiments or large-scale simulations are involved. Multi-fidelity Bayesian optimization is the new approach to further improve the sampling efficiency by reducing the number of expensive samples at the highest fidelity level and supplementing them with less expensive ones at low-fidelity levels. In this paper, a new consequential improvement (CI) acquisition function is proposed to allow for the simultaneous selection of the solution and the fidelity level in problems with a known hierarchy of fidelity levels. The new CI acquisition function incorporates the consequential effectiveness of objective improvement with the considerations of cost, accuracy, and validity differences between high- and low-fidelity samples in engineering practice. The new method of multi-fidelity Bayesian optimization based on the CI is demonstrated with several analytical and simulation-based design examples. In the simulation-based design optimization example, the results show that the CI acquisition function has a decisive advantage in the sampling efficiency over the other methods of multi-fidelity Bayesian optimization with simultaneous selection. The results indicate that the proposed method is particularly advantageous in solving high-dimensional problems and when large cost ratios between high- and low-fidelity

✉ Yan Wang
yan-wang@gatech.edu

1 Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0405, USA

2 Department of Civil Engineering, Akdeniz University, Antalya 07058, Turkey

3 College of Engineering, Valparaiso University, Valparaiso, Indiana 46383-6493, USA

4 Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, USA

5 School of Data Science, Chinese University of Hong Kong, Shenzhen, China

Published online: 27 March 2026

evaluations exist and high-fidelity validation is mandatory. Furthermore, the method robustly avoids the prevalent issue of over sampling at low-fidelity levels.

Keywords Multi-fidelity Bayesian optimization · Active learning · Gaussian process

1 Introduction

In the iterative engineering design optimization process, the objective function needs to be evaluated at sample design solutions. When the number of design variables is large and the dimension of the feasible solution space is high, many samples are usually required before the optimal solution is obtained. Therefore, increasing the sampling efficiency is important for design optimization. Surrogate-based optimization methods such as Bayesian optimization (BO) (Shahriari et al. 2016) can increase the sampling efficiency. In surrogate-based optimization, a surrogate of the objective function is constructed and iteratively updated based on new samples. The surrogate predicts the objective value of unsampled design solutions. Specifically in BO, the surrogate, typically Gaussian process regression model, is used to identify the most promising sample for the next iteration based on an acquisition function. In each iteration, the solution that maximizes the acquisition function is selected as the next sample to be evaluated. The sampling efficiency of surrogate-based optimization is known to be higher than those of traditional population-based optimization methods such as genetic algorithms and particle swarm optimization. Hybrid surrogate-assisted evolutionary optimization methods have also been proposed to improve the sampling efficiency by introducing surrogates into population-based optimization (Liu et al. 2024).

Although the surrogate-based approach improves the sampling efficiency through the guidance of the surrogate, the cost for each evaluation of the objective function can still be very high, especially when physical experiments or large-scale simulations are involved. Recently, the multi-fidelity approach has attracted researchers' attention. In multi-fidelity Bayesian optimization (MFBO), samples at low-fidelity (LF) levels are used to supplement expensive samples at the highest-fidelity (HF) level. Among fidelity levels, evaluations of the objective differ in expense and accuracy. Evaluations at the LF levels are less expensive but also less accurate.

In engineering design, the most trustworthy technique to evaluate a design is usually the physical experiment or prototype which would be regarded as a HF sample and eventually needed for validation. That is, the unique role of the HF samples in engineering design is that they are also used for the purpose of validation ultimately. No further validation is needed if an HF sample is chosen as the final optimal design after the iterative search. In contrast, when an LF sample is selected as the optimum, it is implied that a corresponding HF level sample is still needed as the validation of the design, which is the hidden and deferred cost associated with the LF sample. When the multi-fidelity method is used to combine experimental and simulation data to perform design, even though simulation-based design can substantially reduce the cost, experiments as the HF-level evaluations are still required in engineering practice to validate the optimum and finalize the design. When the multi-fidelity method is used with simulations only, the prediction from the highest fidelity and most accurate

simulation model is typically used for validation because it is the most trustworthy one. With the consideration of accuracy and validity differences between HF and LF samples in engineering practice, a new LF sample does not effectively improve the objective value with high confidence if the best design estimated so far has been just sampled at the LF levels. Samples at both HF and LF levels provide valuable information about the portion of the design space where solutions could improve the objective. However, the LF evaluations themselves cannot be trusted as the most reliable values. Therefore, at least one sample at the HF level is necessary to quantify and materialize the actual improvement. The accurate quantification of actual improvement is only possible with trustworthy evaluations at the HF level. Therefore, in the multi-fidelity approach for design optimization, the unique merit of HF evaluations for the validated and materializable design improvement needs to be considered during the sample selection.

The primary contribution of this paper is a novel consequential improvement (CI) acquisition function for MFBO, which uniquely addresses the limitations of existing MFBO methods by simultaneously selecting the sample solution and fidelity level while accounting for the consequential effectiveness of objective improvement and hidden cost when LF samples are selected. Different from prior MFBO approaches, such as those with pre-scheduled fidelity selections (Khatouri et al. 2020; Perdikaris and Karniadakis 2016) or with disjointed selections of solutions and fidelity levels (Shu et al. 2021; Tran et al. 2020b), the CI acquisition function enables the simultaneous selection of the solution and fidelity level. At the same time, the new concept of hidden and deferred cost of validation associated with LF samples is introduced in this paper. The proposed CI acquisition function explicitly differentiates the roles of HF and LF samples in engineering design. In the existing MFBO approaches including the simultaneous selection methods (Poloczek et al. 2017; Takeno et al. 2020), the roles of LF and HF samples in engineering design are treated similarly and the critical role of validation for HF samples is omitted. In contrast, the proposed CI acquisition function incorporates the hidden cost of validation when LF samples are selected, since the actual improvement can only be proven through HF evaluations. As a result, this approach prevents over-sampling at LF levels and ensures that the final optimal design is validated at the HF level, aligning with the common engineering practice. By synergistically maximizing the CI acquisition function, our method achieves the overall high sampling efficiency, as demonstrated through analytical and simulation-based design examples.

The proposed CI acquisition function incorporates the consequential effectiveness of objective improvement, while the solution and the fidelity level of sampling are selected simultaneously. For a selection at the HF level, the value of the acquisition function is the quotient of the anticipated improvement in the objective and the evaluation cost over the current or immediate iteration. For a selection at an LF level, the value of the acquisition function is the quotient of the anticipated objective improvement and the overall evaluation cost over *both* the current and subsequent iterations. Therefore, the effects of LF samples and the associated 'hidden cost' for the required validation are further differentiated from those of HL samples. The 'hidden cost' of final validation plays a crucial role in engineering, which has not been considered in existing multi-fidelity optimization methods such as MES and KG. The proposed CI

acquisition function is particularly advantageous in solving problems of high dimensionality and when significant cost disparities exist between fidelity levels.

In the remainder of this paper, Sect. 2 provide the background of multi-fidelity surrogates and a review of acquisition functions used in MFBO. In Sect. 3, the new CI acquisition function is formulated and its application to MFBO is described. In Sect. 4, several analytical and simulation-based design examples are used to compare the sampling efficiency of MFBO with the new CI acquisition function to those of other MFBO methods for simultaneous selection. These examples include the multiscale design of 25-bar and 200-bar spatial truss structures.

2 Background

In this section, the most relevant work on MFBO is reviewed. In Sect. 2.1, a background of multi-fidelity surrogate modeling is introduced. In Sect. 2.2, the existing MFBO methods with scheduled selection of the fidelity level are reviewed. In Sect. 2.3, the existing MFBO methods with a disjointed selection of the solution and the fidelity level are summarized. In Sect. 2.4, the MFBO methods with simultaneous selection of the solution and fidelity level are reviewed.

2.1 Multi-fidelity surrogate of objective function

Multi-fidelity surrogate modeling (Gano et al. 2004, 2005) has been adopted as a computationally efficient method for design optimization. Gaussian process regression is the most used surrogate because a localized uncertainty level at each unsampled solution can be estimated in addition to the mean estimate of the objective. At each fidelity level, some additional hyperparameters, called cross-fidelity correlation coefficients, are introduced in the surrogate. The cross-fidelity correlation coefficients are optimized to fit the data obtained up to the respective fidelity level. For instance, the coefficients associated with the third fidelity level are only trained based on the data from the first, second, and third fidelity levels. All cross-fidelity correlation coefficients can be fit simultaneously (Han and Görtz 2012; Han et al. 2010; Zimmermann and Han 2010). The computational expense may be further reduced through an autoregressive model (Kennedy and O'Hagan 2000) with the coefficients fit in the ascending order of fidelity level. More recently, a hierarchical or recursive formulation of the multi-fidelity Gaussian process was proposed to eliminate the need for cross-fidelity correlation coefficients in the bi-fidelity (Han and Görtz 2012) and unlimited fidelity cases (Le Gratiet and Garnier 2014; Parussini et al. 2017). This further reduces the computational expense of training the surrogate. In this paper, this approach is used in constructing the Gaussian process regression models. Some non-hierarchical model fusion approaches (Chen et al. 2017; Foumani et al. 2023) were also proposed.

2.2 MFBO with scheduled selection of the fidelity level of sampling

In some methods of MFBO, the fidelity level of each sample is not selected based on the acquisition function. In the method developed by (Rezende et al. 2021), the fidelity level of sampling is predetermined, with a batch of samples taken at an LF level alongside one sample at the HF level. The selection of a batch of samples collectively spanning the fidelity levels has also been used in the work of Honarmandi et al. (Honarmandi et al. 2022). In other methods (Khatouri et al. 2020; Nachar et al. 2020; Perdikaris and Karniadakis 2016), the fidelity level is fixed between subsequent iterations. Alternatively, in the method developed by (Song et al. 2019), the sampling is divided into a phase of exploration prior to an iterative phase of exploitation, in which sampling is limited to the HF level. Similarly, in the method developed by (Frey et al. 2022), new samples are limited to the HF level. Because of the inflexibility in selecting the fidelity level of each sample, these methods provide only a limited advantage in sampling efficiency.

2.3 MFBO with disjointed selection of the fidelity level of sampling

The sampling efficiency of MFBO can be increased by selecting both the solution and the fidelity level of sampling in each iteration. In the methods of disjointed selection, the solution is selected before selecting the fidelity level at which the solution is sampled. In the methods of (Zhang et al. 2019) and (Korondi et al. 2021), the solution is selected to maximize the anticipated improvement of the objective at an LF level prior to the selection of the fidelity level. Due to the discrepancy between evaluations of the objective at different fidelity levels, these methods of selecting the solution could lead to redundant sampling.

The issue of redundant sampling is resolved in the following methods, where solutions are selected to maximize various metrics related to the anticipated improvement of the objective value at the HF level. For example, the fidelity level associated with the lowest cost can be selected from the subset of fidelity levels that meet a specified uncertainty threshold at the selected solution (Kandasamy et al. 2016, 2019, 2017; Sorourifar et al. 2023). In the methods developed by (Tran et al. 2020a, 2020b, 2020), the fidelity level is selected to maximize the ratio between the anticipated reduction in the uncertainty of the surrogate at the HF level and the cost of evaluation at the respective fidelity level. In the method developed by (Shu et al. 2021), the fidelity level is selected to maximize a metric that is specific to each fidelity level. In this method, the anticipated improvement quantifies the benefit of sampling at the HF level. The benefit of an LF sample is quantified as the difference between the anticipated improvement at the HF level after the same solution is evaluated at the current LF level and the anticipated improvement when directly evaluated at the HF level.

In the above methods, better combinations of solutions and fidelity levels may be missed during the disjointed selection because the metrics of selecting the fidelity level are specific to the selected solution. Therefore, direct extensions of these acquisition functions for simultaneous selection do not provide the best outcome.

2.4 MFBO with simultaneous selection of the fidelity level of sampling

The sampling efficiency of MFBO can be further increased by selecting both the solution and fidelity level simultaneously, rather than sequentially, in each iteration. The acquisition function in the methods of simultaneous selection is defined as a function of both sample location and fidelity level.

In some methods (Belakaria et al. 2020; Li et al. 2021, 2020; Takeno et al. 2020), the simultaneous selection is done by maximizing the ratio between a multi-fidelity formulation of max-value (output-space) entropy search (MES) and the evaluation cost. MES (Hoffman and Ghahramani 2015; Wang and Jegelka 2017) quantifies the anticipated reduction in the uncertainty associated with the optimum value of the surrogate at the HF level. In other methods (Imani and Ghoreishi 2022; Imani et al. 2019; Poloczek et al. 2017; Wu et al. 2020), the simultaneous selection is done by maximizing the ratio between a multi-fidelity formulation of the knowledge gradient (KG) and the evaluation cost. The KG (Frazier et al. 2009) quantifies the anticipated improvement in the optimum value of the surrogate at the HF level. However, since the surrogate optimum is used in place of the optimum objective value, the multi-fidelity extensions of the MES and KG do not account for the discrepancy between the surrogate and the original objective. Some additional evaluation cost is required to verify the objective at the surrogate optimum. Due to the mismatch in the ratio of benefit to cost, multi-fidelity extensions of the MES and KG could lead to over-sampling at LF levels. The mismatch is due to the overly similar treatment of the HF level versus the LF levels. An important feature of the multi-fidelity strategy for engineering design is that the objective improvement by an evaluation at a LF level cannot be validated. The actual improvement with validation is only possible at the HF level.

In this paper, a new acquisition function, called consequential improvement, is proposed to target the actual and validated improvement. When an LF sample is chosen, there is a hidden cost of validation, which has not been considered in the existing MFBO methods.

3 MFBO with proposed CI acquisition function

In this section, the new methodology of MFBO is described. For the sake of completeness and ease of reference, the formulation of the hierarchical multi-fidelity surrogate of the objective function without cross-fidelity correlation coefficients is described in Sect. 3.1. The proposed CI acquisition function is described in Sect. 3.2.

3.1 MFBO with simultaneous selection

We extend the bi-fidelity formulation of (Han and Görtz 2012) to allow for any number of fidelity levels. The extension is accomplished through a generalization of the relationship between consecutive levels of fidelity. The evaluation of the objective

function $f(\boldsymbol{x}, a)$ depends on both the solution $\boldsymbol{x}$ and the fidelity level a. Each solution within the solution domain Ω_1 is represented by a vector $\boldsymbol{x} \in \mathbb{R}^{1\times d}$ containing d components $x_1, \ldots, x_d$. The fidelity level a is an integer in the fidelity level domain $\Omega_2 = \{A - L, \ldots, A\}$, where the HF level is denoted as A and there are L different LF levels. In this paper, the minimization problem is used to illustrate our new method. The methodology can be applied to solve maximization problems. The ultimate goal of MFBO is to minimize the objective function at the HF level by solving

$$\min_{\boldsymbol{x}\in\Omega_1} f(\boldsymbol{x}, A) \tag{1}$$

MFBO is a double-loop searching process. The evaluation of the objective function in each iteration is done by maximizing acquisition function $r(\boldsymbol{x}, a)$ in the inner loop, as a mixed-integer optimization problem

$$\max_{\boldsymbol{x}, a} r(\boldsymbol{x}, a) \tag{2}$$

such that a particular sample solution $\boldsymbol{x}$ and a particular fidelity level a are selected for each outer-loop iteration. That is, the acquisition function helps select both the fidelity level and the sample location to obtain the evaluation $y_{\boldsymbol{x}}^{(a)} = f(\boldsymbol{x}, a)$. The cost of an evaluation at fidelity level a is $\Gamma(a)$. Due to the recursive nature of building the surrogate, some initial samples are required at each fidelity level. Afterward, a new solution and a fidelity level are selected to maximize the acquisition function. The covariance between the solutions that have been evaluated at fidelity level a are contained in the matrix

$$\mathbf{C}^{(a)} = \left[\boldsymbol{c}\left(\boldsymbol{x}_1^{(a)}, a\right), \ldots, \boldsymbol{c}\left(\boldsymbol{x}_{\mathcal{N}_a}^{(a)}, a\right)\right] \in \mathbb{R}^{\mathcal{N}_a\times\mathcal{N}_a} \tag{3}$$

where the j-th solution to be evaluated at fidelity level a is $\boldsymbol{x}_j^{(a)}$. The total number of evaluations at fidelity level a is $\mathcal{N}_a$. The column vector in the covariance matrix in Eq. (3) is calculated based on the function

$$\boldsymbol{c}(\boldsymbol{x}, a) = \left[c\left(\boldsymbol{x}_1^{(a)}, \boldsymbol{x}, a\right), \ldots, c\left(\boldsymbol{x}_{\mathcal{N}_a}^{(a)}, \boldsymbol{x}, a\right)\right]^T \in \mathbb{R}^{\mathcal{N}_a\times 1} \tag{4}$$

where $c\left(\boldsymbol{x}', \boldsymbol{x}, a\right)$ is a covariance kernel between solutions $\boldsymbol{x}'$ and $\boldsymbol{x}$ for fidelity level a. For example, Gaussian kernel (Gano et al. 2004, 2005) between two solutions, $\boldsymbol{x}'$ and $\boldsymbol{x}$, is.

$$c\left(\boldsymbol{x}', \boldsymbol{x}, a\right) = e^{-\theta_a \upsilon(\boldsymbol{x}', \boldsymbol{x})} \tag{5}$$

where $\upsilon\left(\boldsymbol{x}', \boldsymbol{x}\right)$ is an instance of the similarity metric evaluated between two solutions $\boldsymbol{x}'$ and $\boldsymbol{x}$. The similarity metric is defined to suit the problem of interest. At each

fidelity level a, the similarity metric is scaled by a hyperparameter, θ_a. The hyperparameters are tuned to maximize the fit of the surrogate. Each hyperparameter θ_a is tuned in the order of ascending fidelity level (Han and Görtz 2012; Le Gratiet and Garnier 2014) by maximizing the log-likelihood or equivalently solving

$$\min_{\theta_a} \mathcal{N}_a \log \mathbb{V}_a + \log\left|\mathbf{C}^{(a)}\right| \tag{6}$$

where

$$\mathbb{V}_a = \frac{1}{\mathcal{N}_a}\left(y^{(a)} - \gamma_a \boldsymbol{G}^{(a)}\right)^T \left(\mathbf{C}^{(a)} + \lambda \mathbf{I}\right)^{-1} \left(\boldsymbol{y}^{(a)} - \gamma_a \boldsymbol{G}^{(a)}\right) \tag{7}$$

is the Gaussian process variance for fidelity level a. $\mathbf{I}$ is the identity matrix with the same dimensions as $\mathbf{C}^{(a)}$. The noise term λ ensures the invertibility of the covariance matrices. Each evaluation at fidelity level a is listed in a column vector $\boldsymbol{y}^{(a)} \in \mathbb{R}^{\mathcal{N}_a \times 1}$.

$$\gamma_a = \frac{\left(\boldsymbol{G}^{(a)}\right)^T \left(\mathbf{C}^{(a)} + \lambda \mathbf{I}\right)^{-1} \boldsymbol{y}^{(a)}}{\left(\boldsymbol{G}^{(a)}\right)^T \left(\mathbf{C}^{(a)} + \lambda \mathbf{I}\right)^{-1} \boldsymbol{G}^{(a)}} \tag{8}$$

is the scaling factor that indicates the degree of correlation between evaluations at different fidelity levels. The scaling factor eliminates the need for fitting cross-fidelity correlation coefficients (Han and Görtz 2012). The recursive information fusion (Han and Görtz 2012; Kennedy and O'Hagan 2000; Le Gratiet and Garnier 2014) is based on the estimation of

$$\boldsymbol{G}^{(a)} = \left[\hat{f}\left(\boldsymbol{x}_1^{(a)}, a-1\right), \ldots, \hat{f}\left(\boldsymbol{x}_{\mathcal{N}_a}^{(a)}, a-1\right)\right]^T \in \mathbb{R}^{\mathcal{N}_a \times 1} \tag{9}$$

which contains the surrogate predictions at the immediately lower fidelity level, $a-1$, for the solutions sampled at the current fidelity level, a. The surrogate prediction of the objective for a solution $\boldsymbol{x}$ and fidelity level a is (Han and Görtz 2012; Le Gratiet and Garnier 2014)

$$\hat{f}(\boldsymbol{x}, a) = \gamma_a \hat{f}(\boldsymbol{x}, a-1) + \mathbf{c}^T(\boldsymbol{x}, a)\left(\mathbf{C}^{(a)} + \lambda I\right)^{-1}\left(\boldsymbol{y}^{(a)} - \gamma_a \boldsymbol{G}^{(a)}\right) \tag{10}$$

Note that the recursive information fusion is enforced by $\boldsymbol{G}^{(a)}$ defined in Eq. (9). A prediction for fidelity level a in Eq. (10) is recursively calculated based on all of the evaluations so far at or below that fidelity level a. As a stipulation, the identity scalar $\widehat{f}(\boldsymbol{x}, A-L-1) = 1$ allows for predictions at the lowest-fidelity level, $a = A - L$, based on the Gaussian process. The mean squared error of $\widehat{f}(\boldsymbol{x}, a)$ is (Han and Görtz 2012)

$$\mathbb{S}\left(\hat{f}(\boldsymbol{x}, a)\right) = \mathbb{V}_a^{1/2}\left[\frac{\left(\boldsymbol{c}^T(\boldsymbol{x}, a)\left(\mathbf{C}^{(a)} + \lambda \mathbf{I}\right)^{-1} \boldsymbol{G}^{(a)} - \hat{f}(\boldsymbol{x}, a-1)\right)^2}{\left(\boldsymbol{G}^{(a)}\right)^T \left(\mathbf{C}^{(a)} + \lambda \mathbf{I}\right)^{-1} \boldsymbol{G}^{(a)}}\right.$$

$$+1-\boldsymbol{c}^T(\boldsymbol{x},\,a)\left(\mathbf{C}^{(a)}+\lambda\mathbf{I}\right)^{-1}\boldsymbol{c}(\boldsymbol{x},\,a)\Bigg]^{1/2} \tag{11}$$

3.2 Proposed CI acquisition function

In this section, the new MFBO methodology with the proposed CI acquisition function is described. The existing MFBO methods only consider the immediate evaluation costs and overlook the additional HF evaluation required to validate an LF sample's improvement in engineering design. In contrast, the CI acquisition function incorporates the hidden and deferred cost of validation, where the function value for a LF selection as the quotient of the anticipated objective improvement and the overall evaluation cost, including both the current LF evaluation and the subsequent HF evaluation for validation. For a HF selection, the acquisition function is defined only with the current iteration's cost and improvement. This differentiation ensures that the cost–benefit analysis reflects the practical need for HF validation, which prevents over-sampling at LF levels and prioritizes validated improvements. This approach is significant because it aligns with engineering requirements for trustworthy and validated designs. At the same time, the CI acquisition function enhances sampling efficiency, particularly in high-dimensional or high-cost scenarios, as shown in our analytical and simulation-based examples in Sect. 4.

3.2.1 Improvement in the current iteration of the HF case

Let $\mathfrak{D}=\left\{\boldsymbol{y}^{(a)}\middle|a=A-L,\ldots,A\right\}$ denote the set of all evaluations. With the additional solution $\boldsymbol{x}^*$ at fidelity level a^* in the current iteration, the dataset is updated to $\mathfrak{D}^*=\mathfrak{D}\cup\left\{y_{\boldsymbol{x}^*}^{(a^*)}\right\}$. The improvement I of the optimal objective value with the latest sample at the HF level and the updated dataset is

$$I\left(\boldsymbol{x}^*,\,a^*|\mathfrak{D}^*\right)=\delta_{a^*,A}\max\left(0,\,y_{min}-y_{\boldsymbol{x}^*}^{(a^*)}\right) \tag{12}$$

for the minimization problem. Here, the minimization problem is used to illustrate, and y_{min} is the minimum (best) objective that has been evaluated at the HF level A. $\delta_{a^*,A}=1$ if $a^*=A$ and otherwise $\delta_{a^*,A}=0$. That is, the actual improvement only occurs when sampling at the highest fidelity level. For the maximization problem, the improvement is similarly defined as

$$I\left(\boldsymbol{x}^*,\,a^*|\mathfrak{D}^*\right)=\delta_{a^*,A}\max\left(0,\,y_{x^*}^{(a^*)}-y_{\max}\right) \tag{13}$$

where y_{max} is the maximum objective that has been evaluated at the HF level. Note that $I(\boldsymbol{x}^*,\,a^*|\mathfrak{D}^*)$ is also conditional on the evaluation $y_{\boldsymbol{x}^*}^{(a^*)}$ which is probabilistic and not

precisely known. The conditional expectation of the improvement over the unknown evaluation $y_{\boldsymbol{x}^*}^{(a^*)}$ is $\mathbb{E}_*(I(\boldsymbol{x}^*, a^*|\mathfrak{D}^*))$. To simplify the notation, we use $\mathbb{E}_*(\bullet)$ to denote the conditional expected improvement $\mathbb{E}_*(\bullet|\mathfrak{D}^*)$ over the latest sample evaluation $y_{\boldsymbol{x}^*}^{(a^*)}$. In this paper, $\mathbb{E}_*(I(\boldsymbol{x}^*, a^*))$ is derived as a multi-fidelity generalization of the expected improvement. The conditional expectation of the improvement for the optimal objective value that has been sampled at the HF level is calculated as

$$\mathbb{E}_*\big(I\big(x^*, a^*\big)\big) = \int_{-\infty}^{\infty} \Big(I\Big(x^*, a^*|\mathfrak{D}, y_{x^*}^{(a^*)} = z\Big)\mathbb{P}\Big(y_{x^*}^{(a^*)} = z\Big)\Big)dz \tag{14}$$

where z is a possible value of the objective function evaluation at solution $\boldsymbol{x}^*$ and fidelity level a^*, and the probability of obtaining such value $\mathbb{P}\Big(y_{\boldsymbol{x}^*}^{(a^*)} = z\Big)$ is calculated based on the Gaussian process surrogate's predictive distribution. Since the improvement I is zero unless $a^* = A$, $\mathbb{E}_*(I(\boldsymbol{x}^*, a^*))$ is only non-zero when $a^* = A$. This property reflects the nature of engineering design, in that a LF selection cannot directly yield the actual and materializable improvement. Only a validated HF evaluation can do so. Nevertheless, LF samples can still be used to inform the decision of sampling in the subsequent iteration. Therefore, the benefit of an LF selection may be more accurately quantified from the subsequent iteration.

For a selection at the HF level, the analytical form of conditional expectation is approximated based on the Gaussian distribution. For instance, the expectation of a function g conditional on u is calculated as

$$\mathbb{E}(g|u) = \int_{-\infty}^{\infty} ((g|u = z)\mathbb{P}(u = z))dz \approx \mathbb{M}(g|u)\Phi\left(\frac{\mathbb{M}(g|u)}{\mathbb{S}(g|u)}\right) + \mathbb{S}(g|u)\varphi\left(\frac{\mathbb{M}(g|u)}{\mathbb{S}(g|u)}\right) \tag{15}$$

where $\Phi(\bullet)$ is the cumulative distribution function of the standard normal distribution, and $\varphi(\bullet)$ is the probability density function. $\mathbb{M}(\bullet)$ is the mean function associated with the Gaussian distribution. $\mathbb{S}(\bullet)$ indicates the square root of the mean squared error. The mean of the improvement I for a minimization problem is approximated as

$$\mathbb{M}\big(I\big(\boldsymbol{x}^*, A|\mathfrak{D}^*\big)\big) \approx y_{\min} - \hat{f}\big(\boldsymbol{x}^*, A\big) \tag{16}$$

The square root of the mean squared error of the improvement I is approximated as

$$\mathbb{S}\big(I\big(\boldsymbol{x}^*, A|\mathfrak{D}^*\big)\big) \approx \mathbb{S}\Big(\hat{f}\big(\boldsymbol{x}^*, a^*\big)\Big) \tag{17}$$

and calculated based on Eq. (11).

Given the latest dataset $\mathfrak{D}^*$, the best objective value that has been evaluated at the HF level is updated as

$$y_{min}^* = \big(1 - \delta_{a^*, A}\big)y_{min} + \delta_{a^*, A}\min\Big(y_{min}, y_{\boldsymbol{x}^*}^{(a^*)}\Big) \tag{18}$$

for a minimization problem. The update for a maximization problem will be similar.

3.2.2 Consequential improvement in the subsequent iteration for LF case

When the next sample $\boldsymbol{x}^*$ is not at the HF level, there is no actual improvement. The future improvement of the subsequent sample needs to be considered. Let $X^\dagger$ denote the subsequent sample at the fidelity level $a^\dagger$ in the subsequent iteration. The dataset is further updated to $\mathfrak{D}^\dagger = \mathfrak{D}^* \cup \left\{ y_{x^\dagger}^{(a^\dagger)} \right\}$. The conditional expectation of the subsequent improvement $I\left(\boldsymbol{x}^\dagger, a^\dagger | \mathfrak{D}^\dagger\right)$ in the subsequent iteration over the unknown evaluation $y_{x^\dagger}^{(a^\dagger)}$ is $\mathbb{E}_\dagger\left(I\left(\boldsymbol{x}^\dagger, a^\dagger | \mathfrak{D}^\dagger\right)\right)$, which can be calculated similarly to Eq. (14). To simplify the notation, we use $\mathbb{E}_\dagger(\cdot)$ to denote $\mathbb{E}_\dagger\left(\cdot | \mathfrak{D}^\dagger\right)$ We also use $\mathbb{E}_\dagger(I)$ to denote the $\mathbb{E}_\dagger\left(I\left(\boldsymbol{x}^\dagger, a^\dagger | \mathfrak{D}^\dagger\right)\right)$.

The input–output relationship between the inputs, $\boldsymbol{x}^\dagger$ and, $a^\dagger$ and the response value $\mathbb{E}_\dagger(I)$ is unknown because the subsequent improvement $I\left(\boldsymbol{x}^\dagger, a^\dagger | \mathfrak{D}^\dagger\right)$ is conditional on the unknown evaluation $y_{x*}^{(a*)}$ during the current iteration. The value of the expectation function $\mathbb{E}_\dagger(I)$ is approximated as its expected value, $\mathbb{E}_*\left(\mathbb{E}_\dagger(I)\right)$, further conditional on the unknown evaluation $y_{x^*}^{(a^*)}$.

The exact calculation of $\mathbb{E}_*\left(\mathbb{E}_\dagger(I)\right)$ would require a continuous distribution of possible values for $\mathbb{E}_\dagger(I)$ based on the potential evaluations in place of the unknown evaluation $y_{x^*}^{(a^*)}$. To simplify, $\mathbb{E}_*\left(\mathbb{E}_\dagger(I)\right)$ can be numerically approximated based on a finite number, $2m+1$, of potential evaluations. The t-th evaluation is calculated as

$$y_{t,\boldsymbol{x}^*}^{(a^*)} = \hat{f}\left(\boldsymbol{x}^*, a^*\right) + \alpha_t \mathbb{S}\left(\hat{f}\left(\boldsymbol{x}^*, a^*\right)\right) \tag{19}$$

where $\alpha_t \in [-\alpha_m, -\alpha_{m-1}, \ldots, -\alpha_1, 0, \alpha_1, \ldots, \alpha_{m-1}, \alpha_m]$ is a scalar value that equally samples the range of evaluations.

For each potential evaluation, there is a possible value $\mathbb{E}_\dagger\left(I\left(\boldsymbol{x}^\dagger, a^\dagger | \mathfrak{D}_t^\dagger\right)\right)$ which depends on the t-th possible dataset $\mathfrak{D}_t^\dagger = \mathfrak{D} \cup \left\{ y_{t,\boldsymbol{x}^*}^{(a^*)} \cup y_{\boldsymbol{x}^\dagger}^{(A)} \right\}$. To simplify the notation, we use an additional subscript (t) to denote the situation that a conditional expectation is dependent on the t-th potential dataset instead of an actual set. For example, we use $\mathbb{E}_{\dagger,t}(I)$ to represent $\mathbb{E}_{\dagger,t}\left(I\left(\boldsymbol{x}^\dagger, a^\dagger | \mathfrak{D}_t^\dagger\right)\right)$. Collectively, the potential evaluations are used to approximate the mean of $\mathbb{E}_\dagger(I)$ as

$$\mathbb{M}\left(\mathbb{E}_\dagger(I)\right) \approx \frac{1}{2m+1} \sum_{t=1}^{2m+1} \mathbb{E}_{\dagger,t}(I) \tag{20}$$

and the square root of the mean squared error as

$$\mathbb{S}\big(\mathbb{E}_{\dagger}(I)\big) \approx \left[\frac{1}{2m+1}\sum_{t=1}^{2m+1}\big(\mathbb{E}_{\dagger,t}(I) - \mathbb{M}\big(\mathbb{E}_{\dagger}(I)\big)\big)^2\right]^{1/2} \tag{21}$$

Then, $\mathbb{E}_{*}\big(\mathbb{E}_{\dagger}(I)\big)$ is approximated based on a Gaussian distribution as

$$\mathbb{E}_{*}\big(\mathbb{E}_{\dagger}(I)\big) \approx \mathbb{M}\big(\mathbb{E}_{\dagger}(I)\big)\Phi\left(\frac{\mathbb{M}\big(\mathbb{E}_{\dagger}(I)\big)}{\mathbb{S}\big(\mathbb{E}_{\dagger}(I)\big)}\right) + \mathbb{S}\big(\mathbb{E}_{\dagger}(I)\big)\varphi\left(\frac{\mathbb{M}\big(\mathbb{E}_{\dagger}(I)\big)}{\mathbb{S}\big(\mathbb{E}_{\dagger}(I)\big)}\right) \tag{22}$$

3.2.3 Other options to quantify consequential improvement

In the new MFBO approach, the acquisition function $r(\mathbf{x}^*, a^*)$ as a function of the next sample $\mathbf{x}^*$ and its fidelity level a^* that quantifies the expected consequential improvement per unit cost. Different choices of $r(\mathbf{x}^*, a^*)$ are possible. Here, two options are first described before our proposed one is given in Sect. 3.2.4.

The first option of the acquisition function is directly defined as the quotient of the expected current and consequential improvements and cost over both the current and subsequent iterations, as.

$$r\big(\boldsymbol{x}^*, a^*\big) = \max_{\boldsymbol{x}^{\dagger}\in\Omega_1, a^{\dagger}\in\Omega_2}\left(\frac{\mathbb{E}_{*}(I(\boldsymbol{x}^{*}, a^*)) + \mathbb{E}_{*}\big(\mathbb{E}_{\dagger}\big(I\big(\boldsymbol{x}^{\dagger}, a^{\dagger}\big)\big)\big)}{\Gamma(a^*) + \Gamma\big(a^{\dagger}\big)}\right) \tag{23}$$

Where $\mathbb{E}_{*}(\cdot)$ and $\mathbb{E}_{\dagger}(\cdot)$ are calculated based on Eq. (14). Particularly, $\mathbb{E}_{*}(I(\boldsymbol{x}^{*}, a^*))$ represents the expected improvement directly from the next sample $\mathbf{x}^*$ conditional on the current evaluation $y_{x^*}^{(a^*)}$, whereas $\mathbb{E}_{*}\big(\mathbb{E}_{\dagger}\big(I\big(\boldsymbol{x}^{\dagger}, a^{\dagger}\big)\big)\big)$ indicates the expected consequent improvement in the subsequent iteration, conditional on both the current evaluation $y_{x*}^{(a*)}$ and the subsequent evaluation $y_{x^{\dagger}}^{(a^{\dagger})}$, where $\mathbf{x}^{\dagger}$ and $a^{\dagger}$ are the solution and fidelity level chosen in the subsequent iteration. The next sample in the outer-loop iteration is selected by $r(x^*, a^*)$ maximizing within domain $\Omega_1 \times \Omega_2$ in the inner loop.

The acquisition function in Eq. (23), however, is not sensitive to the values of sample solution $\mathbf{x}^*$ and fidelity level a^*. For example, in the case that a subsequent selection is at the LF level $a^{\dagger}$ the quotient is zero for all possible selections of a solution $\mathbf{x}^*$ and LF level a^* This incorrectly suggests that there is no benefit of sampling.

The second option that can avoid the above issue is to apply the maximization directly to the conditional expectation of the improvement before the quotient is applied, as in

$$r\big(\mathbf{x}^*, a^*\big) = \frac{\mathbb{E}_{*}(I(\mathbf{x}^*, a^*)) + \mathbb{E}_{*}\left(\max\limits_{\mathbf{x}^{\dagger}, a^{\dagger}}\big(\mathbb{E}_{\dagger}\big(I\big(\mathbf{x}^{\dagger}, a^{\dagger}\big)\big)\big)\right)}{\Gamma(a^*) + \Gamma(A)} \tag{24}$$

This is done based on the presumption that the subsequent selection will be done to maximize the improvement.

For a subsequent selection of solution $x^{\dagger}$ and fidelity level $a^{\dagger}$ $\max\limits_{\boldsymbol{x}^{\dagger}, a^{\dagger}}\left(\mathbb{E}_{\dagger}\left(I\left(\boldsymbol{x}^{\dagger},\, a^{\dagger}\right)\right)\right)$ is the maximum value that could be obtained for $\mathbb{E}_{\dagger}\left(I\left(\boldsymbol{x}^{\dagger},\, a^{\dagger}\right)\right)$. The improvement is non-negative and may only be non-zero in the subsequent iteration when $a^{\dagger} = A$. As a result, the maximum value calculation in Eq. (24) may be simplified as

$$\max_{\boldsymbol{x}^{\dagger}, a^{\dagger}}\left(\mathbb{E}_{\dagger}\left(I\left(\boldsymbol{x}^{\dagger},\, a^{\dagger}\right)\right)\right) = \max_{\boldsymbol{x}^{\dagger}}\left(\mathbb{E}_{\dagger}\left(I\left(\boldsymbol{x}^{\dagger},\, A\right)\right)\right) \tag{25}$$

To further simplify notation, the arguments of the maximization function and the encompassed improvement function in the subsequent iteration are implied by their arrangement such that we use $\max\left(\mathbb{E}_{\dagger}(I)\right)$ to denote $\max\limits_{\boldsymbol{x}^{\dagger}}\left(\mathbb{E}_{\dagger}\left(I\left(\boldsymbol{x}^{\dagger},\, A\right)\right)\right)$

An interpretation is that the maximum value $\max\left(\mathbb{E}_{\dagger}(I)\right)$ is the greatest expected improvement that is available among possible selections in the subsequent iteration. Because the expectation $\mathbb{E}_{\dagger}(I)$ is unknown, its maximum value $\max\left(\mathbb{E}_{\dagger}(I)\right)$ is also unknown. The value of the maximization function $\max\left(\mathbb{E}_{\dagger}(I)\right)$ is approximated as its expected value, $\mathbb{E}_{*}\left(\max\left(\mathbb{E}_{\dagger}(I)\right)\right)$, conditional on the unknown evaluation $y_{x*}^{(a*)}$ The value of $\mathbb{E}_{*}\left(\max\left(\mathbb{E}_{\dagger}(I)\right)\right)$ can be approximated similarly to the ones proposed for $\mathbb{E}_{*}\left(\mathbb{E}_{\dagger}(I)\right)$ in Eqn. (20)-(22). That is, the consequential improvement in the subsequent iteration, $\mathbb{E}_{*}\left(\max\left(\mathbb{E}_{\dagger}(I)\right)\right)$, is approximated based on a Gaussian distribution as

$$\begin{aligned}\mathbb{E}_{*}\left(\max\left(\mathbb{E}_{\dagger}\left(I\right)\right)\right) \approx\, & \mathbb{M}\left(\max\left(\mathbb{E}_{\dagger}\left(I\right)\right)\right)\Phi\left(\frac{\mathbb{M}\left(\max\left(\mathbb{E}_{\dagger}\left(I\right)\right)\right)}{\mathbb{S}\left(\max\left(\mathbb{E}_{\dagger}\left(I\right)\right)\right)}\right) \\ & + \mathbb{S}\left(\max\left(\mathbb{E}_{\dagger}\left(I\right)\right)\right)\varphi\left(\frac{\mathbb{M}\left(\max\left(\mathbb{E}_{\dagger}\left(I\right)\right)\right)}{\mathbb{S}\left(\max\left(\mathbb{E}_{\dagger}\left(I\right)\right)\right)}\right)\end{aligned} \tag{26}$$

Similarly, potential evaluations $y_{t,x^*}^{(a^*)}$'s are used collectively to approximate the mean of $\max\left(\mathbb{E}_{\dagger}(I)\right)$ as

$$\mathbb{M}\left(\max\left(\mathbb{E}_{\dagger}(I)\right)\right) \approx \frac{1}{2m+1}\sum_{t=1}^{2m+1}\max\left(\mathbb{E}_{\dagger,t}(I)\right) \tag{27}$$

and the square root of the mean squared error as

$$\mathbb{S}\left(\max\left(\mathbb{E}_{\dagger}(I)\right)\right) \approx \left[\frac{1}{2m+1}\sum_{t=1}^{2m+1}\left(\max\left(\mathbb{E}_{\dagger,t}(I)\right) - \mathbb{M}\left(\max\left(\mathbb{E}_{\dagger,t}(I)\right)\right)\right)^2\right]^{1/2} \tag{28}$$

such that the consequential improvement is approximated in the subsequent iteration based on a probabilistic distribution of temporary surrogate updates.

In summary, while the first option of acquisition function in Eq. (23) accounts for future costs, it lacks the sensitivity to the immediate sample selection because the signal is diluted by the aggregated future uncertainty. The second option of acquisition function in Eq. (24) attempts to maximize the future improvement but introduces compounding uncertainty with the recursive approximations based on Gaussian distributions, leading to overly conservative or unstable exploration. The proposed CI (as shown in Sect. "Proposed CI acquisition function") strikes the good balance by accounting for the consequential validation cost without compounding the model uncertainty beyond the immediate decision horizon.

3.2.4 Proposed CI acquisition function

Instead of the above two options in 3.2.3, the proposed CI acquisition function is.

$$r\left(\boldsymbol{x}^*,\, a^*\right) = \delta_{a^*,\,A}\frac{\mathbb{E}_*(I(\boldsymbol{x}^{\boldsymbol{*}},\, A))}{\Gamma(A)} + \left(1-\delta_{a^*,\,A}\right)\frac{\mathbb{E}_*\left(\max\left(\mathbb{E}_\dagger(I)\right)\right)}{\Gamma(a^*)+\Gamma(A)} \tag{29}$$

such that the solution $\boldsymbol{x}^{\boldsymbol{*}}$ and fidelity level a^* are selected by solving

$$\max_{\boldsymbol{x}^{\boldsymbol{*}},\,a^*}\left\{\delta_{a^*,\,A}\frac{\mathbb{E}_*(I(\boldsymbol{x}^{\boldsymbol{*}},\, A))}{\Gamma(A)} + \left(1-\delta_{a^*,\,A}\right)\frac{\mathbb{E}_*\left(\max\left(\mathbb{E}_\dagger(I)\right)\right)}{\Gamma(a^*)+\Gamma(A)}\right\} \tag{30}$$

With the CI acquisition function, the subsequent iteration is only considered for the potential selections at LF levels. For each fidelity level, the CI acquisition function is equal to the ratio between improvement and cost over the iterations for which the consequential effects are approximated. The improvement in the current iteration is emphasized by using separate ratios for each fidelity level. This is done to reduce the likelihood that all selections are taken at LF levels. For a potential LF selection, the value of the proposed CI acquisition function is the quotient of the anticipated improvement and evaluation cost for both the current selection and the presumed subsequent selection.

The relationship between the cost and benefit associated with selections in the CI acquisition function is illustrated in Fig. 1. The benefit, which is quantified as the consequential improvement, is depicted along the vertical axis. The evaluation cost required to obtain the improvement is depicted on the horizontal axis. For sample A at the HF level, the evaluation cost and consequential improvement are calculated directly based on the current iteration. In contrast, for sample B at a LF level, there is no improvement in the current iteration. Therefore, the consequential improvement of sample B is calculated by considering the subsequent iteration which is indicated as sample B'. For sample B', the consequential improvement in the subsequent iteration is associated with the evaluation cost accrued over both the current and subsequent iterations.

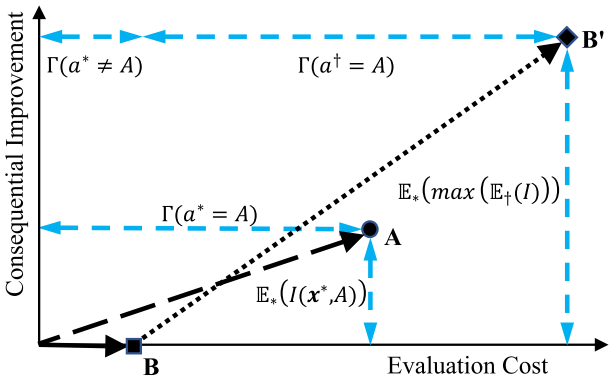


Fig. 1 The consequential improvement and associated evaluation cost are calculated over either the current iteration for a selection at the HF level (long-dashed line) or both the current (solid line) and subsequent (short-dashed line) iterations for a selection at an LF level

For the CI acquisition function in Eq. (29), the effect of an LF selection is quantified based on the presumption, without enforcement, that the subsequent selection will be done to maximize the improvement. This recursive formulation is common in modeling rational decision making such as in dynamic programing. Without the presumption of the subsequent selection being done to maximize the improvement, every possible selection would need to be considered until the sampling budget is exhausted. Because the uncertainty aggregates with compounding approximations based on temporary updates of the surrogate for each future selection, there will be diminishing returns on the viability of the information. Therefore, it is unnecessary to bear the high computational cost of an exhaustive search.

The algorithm of the MFBO with the CI acquisition function is listed in Algorithm 1. The iterative sampling continues until the accumulated evaluation cost exceeds the overall sampling budget. Additionally, the last selection is always at the HF level.

Algorithm 1 MFBO with the CI acquisition function

Inputs: Evaluation costs $\Gamma(a)$'s for $a = A - L, \ldots, A$ Outputs: Optimal solution y_{min} at the HF (validation) level

Set a sampling budget balance;
Take initial samples at each fidelity level $a \in D_2$;
Extract the combined cost of initial evaluations from the sampling budget balance;
WHILE sampling budget balance $\geq \Gamma(A)$:
FOR each $a = A - L, \ldots, A$:
Update hyperparameters θ_a by solving Eqn.(6);
END FOR
Select next sample $\boldsymbol{x}^*$ and fidelity level a^* by solving Eqn.(30);
Calculate $y_{\boldsymbol{x}^*}^{(a^*)}$;
Extract cost $\Gamma(a^*)$ from the sampling budget balance;
END WHILE
Select final sample $\boldsymbol{x}^*$ at fidelity level A.

As mentioned in Sect. 2.4, the major benefit of the CI acquisition function over the multi-fidelity extensions of the KG or MES is the more realistic treatment of LF selections in the actual engineering design process. With the CI acquisition function, there is a consideration that the actual improvement is exclusive to the evaluations at the HF level. In contrast, the multi-fidelity extensions of the KG or MES could lead to over-sampling at LF levels. Furthermore, in the multi-fidelity extensions of the KG and MES, the surrogate needs to be temporarily updated when assessing the acquisition function for each potential selection. In contrast, with the CI acquisition function, temporary updates are only needed when assessing potential selections at LF levels.

4 Demonstrations of the proposed CI acquisition function

In this section, the MFBO based on the proposed CI acquisition function is demonstrated with several analytical and engineering examples. In Sect. 4.1, the setup of the demonstrations is described. In Sect. 4.2, five analytical examples are used to assess the sampling efficiency of the new MFBO method, including two benchmark objective functions and three engineering design examples. In these five examples, the sensitivity of the methodology is analyzed by modifying the evaluation cost ratios of LF levels synthetically. In Sect. 4.3, the simulation-based design of a 25-bar spatial truss and a 200-bar planar truss are used to analyze the sampling efficiency of the CI acquisition function in a situation where the actual costs of evaluations differ between fidelity levels.

4.1 Setup of the MFBO for demonstrations

Here, the CI acquisition function is compared with the multi-fidelity extensions of the KG and the MES acquisition functions for simultaneous selection. With every optimization example in this paper, the goal is to minimize the objective function at the HF level as in Eq. (1). In each example, there are two initial samples at each fidelity level. The MFBO algorithms used in the multi-fidelity extensions of the KG and the MES acquisition functions are similar to the ones in Algorithm 1. In each iteration of MFBO, $\boldsymbol{x}^{*}$ and a^{*} are selected from domains Ω_1 and Ω_2, respectively, to maximize the given acquisition function, which is either the CI, the multi-fidelity extension of the KG, or the multi-fidelity extension of the MES. In the examples, a maximum of 40 iterations are allowed regardless of the sampling cost. The number of initial samples and number of iterations are chosen to allow us to analyze the performance of the new acquisition function. The small number of initial samples matches the lowest possible number of samples to construct a surrogate in real-world applications. The maximum number of iterations is beyond what is usually feasible in engineering design, where each trusted evaluation or physical experiment incurs a high cost. For real-world engineering design applications, the HF level is the one that could be trusted to be used as the validation of the results from other LF evaluations. For demonstration, the HF evaluations in the analytical and engineering design examples in this paper are regarded as ground truth.

The new MFBO method is implemented in Python. The optimizations of the acquisition functions are based on the Broyden–Fletcher–Goldfarb–Shanno algorithm in the SciPy library (Virtanen et al. 2020). The sequential least squares constrained programming algorithm (Kraft 1988) is used when equality and inequality constraints are involved. Additionally, the normalized similarity metric

$$\upsilon_q\left(\mathbf{x},\, \mathbf{x}'\right) = \delta_{q,i}|x_i - x_i\prime|/\mathcal{B}_i \tag{31}$$

is applied to the i-th component in $\boldsymbol{x}$, where $\mathcal{B}_i$ is the range of values possible for the i-th dimension of the solutions. To further reduce the computational expense of hyperparameter fitting, each hyperparameter θ_a associated with the same fidelity level a is set to the same value (Wang 2021). This method reduces the computational expense of hyperparameter fitting. Thus, the hyperparameters only depend on the fidelity level and are independent of solution components. Only two potential evaluations are used when evaluating the acquisition functions based on the subsequent iteration. Additionally, hyperparameters are not updated during the temporary updates of the surrogate. These procedures are done to reduce the computational expense of evaluating the acquisition functions.

The maximum effective expenditure of 60 was selected to demonstrate and represent a typical resource-constrained simulation-based engineering design environment (e.g., based on high-fidelity computational fluid dynamics or multiscale multiphysics simulations), where the total number of allowable high-fidelity evaluations can be extremely severely limited.

4.2 Analytical benchmark examples

In this section, five analytical examples are used to assess the sampling efficiency of the new MFBO method, including two benchmark objective functions and three engineering examples. The two benchmark objective functions include the Forrester and Currin functions. The three engineering examples include the analytical designs of a water tank, a borehole, and a welded beam. The objective functions and constraints of the five analytical benchmark examples are provided in the Appendix. The LF and HF objective functions in these benchmark examples are simply assumed. In engineering practice, LF and HF models can be obtained in different scenarios. An HF model can be the physical experiment while LF models are simulations. The LF and HF models can also correspond to simulations with different fidelity levels (e.g., full-scale vs. coarse-grained and reduced-order simulations, simulation vs. analytical models). When an HF model corresponds to a constrained optimization problem, an LF model can be the one where certain expensive-to-evaluate constraints are omitted.

For the first three analytical examples (the Forrester and Currin functions, and the water tank design), two fidelity levels are chosen, and the evaluation cost ratio is either low, $\Gamma(A) = 2\Gamma(A-1)$, medium, $\Gamma(A) = 4\Gamma(A-1)$, or high, $\Gamma(A) = 10\Gamma(A-1)$. The medium and high-cost ratios match those used recently by (Shu et al. 2021). The low-cost ratio is used for insight into the suitability of the methods for problems where cheaper evaluation methods are not available. The cost ratios for the bi-fidelity problems are also used between the lowest fidelity and highest fidelity levels for the analytical benchmark problems with more than two fidelity levels. These synthetic cost ratios were chosen to represent common engineering scenarios. For instance, a high ratio (e.g., 10 to 100) corresponds to the scenarios of time-consuming simulations vs. rapid analytical calculations, whereas a low to medium ratio (e.g., 2 to 5) represents fine-mesh vs. coarse-mesh finite element models. For the borehole example, three fidelity levels are chosen and the evaluation cost ratio is either low, $\Gamma(A) = 2^{(1/2)}\Gamma(A-1) = 2\Gamma(A-2)$ medium, $\Gamma(A) = 4^{1/2}\Gamma(A-1) = 4\Gamma(A-2)$, or high, $\Gamma(A) = 10^{(1/2)}\Gamma(A-1) = 10\Gamma(A-2)$ For the welded beam example, four fidelity levels are chosen and the evaluation cost ratio is either low, $\Gamma(A) = 2^{(1/3)}\Gamma(A-1) = 2^{(2/3)}\Gamma(A-2) = 2\Gamma(A-3)$medium, $\Gamma(A) = 4^{1/3}\Gamma(A-1) = 4^{2/3}\Gamma(A-2) = 4\Gamma(A-3)$, or high, $\Gamma(A) = 10^{1/3}\Gamma(A-1) = 10^{2/3}\Gamma(A-2) = 10\Gamma(A-3)$. In these five examples, the synthetic cost ratios are chosen to be low enough to prevent the scenario where it is always favorable to sample at the LF level and also to be high enough that the samples at the LF levels with low accuracy still provide useful information about the design solutions. In practical engineering scenarios, the cost ratios of LF and HF sample evaluations can be precisely calculated based on the computational time required to finish an LF or HF simulation, or the time or financial cost of performing a physical experiment.

4.2.1 Convergence behavior of the MFBO

For the analysis of MFBO convergence, 20 runs of optimization with random seeds were taken for each of the five analytical examples. Figures 2, 3, 4 5 and 6 show the search histories for the Currin, Forrester, water tank, borehole, and welded beam design examples, respectively. The same initial samples are taken to compare the three methods of MES (Takeno et al. 2020), KG (Poloczek et al. 2017), and the proposed CI. In each of the figures, the x-axis indicates the effective expenditure, which is the accumulated sampling cost from all prior evaluations. The maximum effective expenditure value in each run was set as 60, except for the borehole example, so that the convergence patterns can be observed. The respective low, medium, and high cost ratios correspond to the different multi-fidelity scenarios that were specified previously.

For the Currin example, the convergence curves of average objective function values are compared in Fig. 2a, and the standard deviations are shown in Fig. 2b. The MES acquisition function performs better than the KG and the proposed CI. Not only the average objective value obtained from the MES converges faster, but also the standard deviations are lower. MES shows the most stable performance. As the cost ratio increases, the performances of CI and KG improve and become closer to that of MES. The results of the one-dimensional Currin example show that CI does not consistently outperform MES or KG when the cost ratio is modest such as $\Gamma(A) = 2\Gamma(A-1)$. This is because the smooth objective in a simple one-dimensional search space can be easily approximated, which reduces the need for extensive LF exploration. When HF

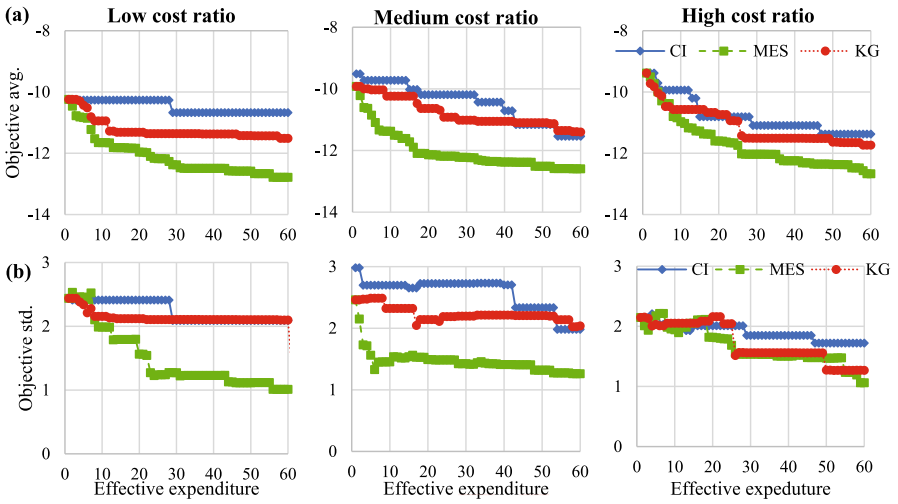


Fig. 2 The comparison of sampling efficiency in reaching the optimal HF solution among CI, KG, and MES for the Currin example with three scenarios of cost ratios (low cost ratio: $\Gamma(A) = 2\Gamma(A-1)$, medium cost ratio: $\Gamma(A) = 4\Gamma(A-1)$, and high cost ratio: $\Gamma(A) = 10\Gamma(A-1)$). **a** Average objective function value, **b** Standard deviation of objective function values

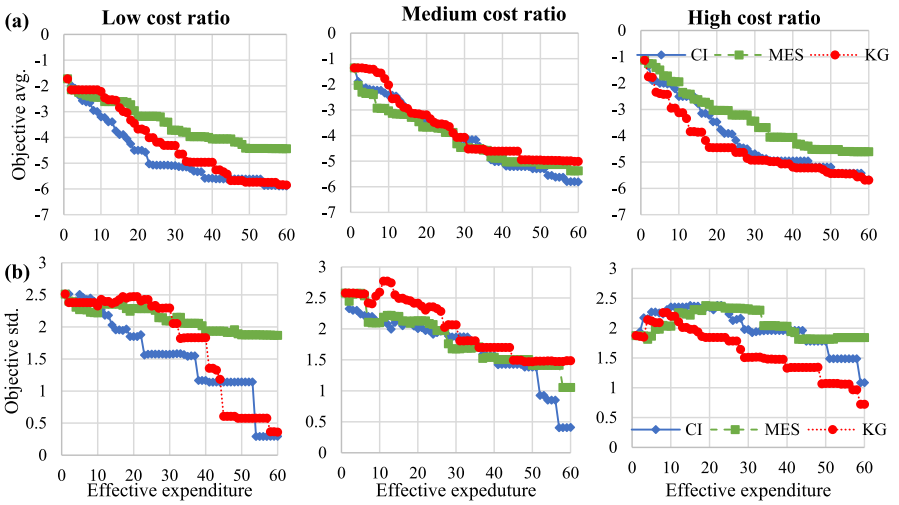


Fig. 3 The comparison of sampling efficiency in reaching the optimal HF solution among CI, KG, and MES for the Forrester example with three scenarios of cost ratios (low cost ratio: $\Gamma(A) = 2\Gamma(A-1)$, medium cost ratio: $\Gamma(A) = 4\Gamma(A-1)$, and high cost ratio: $\Gamma(A) = 10\Gamma(A-1)$). **a** Average objective function value, **b** Standard deviation of objective function values

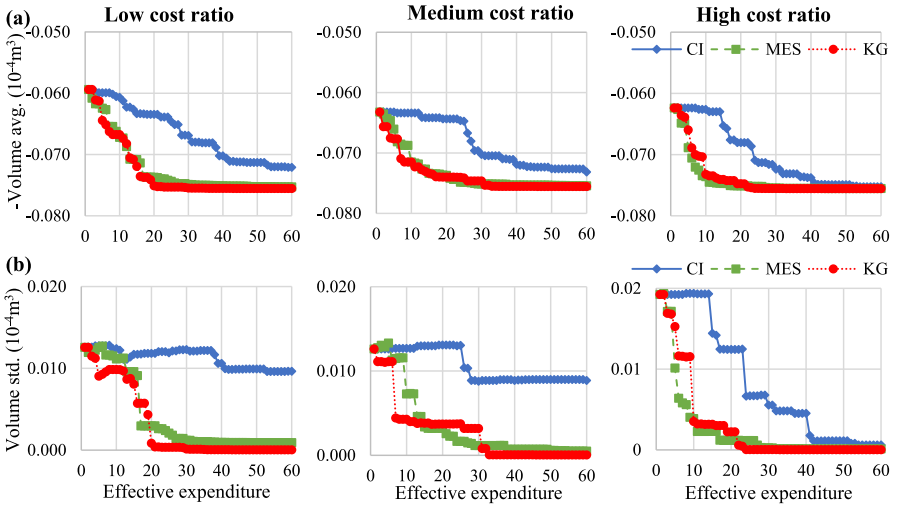


Fig. 4 The comparison of sampling efficiency in reaching the optimal HF solution among CI, KG, and MES for the water tank example with three scenarios of cost ratios (low cost ratio: $\Gamma(A) = 2\Gamma(A-1)$, medium cost ratio: $\Gamma(A) = 4\Gamma(A-1)$, and high cost ratio: $\Gamma(A) = 10\Gamma(A-1)$). **a** Average objective function value, **b** Standard deviation of objective function value

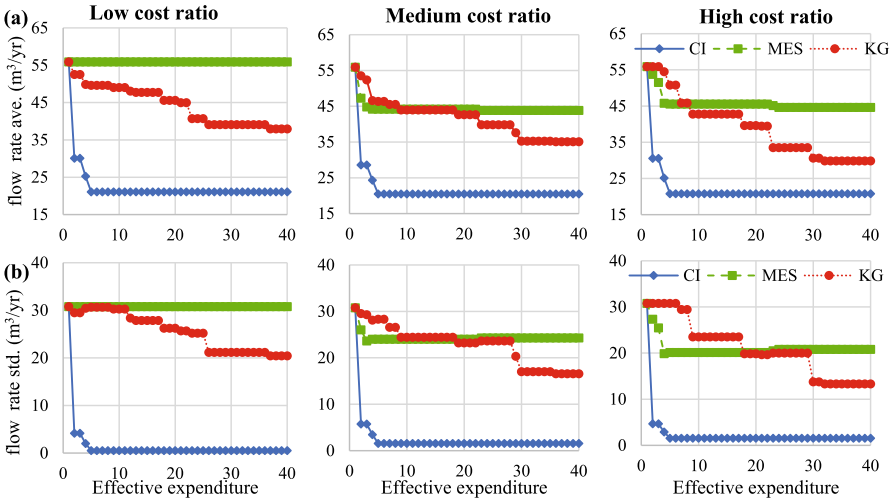


Fig. 5 The comparison of sampling efficiency in reaching the optimal HF solution among CI, KG, and MES for the borehole example with three scenarios of cost ratios (low cost ratio: $\Gamma(A) = 2^{1/2}\Gamma(A-1) = 2\Gamma(A-2)$, medium cost ratio: $\Gamma(A) = 4^{1/2}\Gamma(A-1) = 4\Gamma(A-2)$, and high cost ratio: $\Gamma(A) = 10^{1/2}\Gamma(A-1) = 10\Gamma(A-2)$). **a** Average objective function value, **b** Standard deviation of objective function values

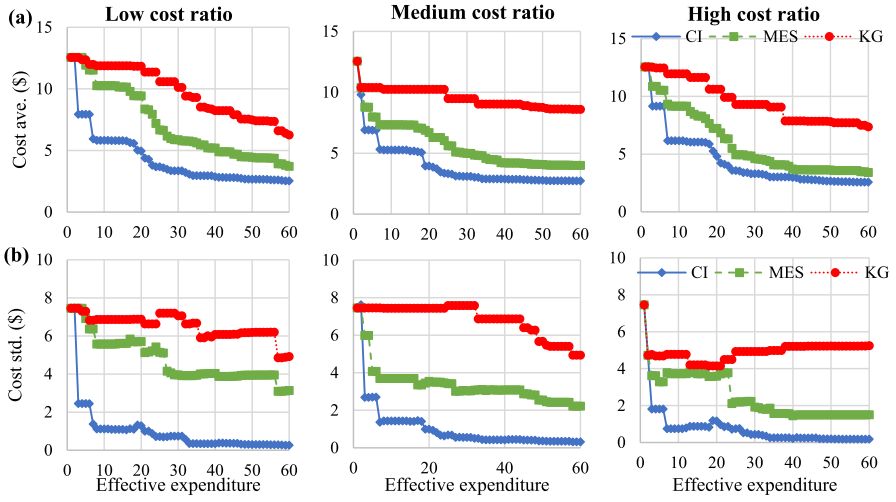


Fig. 6 The comparison of sampling efficiency in reaching the optimal HF solution among CI, KG, and MES for the welded beam example with three scenarios of cost ratios (low cost ratio: $\Gamma(A) = 2^{1/3}\Gamma(A-1) = 2^{2/3}\Gamma(A-2) = 2\Gamma(A-3)$, medium cost ratio: $\Gamma(A) = 4^{1/3}\Gamma(A-1) = 4^{2/3}\Gamma(A-2) = 4\Gamma(A-3)$, and high cost ratio: $\Gamma(A) = 10^{1/3}\Gamma(A-1) = 10^{2/3}\Gamma(A-2) = 10\Gamma(A-3)$.). **a** Average objective function value, **b** Standard deviation of objective function values

evaluations are not costly, the direct HF sampling in MES or KG is efficient enough, and the benefit of considering future HF validation cost as in CI is diminished.

For Forrester's example, the convergence of average objective function values for CI is slightly better than KG and MES (Fig. 3a). The distinction is more obvious in the case of the low-cost ratio. The convergence of CI is also stable with the variance quickly reduced (Fig. 3b). The standard deviations of CI are lower than the other two methods, except for the case of a high-cost ratio where KG is slightly better.

In the water tank example (Fig. 4), the convergences of MES and KG are faster than that of CI during the early stage of the search. The CI can speed up the convergence at a later stage and reduce the gap, especially when the cost ratio increases. The CI exhibits higher variability between different runs during the early stage, and the differences between the average objective values of CI and those of MES and KG are within one standard deviation of the CI results. Therefore, it is still possible that CI can find better optima than MES and KG.

In the borehole example (Fig. 5) where the dimension of search space is eight, CI has the best performance and distinct advantages for all three cost ratios. The average objective values of CI are quickly reduced and show swift convergence. The standard deviations are also quickly reduced to a small value. CI consistently finds the optima quickly.

Similarly, CI performs the best among the three acquisition functions in the welded beam example, where the dimension of search space is four (Fig. 6). It consistently finds the best solution quickly for all three cost ratios. Between MES and KG, MES performs better than KG.

In summary, among the tested five examples, MFBO with the CI acquisition function excels in the two high-dimensional examples (borehole and welded beam). CI has a similar performance as MES and KG in the two two-dimensional examples (Forrester and water tank). CI does not perform as well as MES and KG in the one-dimensional example (Currin). Particularly, CI excels in the scenarios of high-cost ratios. Even for the Currin example where CI does not perform as well as MES and KG, the difference becomes small with the high cost ratio. These results show that the proposed CI can avoid unnecessary LF sampling when the cost of HF sampling is significantly higher, which MFBO methods have the tendency to do. The results also indicate that the main benefit of LF sampling is its usage for cost-effective exploration. In the early stage of search, a mixture of LF and HF samples is typically seen, where multiple LF samples may be utilized before a HF sample is taken. As the objective evaluation gets closer to the optimum value and the improvement diminishes, more LF samples are taken to explore the solution space to search for any further improvement. Furthermore, as the dimension of the problem increases, more samples need to be taken, and the advantage of the CI becomes more prominent.

4.2.2 Computational cost comparison

To assess the computational costs of the proposed CI in comparison with KG and MES, we also recorded the CPU time. With the multi-fidelity extension of either the MES or KG acquisition function, each evaluation requires a costly temporary surrogate update. Similarly, each evaluation of the CI acquisition function at the HF

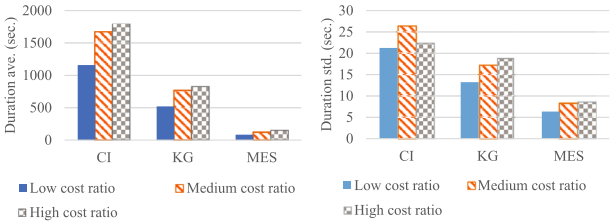


Fig. 7 The comparison of averages and standard deviations of CPU time (in seconds) for the Currin example

level requires a temporary surrogate update, in addition to a maximization procedure. However, neither the temporary surrogate update nor maximization is required for an evaluation of the CI acquisition function at any LF level. Therefore, as additional fidelity levels are added, the increase of computational cost in the CI method is slower than those of the other two methods.

The experiments were conducted on a computer equipped with an Intel Core i5-11400H 2.70GHz CPU and 8.00 GB RAM. The CPU time of CI is about twice as much as the one of KG. For instance, the computational expenses for the Currin example based on the CI, KG, and MES acquisition functions with the low-, medium-, and high-cost ratios are listed in Fig. 7. The computational time for CI in this bi-fidelity example is about 13 times longer than MES and 2.2 times longer than KG. As the cost ratios increase, the computational costs show linear growth for all three acquisition functions. The computational expense of CI is not significantly higher than that of KG. They are in the same order of magnitude.

4.2.3 Effects of initial sample size

We also evaluated the effects of the number of initial samples (NIS) on the performance of MFBO. While a higher NIS expands the exploration capability across a broader search space, it also elongates the solution time and increases the computational costs. To assess the value of adding more initial samples, the borehole example was tested. Four different NIS values (2, 4, 7, 10) were chosen and tested, and 20 optimization runs were performed for each value and each of the three acquisition functions with varied seeds.

Figure 8 depicts the average and standard deviation of objective functions for the Borehole example across different NIS values. Notably, small variations were observed in the optimal objective values due to the stochastic nature of the search. The identified optimal objective values by each MFBO method are consistent for different NIS values, which indicates that the search result is not sensitive to the NIS values. In Fig. 9, the computational costs associated with different NIS values are compared. As expected, the total computational time increases as more initial samples are taken. Therefore, choosing two initial samples to start the MFBO search in this study is cost-effective.

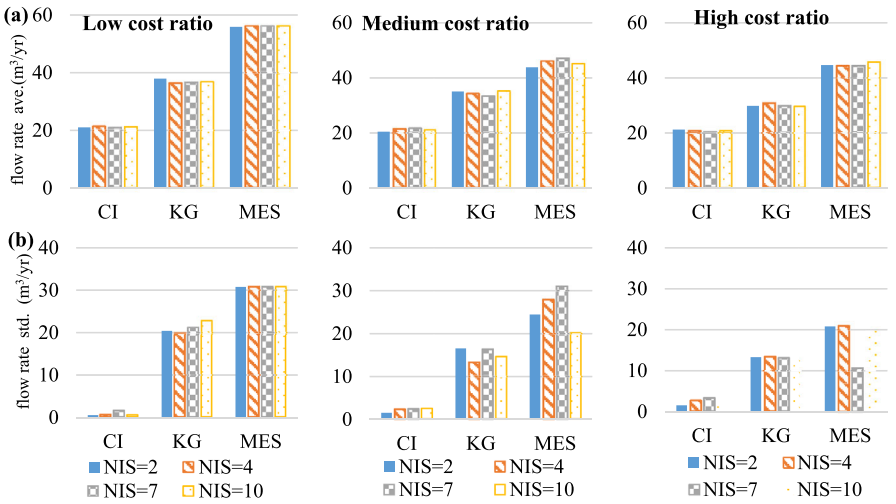


Fig. 8 The optimal objective values with different NIS values for the borehole example. **a** Average objective function value, **b** Standard deviation of objective values

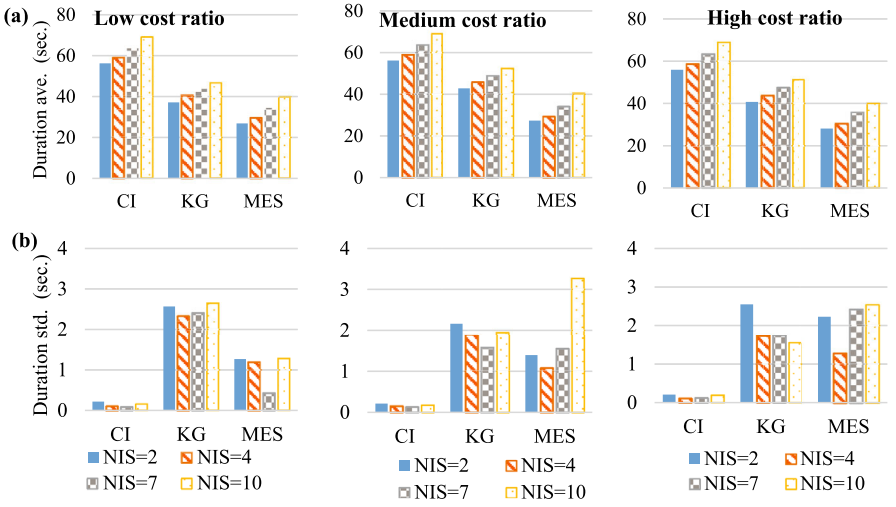


Fig. 9 The computational costs of the MFBO are associated with different NIS values for the borehole example. **a** Average value, **b** Standard deviations

4.3 Simulation-based design of truss

4.3.1 25-bar spatial truss

The simulation-based design of a 25-bar spatial truss used here is a multi-fidelity extension of the single-fidelity version used previously by several other researchers (Awad 2021; Schmit and Farshi 1974; Venkayya 1971). The truss is depicted in Fig. 10.

In this example, two LF levels are added to illustrate a total of three fidelity levels. The black-box objective function for each fidelity level a is

$$f(\boldsymbol{x}, a) = f_0(\boldsymbol{x})\left(\delta_{a,A} P_s(\boldsymbol{x}) + \left(1 - \delta_{a,A-2}\right) P_d(\boldsymbol{x}) + 1\right)^2 \tag{32}$$

where the design variables $\boldsymbol{x}$ include eight grouped cross-sectional areas of the bars, $\Delta_1, \ldots, \Delta_8$. The weight of the truss, $f_0(\boldsymbol{x})$, is penalized by the constraints of displacement $P_d(\boldsymbol{x})$ and stress $P_s(\boldsymbol{x})$. The weight and the penalties are black-box functions of the design solution. Each penalty is the relative extent of the stress or displacement outside of its preferred range.

At each fidelity level, the evaluation of truss weight is performed with the same general procedure. However, the stress penalty is ignored at the middle fidelity level, and both the stress and displacement penalties are ignored at the lowest fidelity level. The calculations of the stress and displacements require finite-element analysis (FEA). Therefore, the computational costs of the objective function evaluations differ between the three fidelity levels.

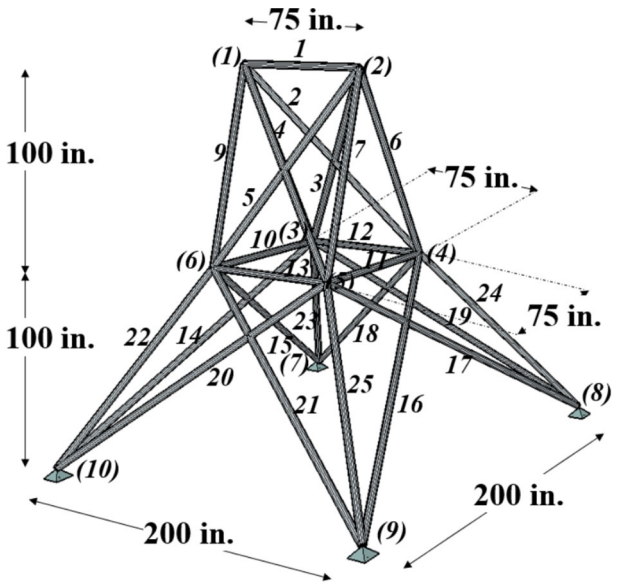


Fig. 10 The 25-bar spatial truss example with the indices of bars (displayed without parenthesis) and the indices of nodes (displayed within parenthesis)

In this example, the design with the lowest objective value is considered the best. A stress or displacement value outside of the preferred range is of acceptable risk as long as the design offers sufficient weight reduction. The square sum of the penalty term in Eq. (32) is to attenuate the influence of the penalties and maintain the balance of the objective function (Homaifar et al. 1994). The goal is to minimize the objective function in Eq. (32) at the HF level by solving

$$\min_{\boldsymbol{x}\in\Omega_1} f_0(\boldsymbol{x})(P_s(\boldsymbol{x}) + P_d(\boldsymbol{x}) + 1)^2 \tag{33}$$

$$s.t.0.01in^2 \leq \Delta_i \leq 240\text{in}^2(i = 1, \ldots, 8) \tag{34}$$

The eight variables, the preferred range for the stress for each group of bars (Venkayya 1971), and the loading conditions on the nodes of the truss (Farshi and Alinia-ziazi 2010) are provided in Table 1. The preferred range for the displacement of each node is between 0 and 0.35. The evaluation cost ratios are calculated based on FEA simulations, which are $\Gamma(A) \cong 1.218546\Gamma(A-1) \cong 11.535356\Gamma(A-2)$. The sensitivity of MFBO performance to the cost and the accuracy differences of fidelity levels are analyzed. Three scenarios are considered. In the first scenario, the performances of MFBO are compared with all three fidelity levels considered simultaneously. In the second scenario, only two fidelity levels are used where the lowest fidelity level is not considered. That is, the evaluation cost ratio for the bi-fidelity case

Table 1 The variable groups and preferred stress range in MPa on the bars (Venkayya 1971) and the loading conditions, with forces in N, on the nodes of the truss (Schmit and Farshi 1974)

Variable group	Indices of bars contained	Preferred range for stress (MPa)
Δ_1	1	− 241.950823 to 275.79029
Δ_2	2–5	− 79.910237 to 275.79029
Δ_3	6–9	− 119.313775 to 275.79029
Δ_4	10–11	− 241.950823 to 275.79029
Δ_5	12–13	− 241.950823 to 275.79029
Δ_6	14–17	− 46.60166 to 275.79029
Δ_7	18–21	− 46.60166 to 275.79029
Δ_8	22–25	− 76.4077003 to 275.79029

Index of node	Condition 1			Condition 2		
	P_x	P_y	P_z	P_x	P_y	P_z
1	4.448	44.48	− 22.24	0	88.96	− 22.24
2	0	44.48	− 22.24	0	− 88.96	− 22.24
3	2.224	0	0	0	0	0
6	2.224	0	0	0	0	0

is $\Gamma(A) \cong 1.218546\Gamma(A-1)$. In the third scenario, the medium level is excluded and only two other fidelity levels are used such that $\Gamma(A) \cong 11.535356\Gamma(A-1)$.

The convergence behaviors of CI, KG, and MES are compared in Fig. 11. The searches converged after 40 iterations. The proposed CI demonstrates the best convergence rates across all fidelity levels with the lowest average objective values and the smallest standard deviations. The overall performance of KG is slightly better than that of MES. Particularly KG performs much better than MES in the case when the medium level is excluded.

Computational costs are compared in Fig. 12, MES exhibits the lowest average CPU time, similar to the Currin example. However, the standard deviation of MES slightly exceeds those of KG and CI. It is seen that the standard deviations of computational

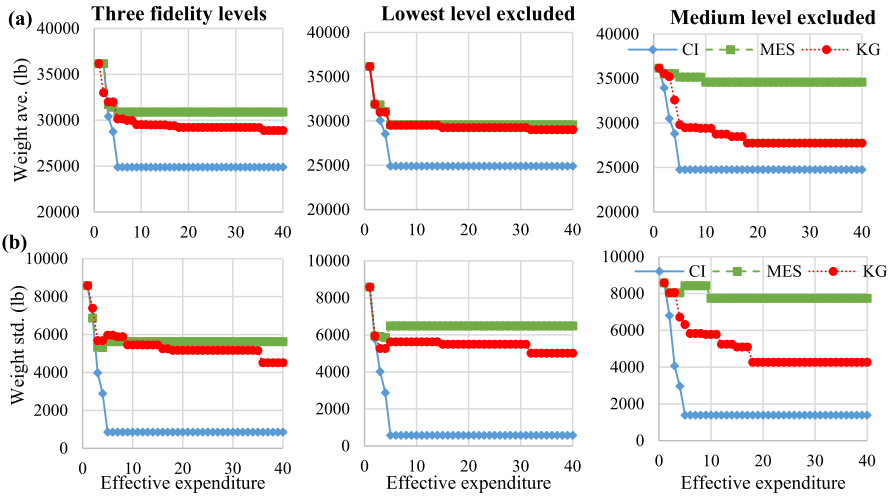


Fig. 11 The best objective value for the 25-bar truss example over the effective expenditure as the accumulated sampling cost for the CI acquisition function, the multi-fidelity extension of the KG, and the multi-fidelity extension of the MES for the cases of three fidelity levels, the lowest fidelity level excluded, and the medium level excluded. **a** Average objective function value, **b** Standard deviation of objective function values

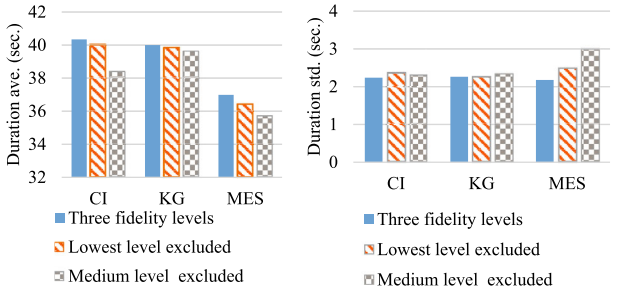


Fig. 12 Comparison of average and standard deviation CPU time (in seconds) for the 25-bat truss example

time for all three MFBO methods are small relative to the average values, suggesting that the performance of MFBO algorithms is fairly consistence in general.

4.3.2 200-bar planar truss.

The 200-bar planar truss structure, shown in Fig. 13, has 77 nodes. Its members are organized into 29 groups, detailed in Table 2. Stress constraints are applied to the members, limited to $\pm$ 10 ksi. The material's unit weight is 0.283 lbs/in^3, and its Young's modulus is 30,000 ksi. The truss is designed to withstand three independent loading scenarios:

1 Case 1:a 1.0 kip load in the positive X-direction applied at nodes 1, 6, 15, 20, 29, 34, 43, 48, 57, 62, and 71
2 Case 2: a 10.0 kip load in the negative Y-direction at nodes 1–6, 8, 10, 12, 14, 15–20, 22, 24, 26, 28, 29–34, 36, 38, 40, 42, and continuing to nodes 70–75 (specifically nodes listed in the original text); and
3 Case 3: a combined loading condition where both loadings in Cases 1 and 2 are applied simultaneously.

In the MFBO test, FEA simulations for all three cases are conducted for each objective function evaluation and the maximum stress and displacement values across the entire structure are identified. These worst-case values are then used to calculate the constraint penalties ($P_s(\boldsymbol{x})$ and $P_d(\boldsymbol{x})$) to ensure that the final design is robust to handle the most critical conditions. The calculation of fidelity levels for the 200-bar truss problem follows the same methodology as the 25-bar truss example. Specifically, the formulation of the objective function and the methodology for defining fidelity levels, as described by Eq. (33) in the 25-bar truss example, is directly applied. The main difference lies in the numerical parameters and specific details of the 200-bar truss, including the numbers of members and nodes, and potentially the constraints or loading conditions. Similar to the 25-bar truss, the fidelity levels are differentiated by the inclusion or exclusion of stress and displacement penalties in the objective function, which causes different computational costs due to the underlying FEA. Therefore, while the fundamental multi-fidelity approach is similar, the specific implementation is tailored to the unique parameters and scale of the 200-bar truss problem.

Figure 14 illustrates the convergence behaviors of CI, KG, and MES for the different fidelity levels. CI generally exhibits the most rapid convergence for the average objective value and achieves the best performance in all scenarios. Furthermore, CI demonstrates the most stable and consistent convergence, characterized by a quick initial decline followed by an early plateau. MES shows a slightly slower convergence than KG at the early stage, but it ultimately reaches the objective values that are comparable to KG. When the lowest fidelity level or medium fidelity level excluded, the performances of MES and KG are slightly improved in comparison with the scenario with all three fidelity levels. Notably, the final average objective values achieved by all three methods across all fidelity level configurations are quite similar, suggesting a less pronounced effect of fidelity level selection than initially anticipated.

Figure 15 compares the computational costs. CI demonstrates significantly shorter

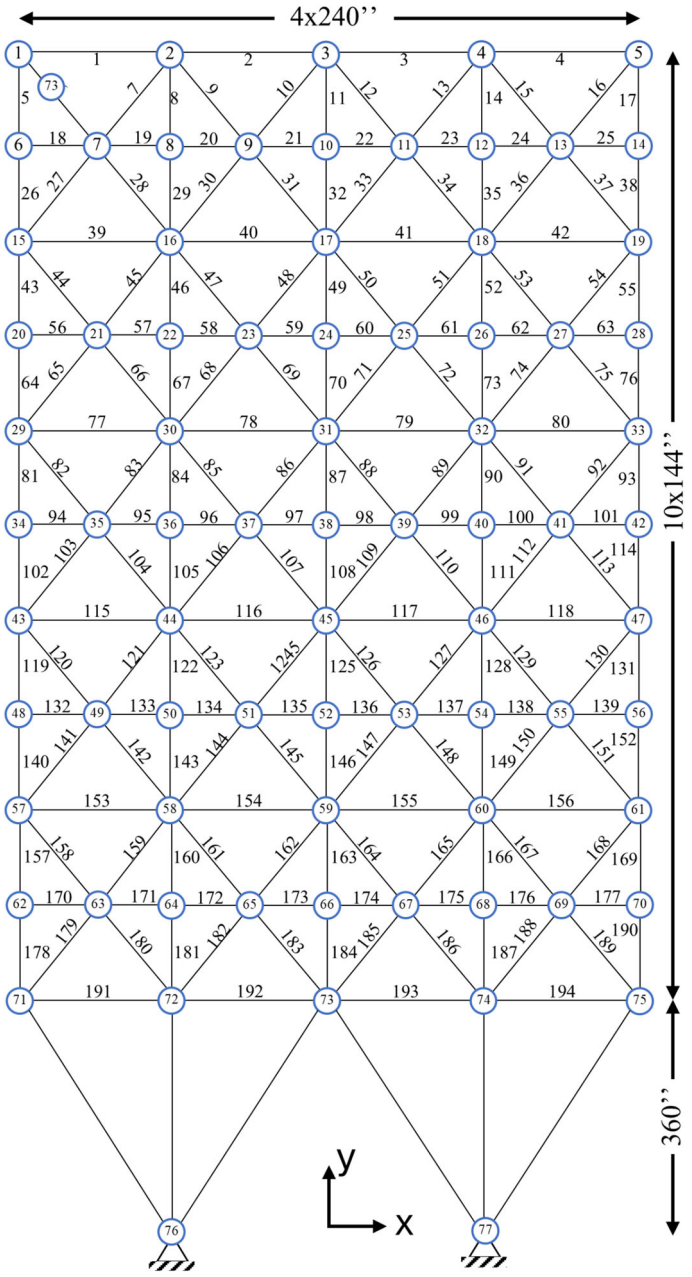


Fig. 13 The 200-bar truss example with the indices of bars and the indices of nodes

Table 2 Group details for the 200-bar truss example

Group	Members	Group	Members
1	M_1, M_2, M_3, M_4	16	M_{82}, M_{83}, M_{85}, M_{86}, M_{88}, M_{89}, M_{91}, A_{92}, A_{103}, A_{104}, A_{106}, A_{107}, A_{109}, A_{110}, A_{112}, A_{113}
2	A_5, A_8, A_{11}, A_{14}, A_{17}	17	A_{115}, A_{116}, A_{117}, A_{118}
3	A_{19}, A_{20}, A_{21}, A_{22}, A_{23}, A_{24}	18	A_{119}, A_{122}, A_{125}, A_{128}, A_{131}
4	A_{18}, A_{25}, A_{56}, A_{63}, A_{94}, A_{101}, A_{132}, A_{139}, A_{170}, A_{177}	19	A_{133}, A_{134}, A_{135}, M_{136}, M_{137}, M_{138}
5	M_{26}, M_{29}, M_{32}, M_{35}, M_{38}	20	M_{140}, M_{143}, M_{146}, M_{149}, M_{152}
6	M_6, M_7, M_9, M_{10}, M_{12}, M_{13}, M_{15}, M_{16}, M_{27}, M_{28}, M_{30}, M_{31}, M_{33}, M_{34}, M_{36}, M_{37}	21	M_{120}, M_{121}, M_{123}, M_{124}, M_{126}, M_{127}, M_{129}, M_{130}, M_{141}, M_{142}, M_{144}, M_{145}, M_{147}, M_{148}, M_{150}, M_{151}
7	M_{39}, M_{40}, M_{41}, M_{42}	22	M_{153}, M_{154}, M_{155}, M_{156}
8	M_{43}, M_{46}, M_{49}, M_{52}, M_{55}	23	M_{157}, M_{160}, M_{163}, M_{166}, M_{169}
9	M_{57}, M_{58}, M_{59}, M_{60}, M_{61}, M_{62}	24	M_{171}, M_{172}, M_{173}, M_{174}, M_{175}, M_{176}
10	M_{64}, M_{67}, M_{70}, M_{73}, M_{76}	25	M_{178}, M_{181}, M_{184}, M_{187}, M_{190}
11	M_{44}, M_{45}, M_{47}, M_{48}, M_{50}, M_{51}, M_{53}, M_{54}, M_{65}, M_{66}, M_{68}, M_{69}, M_{71}, M_{72}, M_{74}, M_{75}	26	M_{158}, M_{159}, M_{161}, M_{162}, M_{164}, M_{165}, M_{167}, M_{168}, M_{179}, M_{180}, M_{182}, M_{183}, M_{185}, M_{186}, M_{188}, M_{189}
12	M_{77}, M_{78}, M_{79}, M_{80}	27	M_{191}, M_{192}, M_{193}, M_{194}
13	M_{81}, M_{84}, M_{87}, M_{90}, M_{93}	28	M_{195}, M_{197}, M_{198}, M_{200}
14	M_{95}, M_{96}, M_{97}, M_{98}, M_{99}, M_{100}	29	M_{196}, M_{199}
15	M_{102}, M_{105}, M_{108}, M_{111}, M_{114}		

time than KG and MES across all fidelity level configurations. KG exhibits substantially longer computational time than CI, while MES is the most costly among the three. Utilizing all three fidelity levels also requires much longer computational time than the other two configurations. The computational costs for excluding the lowest or the medium level are comparable.

4.3.3 Impact of problem size on performance

The results from the 25-bar and 200-bar truss examples show the stable and reliable performance of CI. As the dimension of the problem increases, the advantage of CI in comparison with KG and MES is even more prominent. The exploration capability in KG is particularly beneficial for the high-dimensional design space of the 200-bar truss, although its convergence remains more fluctuating than CI. MES shows some performance improvement as the problem size increases. In the 200-bar truss, the performance of MES is much closer to that of KG in comparison with the 25-bar truss example, even if its initial convergence is generally slower than CI and KG.

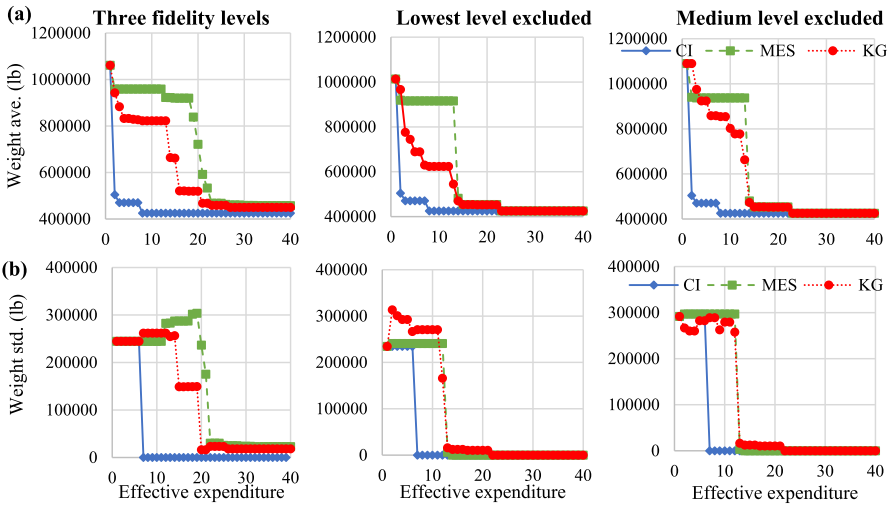


Fig. 14 The best objective value for the 200-bar truss example over the effective expenditure as the accumulated sampling cost for the CI acquisition function, the multi-fidelity extension of the KG, and the multi-fidelity extension of the MES for the cases of three fidelity levels, the lowest fidelity level excluded, and the medium level excluded. **a** Average objective function value, **b** Standard deviation of objective function values

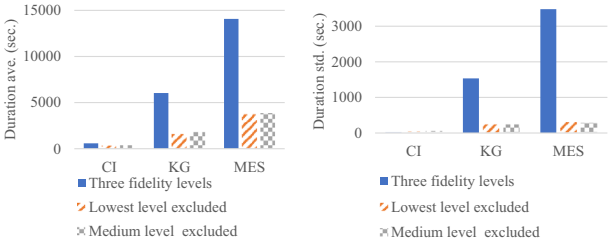


Fig. 15 Comparison of average and standard deviation CPU time (in seconds) for the 200-bar truss example

The role of fidelity levels shows some variation with problem size. For both truss examples, applying three fidelity levels does not guarantee the improved final objective values. However, it can influence the convergence behaviors. Excluding the lowest fidelity level appears to be less impactful than removing the medium level for both problem sizes, which indicates the value of information obtained at the medium fidelity level evaluation. Overall, the computational saving by excluding one fidelity level becomes more prominent for the 200-bar truss due to its increased complexity. Interestingly, the final average objective values are quite similar across different fidelity level configurations for the 200-bar truss. For large problems, the choice of fidelity levels might mostly differentiate the computational costs.

Note that the multi-fidelity approach to omit certain constraints is practical in complex engineering scenarios when the evaluations of some constraints require expensive

simulations or experiments. When only LF models with fewer constraints are evaluated, it may find infeasible solutions based on the LF samples. The risk of finding and treating an infeasible solution as the optimum is inherently mitigated in the proposed CI, because CI treats HF samples as the ultimate way to obtain the optimal solution and prevents the over-sampling at LF levels.

One major limitation of BO is the scalability of Gaussian process regression model that is used as the surrogate. The update of Gaussian process model requires the calculation of the inverse of the sample covariance matrix, which has the time complexity of $O(\mathcal{N}^3)$ in terms of the sample size $\mathcal{N}$. As the dimension of searching space increases, the required number of samples grows exponentially. Therefore, solving large-scale problems has been a challenge for BO. New approaches such as sparse Gaussian process based BO (Aydogdu and Wang 2024) can significantly improve the scalability.

4.3.4 Comparison of conventional optimization methods

To contextualize the performance of the proposed CI-based MFBO methodology, a comparative study was conducted against two well-established, conventional optimization algorithms: a gradient-based Quasi-Newton method and a population-based Genetic Algorithm (GA). These methods were chosen to represent two fundamentally different optimization philosophies: the Quasi-Newton method exemplifies a "white-box" local optimization strategy that leverages detailed gradient information, while the GA represents a "black-box" global search metaheuristic that operates without such information.

Both the Quasi-Newton and GA methods were implemented in the MATLAB environment using the highly-regarded *fmincon* and *ga* solvers from the Optimization (The MathWorks 2025b) and Global Optimization (The MathWorks 2025a) Toolboxes, respectively. The use of these standardized tools ensures a robust and reproducible benchmark. To enable the efficient application of the gradient-based method, an analytical sensitivity analysis was implemented within the Finite Element Method (FEM) solver to compute the exact gradients of the objective and constraint functions. The accuracy of these analytical gradients was verified against numerical finite-differencing to ensure the reliable performance of the Quasi-Newton algorithm. The source codes for these implementations have been made publicly available on MATLAB Central File Exchange for transparency and future research (Aydogdu 2025a, b). The additional feasibility and performance reports for these benchmark runs are provided in the Supplementary Material.

A critical aspect of this comparison was maintaining experimental fairness. As the true bottleneck in complex engineering design is the computational cost of each function evaluation (i.e., each FEM simulation), a strict budget was imposed. The maximum number of function evaluations was set to 40 for both methods, directly corresponding to the "Effective expenditure" budget used for the MFBO methods. For the GA, this budget was implemented with a population size of 4 over 10 generations. To ensure a direct and unambiguous comparison, the CI-MFBO results used in this section correspond to the single high-fidelity case (utilizing all three fidelity levels), as both the Quasi-Newton and GA methods operate on the fully constrained, high-fidelity problem definition.

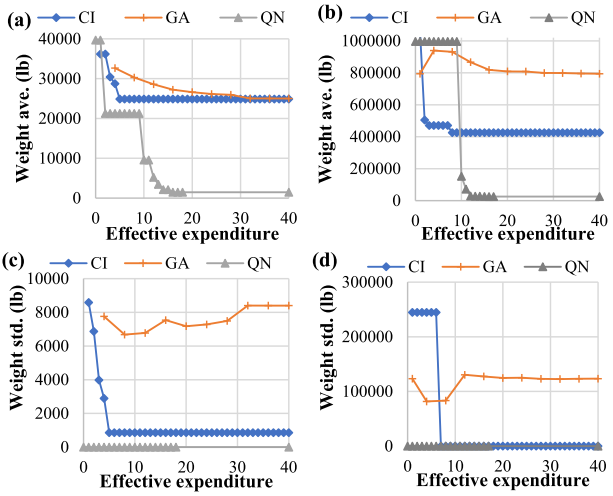


Fig. 16 Comparison of convergence histories for the CI-MFBO, Quasi-Newton (QN), and Genetic Algorithm (GA) methods over 10 independent runs. **a** The average penalized objective value for the 25-bar truss, **b** The average penalized objective value for the 200-bar truss, **c** The standard deviation of the objective value for the 25-bar truss, **d** The standard deviation for the 200-bar truss

The optimization process for both benchmark methods was repeated 10 times with different random seeds to account for any stochastic effects and ensure statistical robustness. The average and standard deviation of the penalized objective function over the 10 runs were recorded. The convergence histories for the 25-bar and 200-bar truss examples are compared in Fig. 16, where the mean objective values are shown in Fig. 16(a–b) and the standard deviations are shown in Fig. 16(c–d), respectively. It is important to note that a direct comparison of computational time (CPU time) was omitted, as the CI-MFBO was implemented in Python while the benchmark methods were in MATLAB, making such a direct comparison scientifically unsound. In Fig. 16(b), the Quasi-Newton objective values appear significantly lower (~ 25,000 lb) than the GA results (~ 794,000 lb), because the Quasi-Newton method converged to marginally infeasible solutions with approximately 0.03% constraint violation, whereas the GA solutions were strictly feasible. The standard deviation of zero for the Quasi-Newton method in Fig. 16(c–d) reflects its deterministic nature.

The convergence plots in Fig. 16(a–b)reveal stark differences in the performance of the three algorithms. For the 200-bar truss problem, the Quasi-Newton method demonstrates exceptionally rapid convergence to a solution with a penalized objective value significantly lower than that of the other two methods. The CI-MFBO method also shows strong sample efficiency, quickly finding a high-quality solution region and outperforming the GA, which makes slow and incremental progress. The standard deviation plots in Fig. 16(c–d) highlight the deterministic nature of the Quasi-Newton method (zero variance) and the initially exploratory, higher-variance search of the CI-MFBO. A similar trend is observed for the 25-bar truss, where the Quasi-Newton

method again finds the best final solution, though the CI-MFBO's initial convergence is competitively fast.

It should be noted that the penalized objective value can be misleading, as it combines both the true weight and constraint violations. A final feasibility analysis was performed on the terminal solution from each of the 10 runs for each algorithm. For the 200-bar truss example, the GA achieved a 100% feasibility rate, whereas the Quasi-Newton method converged to solutions that were marginally infeasible, with a 0.03% stress constraint violation. This slight violation prevented the objective function from incurring a heavy penalty, which is the reason that the Quasi-Newton results in Fig. 16(b) appear significantly lower (~ 25,000 lb) than the strictly feasible GA results (~ 794,000 lb). For the more complex 25-bar spatial truss in Fig. 16(a), a critical distinction emerged, where the Quasi-Newton method failed to find a single feasible solution in any of the 10 runs (0% feasibility rate). The search was trapped in a deep but significantly infeasible local minimum (with constraint violations exceeding 30%). In contrast, the GA achieved a 90% feasibility rate, and the CI-MFBO method is shown to have a similarly high success rate, underscoring its ability to navigate complex constraint boundaries.

This feasibility result is crucial for a correct interpretation. The seemingly superiority of the Quasi-Newton method in terms of objective values is illusory due to its inability to strictly satisfy the design constraints during the search with restricted resources. Its improved performance on the 200-bar problem is because of the relatively smooth constraint surfaces which allow the search to reach the constraint boundary. Yet, the search failed to strictly follow the constraint boundary and satisfy the constraints. In problems with highly nonlinear constraints, discrete variables, or highly nonlinear objective functions, which are very common in real-world engineering, such gradient-based methods can easily fail.

Furthermore, the comparison between GA and CI-MFBO demonstrates the core advantage of Bayesian Optimization. GA's methodology requires a large number of objective function evaluations for its evolutionary process to be effective and is thus ill-suited for a low-budget scenario. The CI-MFBO, designed specifically for the maximum sample efficiency in computationally expensive problems, demonstrates its strength by rapidly learning a surrogate model of the problem and making intelligent and informed decisions, thereby achieving superior performance within the strict evaluation budget. This highlights the critical advantage of surrogate-based Bayesian approaches when simulation cost is a dominant factor.

5 Concluding remarks

In this paper, a new CI acquisition function is proposed to further increase the sampling efficiency of MFBO. With the new CI acquisition function, the sample solution and the fidelity level of the objective evaluation are selected simultaneously. Simultaneous selection is done to synergistically maximize the proposed CI acquisition function. In engineering design, the actual improvement needs to be validated with samples at the HF level, which usually involves physical experiments. The proposed CI acquisition

function incorporates the consequential and actual improvement of the design objective. This implies that the final selection will always be at the HF level in applications. It also avoids the situation that samples are only taken at LF levels.

The proposed CI acquisition function was applied to analytical functions and simulation-based design optimization examples. The sampling efficiency of the CI acquisition function is compared to those of the other MFBO methods for simultaneous selection, including the multi-fidelity extensions of the KG and MES. The sensitivity of the sampling efficiency to different evaluation cost ratios is analyzed. Finally, a multiscale materials and device design example is developed to assess the sampling efficiency in a practical engineering design situation, where the computational effort of surrogate construction is justified by the substantial cost savings from the expensive simulations. In the multiscale design example, a multi-fidelity extension of a 25-bar spatial truss is solved. For simple analytical examples, the results indicate that the sampling efficiency depends on the evaluation cost ratios. For the simulation-based design examples where the true evaluation costs are considered, the results show that the CI has a decisive advantage in the sampling efficiency over the other MFBO methods for simultaneous selection.

One limitation of MFBO is the additional computational expense required to evaluate the acquisition function due to the use of multiple surrogate models. In comparison to the multi-fidelity extensions of the KG or MES, the computational expense is reduced for the CI acquisition function because the additional surrogate is not needed for evaluations of potential selections at the HF level. In this paper, the computational expense of fitting the surrogate was reduced by building the multi-fidelity surrogate without cross-fidelity correlation coefficients. The cost of surrogate fitting was further reduced by applying an equal value constraint on the hyperparameters. Furthermore, only two potential evaluations were used when evaluating the acquisition functions based on the subsequent iteration. Additionally, hyperparameters are not updated during the temporary updates of the surrogate. These strategies are taken to reduce the computational expense of evaluating the CI acquisition function. However, the accuracies of the surrogate and the acquisition function evaluations are also reduced. Reducing the accuracies of the surrogate and the acquisition function evaluations could reduce the sampling efficiency of MFBO. Further work on quantifying the impact of these strategies on the sampling efficiency and on the computational expense associated with the algorithm itself is needed.

When the CI is applied in MFBO, the prioritization of the computational expense versus the sampling efficiency can depend on the needs of the optimization problem. While LF evaluations are less expensive, they are also less informative. Therefore, a greater total number of samples are generally required for MFBO than for BO without multiple fidelity levels. The sampling efficiency of MFBO might not be improved over that of single-fidelity BO if LF evaluations do not provide enough cost savings. Conversely, the CI method may not provide significant advantages for simple problems of low dimensionality, when cost differences between fidelity levels are small, or low-fidelity models deviate significantly from high-fidelity ones. In those cases, multi-fidelity approaches are less effective, and LF evaluations may not be cost-effective since they are relatively inaccurate in comparison to the cost saving. Additionally, regardless of the evaluation costs, evaluations at LF levels must still be sufficiently

accurate to be useful for informing the surrogate. Wasteful selections at LF levels may be avoided with the proposed CI acquisition function. However, the resources that have already been expended for the initial samples at LF levels cannot be recouped. Future work can be done to build models to predict whether any of the LF levels are useful based on the initial samples for improving the overall cost-effectiveness. By eliminating ineffective fidelity levels, the computational expenses associated with updating the surrogate and evaluating the acquisition function can be further reduced. Future studies will also include how to extend the CI acquisition function for multi-objective optimization and parallel sampling. However, either of these approaches will decrease the sampling efficiency compared to the single-objective BO with sequential sampling. It is also beneficial to evaluate how the CI acquisition function can handle complex cost structures in real-world engineering scenarios, such as time-varying or stochastic cost ratios, considerations of setup time, parallel or batch evaluations, or uncertainty in the computational or experimental expenses. New strategies of incorporating dynamic cost models into the CI acquisition function, where costs are updated iteratively based on runtime estimates or probabilistic distributions, may improve the robustness of performance.

Appendix

For the Forrester example (Forrester et al. 2007), the multi-fidelity objective function is

$$f(x, a) = (6x_1 - 2)^2 \sin(12x_1 - 4)\,\delta_{a,A} + \left(0.5\,(6x_1 - 2)^2 \sin(12x_1 - 4) + 10\,(x_1 - 0.5) + 5\right)\delta_{a,A-1} \tag{35}$$

The goal is to minimize the objective function in Eq. (35) at the HF level as

$$\min_{x \in \Omega_1} (6x_1 - 2)^2 \sin(12x_1 - 4) \tag{36}$$

$$\text{s.t.} 0 \le x_1 \le 1 \tag{37}$$

For the Currin example (Tran et al. 2020b), the multi-fidelity objective function is

$$f(x, a) = \left(1 - e^{-(2x_2)^{-1}}\right)\left(\frac{2300x_1^3 + 1900x_1^2 + 2092x_1 + 60}{100x_1^3 + 500x_1^2 + 4x_1 + 20}\right)\delta_{a,A} + \left(\frac{f((x_1 + 0.05,\, x_2 + 0.05),\, 2) + f((x_1 + 0.05,\, \max(0,\, x_2 - 0.05)),\, 2)}{4}\right)\delta_{a,A-1} \tag{38}$$

The goal is to minimize the objective function in Eq. (38) at the HF level as

$$\min_{x \in \Omega_1} \left(1 - e^{-(2x_2)^{-1}}\right)\left(\frac{2300x_1^3 + 1900x_1^2 + 2092x_1 + 60}{100x_1^3 + 500x_1^2 + 4x_1 + 20}\right) \tag{39}$$

$$\text{s.t.} 0 \leq x_1,\, x_2 \leq 1 \tag{40}$$

In the water tank example (Belegundu and Chandrupatla 2019), the multi-fidelity objective function is the negative of the volume,

$$f(\boldsymbol{x},\, a) = -\left(\pi d^2 h/4\right)\delta_{a,\,A} - (\pi dh/4)\delta_{a,\,A-1} \tag{41}$$

The scalar design variables, h (height) and d (diameter), are contained. The goal is to minimize the objective function in Eq. (41) at the HF level by solving

$$\min_{\boldsymbol{x}\in\Omega_1} -\left(\pi d^2 h/4\right) \tag{42}$$

$$\text{s.t.} 0 \leq h,\, d \leq 1 \tag{43}$$

$$2d^2 + 4dh = 1 \tag{44}$$

In the borehole example (Tran et al. 2020b), the objective function is the flow rate of water through the borehole,

$$f(\boldsymbol{x},\, a) = \frac{\left(2\pi\delta_{a,\,A} + 5\delta_{a,\,A-1} + 10\delta_{a,\,A-2}\right)x_3(x_4 - x_6)}{ln\left(\frac{x_2}{x_1}\right)\left(\delta_{a,\,A} + 1.5\delta_{a,\,A-1} + 0.1\delta_{a,\,A-2} + \frac{2x_3x_7}{ln\left(\frac{x_2}{x_1}\right)x_1^2x_8} + \frac{x_3}{x_5}\right)} \tag{45}$$

The eight scalar design variables, x_1, ..., x_8, are contained. The goal is to minimize the objective function in Eq. (45) at the HF level by solving.

$$\min_{\boldsymbol{x}\in\Omega_1} \frac{2\pi x_3(x_4 - x_6)}{ln\left(\frac{x_2}{x_1}\right)\left(1 + \frac{2x_3x_7}{ln\left(\frac{x_2}{x_1}\right)x_1^2x_8} + \frac{x_3}{x_5}\right)} \tag{46}$$

$$\text{s.t.} 0.05 \leq x_1 \leq 0.15 \tag{47}$$

$$100 \leq x_2 \leq 50,\, 000 \tag{48}$$

$$63,\, 070 \leq x_3 \leq 115,\, 600 \tag{49}$$

$$990 \leq x_4 \leq 1110 \tag{50}$$

$$63.1 \leq x_5 \leq 116 \tag{51}$$

$$700 \leq x_6 \leq 820 \tag{52}$$

$$1120 \leq x_7 \leq 1,\ 680 \tag{53}$$

$$9855 \leq x_8 \leq 12,\ 045 \tag{54}$$

In the welded beam example (Tran et al. 2020b), the multi-fidelity objective function is the monetary cost of the two-sided weld,

$$\begin{aligned} f(x,\ a) = {} & 10^{-4}h^2 l\left(1047\delta_{a,\,A} + 489\delta_{a,\,A-1} + 5235\delta_{a,\,A-2} + 5584\delta_{a,\,A-3}\right) \\ & + \left(\left(481\delta_{a,\,A} + 224\delta_{a,\,A-1} + 2405\delta_{a,\,A-2} + 2566\delta_{a,\,A-3}\right)(14 + l)\,tb\right) + h^2 l \end{aligned} \tag{55}$$

The scalar design variables are the spatial dimensions, h, l, b, and t. Four design constraints,

$$\frac{0.577hl\left(15\delta_{a,\,A} + 4\left(1 - \delta_{a,\,A}\right)\right)\left((h+t)^2 + l^2\right)}{\sqrt{\frac{\left((h+t)^2+l^2\right)^2}{72} + 2(14 + 0.5l)^2 + \frac{\left((h+t)^2+l^2\right)(14+0.5l)l}{3\sqrt{(h+t)^2+l^2}}}} \geq 18\ \text{in}^2 \tag{56}$$

$$\begin{aligned} & \frac{\left(\frac{1.008}{4.013tb^3} + 5t\left(3\delta_{a,\,A} + 1.4\delta_{a,\,A-1} + \delta_{a,\,A-2} + 1.6\delta_{a,\,A-3}\right)\right)^2}{\left(3\delta_{a,\,A} + 1.4\delta_{a,\,A-1} + \delta_{a,\,A-2} + 1.6\delta_{a,\,A-3}\right)\left(6\delta_{a,\,A} + 3\left(\delta_{a,\,A-1} + \delta_{a,\,A-3}\right) + 2\delta_{a,\,A-2}\right)} \\ & \quad \leq 31360\ \text{in}^2 \end{aligned} \tag{57}$$

$$b \leq \frac{15\delta_{a,\,A} + 4\left(1 - \delta_{a,\,A}\right)}{252t^2} \tag{58}$$

and

$$t^3 b\left(3\delta_{a,\,A} + 1.4\delta_{a,\,A-1} + \delta_{a,\,A-2} + 1.6\delta_{a,\,A-3}\right) \geq 26.3424\text{in}^4 \tag{59}$$

are defined at different levels of fidelity. The goal is to minimize the objective function in Eq. (54) at the HF level by solving

$$\min_{\boldsymbol{x} \in \Omega_1} 0.1047h^2 l \tag{60}$$

$$\text{s.t.} \frac{2.885hl\left((h+t)^2 + l^2\right)}{\sqrt{\frac{\left((h+t)^2+l^2\right)^2}{72} + 2(14 + 0.5l)^2 + \frac{\left((h+t)^2+l^2\right)(14+0.5l)l}{3\sqrt{(h+t)^2+l^2}}}} \geq 6\ \text{in}^2 \tag{61}$$

$$\left(\frac{1.008}{4.013tb^3} + 15t\right)^2 \leq 564480\text{in}^2 \tag{62}$$

$$252bt^2 \leq 15 \tag{63}$$

$$3t^3b \geq 26.3424\text{in}^4 \tag{64}$$

$$0.0625 \leq h \leq 2 \tag{65}$$

$$0.1 \leq l \leq 10 \tag{66}$$

$$2 \leq t \leq 20 \tag{67}$$

$$0.0625 \leq b \leq 2 \tag{68}$$

$$h \leq b \tag{69}$$

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11081-026-10078-5.

Acknowledgements Not applicable

Author contributions Conceptualization: I.A., M.K., J.S., Y. W.; Methodology: I.A., M.K., Y. W.; Formal analysis and investigation: I.A., M.K., Y. W.; Writing—original draft preparation: I.A., M.K., Y. W.; Writing—review and editing: I.A., M.K., J.S., E.Z., H.Z., Y. W.; Funding acquisition: H.Z., Y. W.; Resources: I.A., E.Z., H.Z., Y. W.; Supervision: Y. W.

Funding This work was supported in part by the National Science Foundation under grants CMMI-1663227 and IGERT-1258425, by a Georgia Tech-Oak Ridge National Laboratory (GT-ORNL) Collaboration seed grant. IA was supported by a Fullbright Scholarship. EZ was supported by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U. S. Department of Energy.

Data availability The data and code that support the findings of this study are available upon reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Awad R (2021) Sizing optimization of truss structures using the political optimizer (PO) algorithm. Structures 33:4871–4894. https://doi.org/10.1016/j.istruc.2021.07.027

Aydogdu I, Wang Y (2024) Scalable bayesian optimization based on exploitation-enhanced sparse gaussian process. Struct Multidiscip Optim 67(12):215

Aydogdu I (2025a) Genetic algorithm for truss optimization problem. The mathworks, Inc. https://www.mathworks.com/matlabcentral/fileexchange/182262-genetic-algorithm-for-truss-optimization-problem.

Aydogdu I (2025b) Quasi-Newton method for truss optimization problem. The MathWorks, Inc. https://www.mathworks.com/matlabcentral/fileexchange/182261-quasi-newton-method-for-truss-optimization-problem.

Belakaria S, Deshwal A, Doppa JR (2020)mMulti-fidelity multi-objective Bayesian optimization: an output space entropy search approach. In: Proceedings of the AAAI Conference on artificial intelligence. 34 p 10035–10043

Belegundu AD, Chandrupatla TR (2019) Optimization concepts and applications in engineering. Cambridge University Press

Chen SS, Jiang Z, Yang SX, Chen W (2017) Multimodel fusion based sequential optimization. AIAA J 55(1):241–254. https://doi.org/10.2514/1.J054729

Farshi B, Alinia-ziazi A (2010) Sizing optimization of truss structures by method of centers and force formulation. Int J Solids Struct 47(18–19):2508–2524. https://doi.org/10.1016/j.ijsolstr.2010.05.009

Forrester AI, Sóbester A, Keane AJ (2007) Multi-fidelity optimization via surrogate modelling. Proc R Soc Lond A Math Phys Eng Sci 463(2088):3251–3269

Foumani ZZ, Shishehbor M, Yousefpour A, Bostanabad R (2023) Multi-fidelity cost-aware Bayesian optimization. Comput Methods Appl Mech Eng. https://doi.org/10.1016/j.cma.2023.115937

Frazier P, Powell W, Dayanik S (2009) The knowledge-gradient policy for correlated normal beliefs. INFORMS J Comput 21(4):599–613. https://doi.org/10.1287/ijoc.1080.0314

Frey D, Shin JH, Musco C, Modestino MA (2022) Chemically-informed data-driven optimization (ChIDDO): leveraging physical models and bayesian learning to accelerate chemical research. React Chem Eng 7(4):855–865. https://doi.org/10.1039/d2re00005a

Gano SE, Renaud JE, Sanders B (2005) Hybrid variable fidelity optimization by using a kriging-based scaling function. AIAA J 43(11):2422–2430. https://doi.org/10.2514/1.12466

Gano S, Sanders B, Renaud J (2004) Variable fidelity optimization using a kriging based scaling function. In: 10th AIAA/ISSMO Multidisciplinary analysis and optimization conference. p 4460

Han Z-H, Zimmermann R, Goretz S (2010) A new cokriging method for variable-fidelity surrogate modeling of aerodynamic data. In: 48th AIAA Aerospace sciences meeting including the new horizons forum and Aerospace exposition, p 1225

Han Z-H, Görtz S (2012) Hierarchical kriging model for variable-fidelity surrogate modeling. AIAA J 50(9):1885–1896

Hoffman MW, Ghahramani Z (2015) Output-space predictive entropy search for flexible global optimization. In: NIPS workshop on Bayesian Optimization, p 1–5

Homaifar A, Qi CX, Lai SH (1994) Constrained optimization via genetic algorithms. SIMULATION 62(4):242–253. https://doi.org/10.1177/003754979406200405

Honarmandi P, Attari V, Arroyave R (2022) Accelerated materials design using batch Bayesian optimization: a case study for solving the inverse problem from materials microstructure to process. Comp Mater Sci. https://doi.org/10.1016/j.commatsci.2022.111417

Imani M, Ghoreishi SF, Allaire D, Braga-Neto UM, Aaai (2019) MFBO-SSM: Multi-fidelity Bayesian optimization for fast inference in state-space models. In: 33rd AAAI Conference on artificial intelligence / 31st innovative applications of artificial intelligence conference / 9th AAAI symposium on educational advances in artificial intelligence, Honolulu, AAAI Conference on Artificial Intelligence. p 7858–7865

Imani M, Ghoreishi SF (2022) Scalable inverse reinforcement learning through multifidelity bayesian optimization. IEEE Trans Neural Netw Learn Syst 33(8):4125–4132. https://doi.org/10.1109/tnnls.2021.3051012

Kandasamy K, Dasarathy G, Oliva J, Schneider J, Póczos B (2019) Multi-fidelity gaussian process bandit optimisation. J Artif Intell Res 66:151–196. https://doi.org/10.1613/jair.1.11288

Kandasamy K, Dasarathy G, Oliva J, Schneider J, Póczos B (2016) Gaussian process Bandit optimisation with multi-fidelity evaluations. In: 30th Conference on neural information processing systems (NIPS), Barcelona, SPAIN, 2016. Advances in neural information processing systems, vol 29.

Kandasamy K, Dasarathy G, Schneider J, Póczos B (2017) Multi-fidelity Bayesian optimisation with continuous approximations. In: 34th International Conference on machine learning, Sydney, Australia. Proceedings of machine learning research, vol 70

Kennedy MC, O'Hagan A (2000) Predicting the output from a complex computer code when fast approximations are available. Biometrika 87(1):1–13

Khatouri H, Benamara T, Breitkopf P, Demange J, Feliot P (2020) Constrained multi-fidelity surrogate framework using bayesian optimization with non-intrusive reduced-order basis. Adv Model Simul Eng Sci 7:1–20

Korondi PZ, Marchi M, Parussini L, Poloni C (2021) Multi-fidelity design optimisation strategy under uncertainty with limited computational budget. Optim Eng 22(2):1039–1064. https://doi.org/10.1007/s11081-020-09510-1

Kraft D (1988) A software package for sequential quadratic programming. Forschungsbericht- Deutsche Forschungs- und Versuchsanstalt fur Luft- und Raumfahrt

Le Gratiet L, Garnier J (2014) Recursive co-kriging model for design of computer experiments with multiple levels of fidelity. Int J Uncertain Quantif 4(5):365–386. https://doi.org/10.1615/Int.J.UncertaintyQuantification.2014006914

Li S, Xing W, Kirby R, Zhe S (2020) Multi-fidelity bayesian optimization via deep neural networks. Adv Neural Inf Process Syst 33:8521–8531

Li S, Kirby R, Zhe S (2021) Batch multi-fidelity bayesian optimization with deep auto-regressive networks. Adv Neural Inf Process Syst 34:25463–25475

Nachar S, Boucard PA, Néron D, Rey C (2020) Multi-fidelity bayesian optimization using model-order reduction for viscoplastic structures. Finite Elem Anal and des. https://doi.org/10.1016/j.finel.2020.103400

Parussini L, Venturi D, Perdikaris P, Karniadakis GE (2017) Multi-fidelity gaussian process regression for prediction of random fields. J Comput Phys 336:36–50. https://doi.org/10.1016/j.jcp.2017.01.047

Perdikaris P, Karniadakis GE (2016) Model inversion via multi-fidelity Bayesian optimization: a new paradigm for parameter estimation in haemodynamics, and beyond. J R Soc Interface. https://doi.org/10.1098/rsif.2015.1107

Poloczek M, Wang JL, Frazier PI (2017) Multi-information source optimization. In: 31st Annual Conference on neural information processing systems (NIPS), Long Beach, Advances in neural information processing systems, vol 30.

Rezende RS, Hadziefendic M, Hansen J, Schuhmann R, Ieee (2021) Multi-output variable-fidelity bayesian optimization of a common mode Choke. In: IEEE International joint symposium on electromagnetic compatibility (EMC), Signal and Power Integrity (SIPI) and EMC Europe, Electr Network, Jul 26-Aug 20 2021. IEEE International Symposium on Electromagnetic Compatibility. p 1093–1098

Schmit LA, Farshi B (1974) Some approximation concepts for structural synthesis. AIAA J 12(5):692–699. https://doi.org/10.2514/3.49321

Shahriari B, Swersky K, Wang ZY, Adams RP, de Freitas N (2016) Taking the human out of the loop: a review of Bayesian optimization. Proc IEEE 104(1):148–175. https://doi.org/10.1109/jproc.2015.2494218

Shu LS, Jiang P, Wang Y (2021) A multi-fidelity bayesian optimization approach based on the expected further improvement. Struct Multidiscip Optim 63(4):1709–1719. https://doi.org/10.1007/s00158-020-02772-4

Song JL, Chen YX, Yue YS (2019) A general framework for multi-fidelity bayesian optimization with gaussian processes. In: 22nd International Conference on artificial intelligence and statistics (AISTATS), Naha, Proceedings of Machine Learning Research, vol 89.

Sorourifar F, Choksi N, Paulson JA (2023) Computationally efficient integrated design and predictive control of flexible energy systems using multi-fidelity simulation-based bayesian optimization. Optim Control Appl Methods 44(2):549–576. https://doi.org/10.1002/oca.2817

Takeno S, Fukuoka H, Tsukada Y, et al. (2020) Multi-fidelity Bayesian optimization with max-value entropy search and its parallelization. In: International Conference on Machine Learning, p 9334–9345

The MathWorks I (2025a) Global Optimization Toolbox. 9.3 edn, Natick, MA.

The MathWorks I (2025b) Optimization Toolbox. 9.5 edn, Natick, MA.

Tran A, Tranchida J, Wildey T, Thompson AP (2020a) Multi-fidelity machine-learning with uncertainty quantification and Bayesian optimization for materials design: application to ternary random alloys. J Chem Phys. https://doi.org/10.1063/5.0015672

Tran A, Wildey T, McCann S (2020b) sMF-BO-2CoGP: a sequential multi-fidelity constrained Bayesian optimization framework for design applications. J Comput Inf Sci Eng. https://doi.org/10.1115/1.4046697

Tran A, Wildey T, McCann S, Asme (2020) sBF-BO-2CoGP: A sequential bi-fidelity constrained Bayesian optimization for design applications. In: American-Society-of-Mechanical-Engineers(ASME) International Design Engineering Technical Conferences / 39th American-Society-of-Mechanical-Engineers(ASME) Computers and Information in Engineering Conference (IDETC-CIE), Anaheim, CA

Venkayya V (1971) Design of optimum structures. Comput Struct 1(1–2):265–309

Virtanen P, Gommers R, Oliphant TE et al (2020) Scipy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods 17(3):261–272. https://doi.org/10.1038/s41592-019-0686-2

Wang Y (2021) Design of trustworthy cyber-physical-social systems with discrete Bayesian optimization. J Mech des. https://doi.org/10.1115/1.4049532

Wang Z, Jegelka S (2017) Max-value entropy search for efficient Bayesian optimization. In: 34th International Conference on machine learning, Sydney. Proceedings of machine learning research, vol 70.

Wu J, Toscano-Palmerin S, Frazier PI, Wilson AG (2020) Practical multi-fidelity bayesian optimization for hyperparameter tuning. In: 35th Uncertainty in Artificial Intelligence (UAI) Conference, Tel Aviv, ISRAEL. Proceedings of Machine Learning Research, vol 115. p 788–798

Zhang SH, Lyu WL, Yang F, et al. (2019) An efficient multi-fidelity bayesian optimization approach for analog circuit synthesis. In: 56th ACM/EDAC/IEEE Design Automation Conference (DAC), Las Vegas.

Zimmermann R, Han Z (2010) Simplified cross-correlation estimation for multi-fidelity surrogate cokriging models. Adv Appl Math Sci 7:181–202

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.