Physics-aware multifidelity Bayesian optimization: A generalized formulation

Computers and Structures 296 (2024) 107302

Contents lists available at ScienceDirect

Computers and Structures

journal homepage: www.elsevier.com/locate/compstruc

Physics-aware multifidelity Bayesian optimization: A generalized formulation

Francesco Di Fiore [a,*], Laura Mainini [b,a]

[a] Department of Mechanical and Aerospace Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, Turin, 10129, Italy
[b] Department of Aeronautics, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom

ARTICLE INFO

Keywords:
Domain-aware efficient optimization
Multifidelity Bayesian optimization
Physics-informed machine learning
Structural health monitoring
Simulation-driven design

ABSTRACT

The adoption of high-fidelity models for many-query optimization problems is majorly limited by the significant computational cost required for their evaluation at every query. Multifidelity Bayesian methods (MFBO) allow to include costly high-fidelity responses for a sub-selection of queries only, and use fast lower-fidelity models to accelerate the optimization process. State-of-the-art methods rely on a purely data-driven search and do not include explicit information about the physical context. This paper acknowledges that prior knowledge about the physical domains of engineering problems can be leveraged to accelerate these data-driven searches, and proposes a generalized formulation for MFBO to embed a form of domain awareness during the optimization procedure. In particular, we formalize a bias as a multifidelity acquisition function that captures the physical structure of the domain. This permits to partially alleviate the data-driven search from learning the domain properties on-the-fly, and sensitively enhances the management of multiple sources of information. The method allows to efficiently include high-fidelity simulations to guide the optimization search while containing the overall computational expense. Our physics-aware multifidelity Bayesian optimization is presented and illustrated for two classes of optimization problems frequently met in science and engineering, namely design optimization and health monitoring problems.

1. Introduction

Optimization problems are ubiquitous in science and engineering applications [1]. Those also include the support to engineering tasks that are in increasing demand to meet sustainability goals such as the identification of the best design configurations to maximize the performance and minimize the environmental impact of novel engineering solutions, and the detection and identification of damages or faults to monitor the health condition of complex systems to maximize their useful life and minimize waste of resources.

Over the last decades, the increase of computing power and the advances in computational modeling capabilities made available computer-based models for the accurate analysis and simulation of complex physical systems. This is the case of computational schemes for the numerical solution of governing partial differential equations as computational fluid dynamic solvers to represent viscous fluids, and finite element methods for the analysis of mechanical structures, heath transfer and electromagnetic phenomena. In principle, this computer-based representations can provide a remarkable contribution to enhance the search and identification task in simulation-based optimization. Unfortunately, the extensive adoption of these high-fidelity models during the optimization procedure is hampered by the significant computational cost and time required for their evaluation, potentially in the order of months for a single evaluation on high performance computing platforms. This issue becomes more challenging for many-query optimization problems where the demand for model evaluations grows exponentially with the number of parameters to optimize.

The use of low-fidelity models constitutes a popular approach to reduce the computational resources associated with the solution of optimization problems. Low-fidelity representations introduce assumptions about the physics and/or approximate the solution of the governing equations, and relief the computational expenditure for the evaluation of the response of the system. On one hand, the use of low-fidelity models allows to efficiently acquire and exploit a massive amount of information; on the other hand, their adoption results in a reduction of the overall time required for optimization. However, the main drawback of this strategy is enclosed in the simplified modeling approach which might not be adequate to capture complex physical phenomena

* Corresponding author.
E-mail address: francesco.difiore@polito.it (F. Di Fiore).

https://doi.org/10.1016/j.compstruc.2024.107302
Received 31 March 2023; Accepted 29 January 2024
Available online 7 February 2024

that characterize many advanced technological applications, providing inaccurate responses that potentially lead to the suboptimal identification of solutions.

Multifidelity methods offer the opportunity to efficiently include high-fidelity models during the optimization procedure, combining information elicited from a library of models hierarchically ordered accordingly to the accuracy of the response and associated expense of the computations [2–5]. These methodologies use fast low-fidelity models to speedup the search procedure, and refine the solution identified through of a principled selection of high-fidelity information. The goal is to accelerate the identification of optimal solutions while managing a trade-off between accuracy and computational cost. Multifidelity frameworks have been applied to solve a variety of optimization problems. Examples include and are not limited to the design optimization of aircraft [6], vessels [7], and hybrid vehicle [8], and health assessment of civil structures [9], composite wings [10] and industrial systems [9].

Multifidelity Bayesian Optimization (MFBO) represents a class of computational techniques that relies on a Bayesian framework to address optimization problems, and combines data from sources of information at different levels of fidelity [11–13]. The Bayesian methodology defines a scheme for the optimization of black-box functions where a probabilistic surrogate model is updated at each iteration using evaluations of models prescribed by an acquisition function [14–16]. The surrogate model provides an approximation of the objective function over the domain while the acquisition function determines a policy to measure the utility of evaluating the objective function in certain locations of the domain. Whether a library of models is available, multifidelity Bayesian optimization synthesizes data elicited through multiple sources of information into a unique surrogate model, and guides the search through an adaptive sampling scheme based on a multifidelity acquisition function that selects the most promising location and the associated level of fidelity to query.

MFBO has been widely adopted to solve optimization problems in science and engineering. Meliani et al. [17] propose a multifidelity Bayesian technique for high-dimensional design optimization problems, and demonstrate the methodology for the aerodynamic shape optimization of an airfoil outperforming a standard single-fidelity Bayesian algorithm. Tran et al. [18] develop a multifidelity Bayesian framework for the inference of the optimal chemical composition in material science, dedicating particular attention to the identification of the optimal concentration of components achieving the desired bulk modulus for ternary random alloys. Serani et al. [19] propose a multifidelity approach based on stochastic radial basis functions to include expensive computational fluid dynamic simulations in global design optimization, and demonstrate the methodology for the optimization of a NACA hydrofoil and a destroyer-type vessel for naval engineering applications. Perdikaris and Karniadakis [20] adopt a multifidelity Bayesian strategy to solve inverse problems in haemodynamics through the identification of system parameters that characterize physiologically correct blood flow simulations.

All these multifidelity frameworks rely on fully data-driven adaptive samplings informed exclusively through probabilistic data extracted from the surrogate model. Accordingly, the search procedure has to learn entirely from data both the surrogate model and the characterization of the discrepancies – frequently non-linear – between the different models over the entire physical domain. This can still require a large amount of high-fidelity information to capture the modeling correlations for each level of fidelity available, and results in intensive computations associated with the massive evaluations of accurate numerical models. In the engineering context, prior knowledge about these discrepancies is at disposal, either because it is formalized by the governing equations that represent the physics of the system or because it derives from the know-how of experts about the distinguishing physical phenomena characterizing the system behavior. In these optimization scenarios, the introduction of the physics-awareness during the search procedure could lead to a principled and efficient use of high-fidelity data according to the prior knowledge about the physics of the system.

This paper aims at capturing this opportunity and proposes an original physics-aware multifidelity Bayesian optimization that incorporates prior domain knowledge to further improve and accelerate the optimization search in multifidelity settings. This permits to partially alleviate the data-driven search from the characterization of the domain structure while trying to identify the optimal combination of parameters that minimizes the objective function. In previous works, we introduced preliminary approaches to capture this prior/expert knowledge in the form of domain-awareness [21,22].

This work proposes a formal generalization of this sort of physics-aware reasoning for a broader relevance and direct applicability to multiple classes of optimization problems in science and engineering. The proposed framework adopts the multifidelity Gaussian process regression to model the belief about the objective function over the entire domain, which is progressively updated through an original acquisition function based on the multifidelity expected improvement. The multifidelity acquisition function is distinctively shaped to combine (i) data-driven information extracted from the surrogate model and (ii) prior/expert knowledge about the structure of the domain encapsulated during the search through a physics-aware utility function. This form of awareness allows to define an adaptive sampling scheme that efficiently manages different sources of information, targeting the balance between computational cost and accuracy demanded along the optimization search.

We demonstrate our physics-aware multifidelity framework for two popular families of optimization problems, namely the identification of the best combination of design parameters to maximize systems performance, and the non-destructive identification of systems damages or faults. In the first case, we consider the specific example of the aerodynamic design optimization of a transonic airfoil and the prior domain knowledge concerns the sensitivity of the fluid dynamic regime to the variation of the Mach number. The health monitoring problem demands for the assessment of the fault condition of a composite structure, and the prior domain knowledge relates to the effects that can be observed for different extensions of the damage. The two forms of scientific and expert domain knowledge are formulated as physics-aware utility functions that bias the query of numerical models depending on the structure of the domain.

This article is organized as follows: Section 2 describes the multifidelity Bayesian scheme and Section 3 presents our original physics-aware multifidelity Bayesian framework in detail. In Section 4 and Section 5, the proposed methodology is applied for an aerodynamic design problem and a structural health monitoring test case, respectively, and the formalization of our physics-aware technique together with the results are discussed. Finally, Section 6 provides concluding remarks.

2. Multifidelity Bayesian optimization: problem setup

The goal of optimization problems is to identify a combination of parameters $\boldsymbol{x}^*$ that minimizes an unknown objective function $f(\boldsymbol{x})$, and is mathematically described as:

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x}) \tag{1}$$

where $\mathcal{X} \in \mathbb{R}^d$ is the domain of the objective function. Frequently in these settings, the objective function is a black-box input/output relationship whose analytical form and derivatives are not explicitly available.

Bayesian optimization (BO) is a popular surrogate-based method that uses the Bayesian probability theory to address the optimization of black-box objective functions [14–16]. BO realizes a sequential derivative-free optimization procedure based on two core elements: a stochastic surrogate model that approximates the response of numerical models and emulates the distribution of the objective function over

Algorithm 1 Multifidelity Bayesian optimization.

Input: Design space $\mathcal{X} \in \mathbb{R}^d$, multifidelity objective function $f^{(l)}(\boldsymbol{x})$ and the prior distribution $P(f^{(l)})$
Output: $\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in\mathcal{X}} f(\boldsymbol{x})$
$\mathcal{D}_1 \leftarrow \{\boldsymbol{x}_n, f^{(l_n)}, l_n\}_{n=1}^{N_1}$ collect initial observations
$i \leftarrow 1$
repeat
  Update the posterior distribution $P(f^{(l)}|\mathcal{D}_i)$ using $\mathcal{D}_i$
  Maximize the acquisition function $\{\boldsymbol{x}_{i+1}, l_{i+1}\} = \max U(\boldsymbol{x}, l)$
  Evaluate $f^{(l_{i+1})}(\boldsymbol{x}_{i+1})$
  Augment the dataset $\mathcal{D}_{i+1} = \mathcal{D}_i \cup \{\boldsymbol{x}_{i+1}, f^{(l_{i+1})}(\boldsymbol{x}_{i+1}), l_{i+1}\}$
  $i + 1 \leftarrow i$
until $B_i \leq B_{max}$
return $\boldsymbol{x}^*$ that minimize $f(\boldsymbol{x})$ over $\mathcal{D}_i$

the domain, and an acquisition function that determines the optimal sequence of future samples to be evaluated. The overall computational efficiency of BO can be improved combining data from multiple sources of information at different levels of fidelity $\{f^{(1)}, ..., f^{(l)}, ..., f^{(L)}\}$: usually in this set the higher the level of fidelity, the more accurate and yet expensive its evaluation. In multifidelity Bayesian optimization, both the stochastic emulator and the acquisition function must accommodate the multiple sources available to compute (approximated) observations of the objective function. In particular, the surrogate model approximates the objective function by synthesizing the information from multiple fidelities into a unique emulator. Commonly, the predictive framework for MFBO is based on the extension of Gaussian processes surrogate models to multiple levels of fidelity through an autoregressive scheme [23–25].

The multifidelity acquisition function defines an adaptive sampling scheme that permits to sequentially decide the best location of the domain and associated source of information to query. This decision making procedure is typically driven entirely by the statistical data in output from the surrogate model. The multitude of multifidelity acquisition functions proposed in literature are often developed leveraging existing acquisition functions adopted in single-fidelity Bayesian optimization, such as the Expected Improvement (EI) [26], Probability of Improvement (PI) [27] and Entropy Search (ES) [28]. Examples of popular multifidelity acquisition functions include the Multifidelity Expected Improvement (MFEI) [29], Variable-Fidelity Probability of Improvement (VFPI) [30], Multifidelity Predictive Entropy Search (MF-PES) [31], and Multifidelity Max-Value Entropy Search (MFMES) [32]. The primary difference between these formulations is that the decision making process is realized through different approaches to measure the reward of unknown samples. The goal is to achieve the trade-off between the exploration of the domain $\mathcal{X}$ in locations where the uncertainty associated with the surrogate model is higher, and the exploitation in regions where the prediction of the emulator indicates that the optimum is likely to be located.

Algorithm 1 illustrates the pseudo-code of the multifidelity Bayesian optimization methodology. MFBO consists of a sequential approach to solve the optimization problem in Equation (1), starting with an initial dataset $\mathcal{D}_1 = \{\boldsymbol{x}_n, f^{(l_n)}(\boldsymbol{x}_n), l_n\}_{n=1}^{N_1}$ of samples and associated values of the objective function evaluated with the l_n level of fidelity. At each iteration i of the optimization procedure, the collected dataset is used to build the multifidelity surrogate model combining the prior distribution $P(f^{(l)})$ with the likelihood function $P(\mathcal{D}_i|f^{(l)})$ to compute the posterior distribution $P(f^{(l)}|\mathcal{D}_i) \propto P(\mathcal{D}_i|f^{(l)})P(f^{(l)})$, which represents the updated emulator of the objective function for each level of fidelity. Then, MFBO induces a multifidelity acquisition function $U(\boldsymbol{x}, l)$ based on the posterior that measures the utility of probing domain locations with certain levels of fidelity: the next design $\boldsymbol{x}_{i+1}$ and level of fidelity l_{i+1} to query are selected by maximizing the acquisition function $\{\boldsymbol{x}_{i+1}, l_{i+1}\} = \arg\max U(\boldsymbol{x}, l)$ over the domain $\mathcal{X}$. After the new observation is collected, the dataset $\mathcal{D}_{i+1} = \mathcal{D}_i \cup \{\boldsymbol{x}_{i+1}, f^{(l_{i+1})}(\boldsymbol{x}_{i+1}), l_{i+1}\}$ is updated and the procedure iterates until a maximum computational budget $B_i = B_{max}$ is reached, where B_i is the cumulative computational cost expended until iteration i.

3. Physics-aware multifidelity Bayesian optimization

The Physics-Aware Multifidelity Bayesian Optimization (PA-MFBO) framework permits to accelerate the optimization procedure alleviating the burden of learning the characterization of the domain structure entirely from data. In particular, PA-MFBO embeds the available prior scientific and expert knowledge about the structure of the physical domain during the search procedure. This formalizes a sort of physics-aware reasoning based on an original formulation of the multifidelity acquisition function (Section 3.2) that biases the query of numerical models according to the specific structures of the domain, while still informed with data from the multifidelity emulator (Section 3.1).

3.1. Multifidelity Gaussian process

Our PA-MFBO framework employs the multifidelity Gaussian process as the emulator of the objective function, which is formalized extending the Gaussian process to multiple levels of fidelity through an autoregressive scheme.

The Gaussian process (GP) is a non-parametric kernel-based statistical model that permits to effectively approximate the black-box relationship between locations of the domain $\boldsymbol{x}$ and the associated values of the objective $f(\boldsymbol{x})$ [23,24]. Accordingly, the Gaussian process regression builds an emulator of the objective function using the knowledge of observations of $f(\boldsymbol{x})$ collected in specific locations, and defines a distribution over functions completely specified by the mean function $\mu(\boldsymbol{x}) : \mathcal{X} \to \mathbb{R}$ and covariance function $\kappa(\boldsymbol{x}, \boldsymbol{x}') : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$. The mean function $\mu(\boldsymbol{x}) = \mathbb{E}[f(\boldsymbol{x})]$ reflects the expected value of the objective at a certain location $\boldsymbol{x}$ and the covariance or kernel function $\kappa(\boldsymbol{x}, \boldsymbol{x}') = \mathbb{E}\left[(f(\boldsymbol{x}) - \mu(\boldsymbol{x}))(f(\boldsymbol{x}') - \mu(\boldsymbol{x}'))\right]$ represents the dependence between the values of the objective function at different locations $\boldsymbol{x}$ and $\boldsymbol{x}'$. This constitutes a predictive framework where the predictor in the form of the mean function approximates the objective function over unknown regions of the domain, and the covariance function quantifies the uncertainty associated with this prediction.

The availability of multiple representations of the objective function $\{f^{(l)}\}_{l=1}^{l=L}$ demands for an emulator capable to approximate the distribution of the objective function and synthesize observations from different sources of information. In this scenario, the Gaussian process regression can be extended to combine the models of the objective function at different levels of fidelity into a single predictive framework. Let us assume we have collected paired input/output observations of the objective function in the form $\mathcal{D}_N = \{\boldsymbol{x}_n, y^{(l_n)}(\boldsymbol{x}_n), l_n\}_{n=1}^{N}$, where the output $\boldsymbol{y} = \{y^{(l_n)}(\boldsymbol{x}_n)\}_{n=1}^{N}$ is normally distributed given $\boldsymbol{f} = \{f_n^{(l_n)}\}_{n=1}^{N}$:

$$\boldsymbol{y} \,|\, \boldsymbol{f}, \sigma_\epsilon^2 \sim \mathcal{N}(\boldsymbol{f}, \sigma_\epsilon^2 \boldsymbol{I}) \tag{2}$$

where the measurement noise is characterized by the same variance σ_ϵ^2 over the available levels of fidelity.

Following the Bayesian inference principle, the prior belief about the objective function $P(f^{(l)})$ is combined with the likelihood function $P(\mathcal{D}_N \mid f^{(l)})$ to compute the posterior distribution $P(f^{(l)} \mid \mathcal{D}_N) \propto P(\mathcal{D}_N \mid f^{(l)})P(f^{(l)})$. This represents the updated probabilistic model of the objective function at a certain level of fidelity l. In the black-box setting, the multifidelity Gaussian process regression considers no prior information about the multiple representations of the objective function. Hence, the lowest-fidelity prior $f^{(1)} \sim GP(0, \kappa_1(\boldsymbol{x}, \boldsymbol{x}'))$ is represented as a Gaussian process with zero mean function $\mu^{(l)} = 0$ and covariance function $\kappa_1(\boldsymbol{x}, \boldsymbol{x}')$ while the higher levels of fidelity are derived recursively through an autoregressive scheme [25]:

$$f^{(l)} = \varrho^{(l-1)} f^{(l-1)}(\boldsymbol{x}) + \delta^{(l)}(\boldsymbol{x}) \quad l = 2, ..., L \tag{3}$$

where $\varrho^{(l-1)}$ is a constant scaling factor that tunes successive levels of fidelity $f^{(l)}$ and $f^{(l-1)}$, and $\delta^{(l)}$ is the discrepancy between two adjoining levels of fidelity modeled as a Gaussian process with mean function $\upsilon(\boldsymbol{x})^T\beta^{(l)}$ and covariance function $\kappa^{(l)}\left(\boldsymbol{x},\boldsymbol{x}'\right)$ where υ is the vector of regression functions and $\beta^{(l)}$ are the regression coefficients. In this work, we adopt the Gaussian correlation model as the covariance function:

$$\kappa(\boldsymbol{x},\boldsymbol{x}') = \varsigma_l^2 \exp\{-\sum_{m=1}^{M} \varpi_l^m (\boldsymbol{x}_m - \boldsymbol{x}'_m)^2\} \tag{4}$$

where $\varpi = (\varpi_l^1, \varpi_l^2, ..., \varpi_l^M)$ is the roughness parameter, and ς_l^2 is the process variance of the l-th level of fidelity.

Accordingly, the posterior distribution of the multifidelity Gaussian process is completely specified by the mean $\mu^{(l)}$ and variance $\sigma^{2(l)}$:

$$\mu^{(l)}(\boldsymbol{x}) = \kappa_N^{(l)}(\boldsymbol{x})^T \left(\boldsymbol{K} + \sigma_\epsilon \boldsymbol{I}\right)^{-1} \boldsymbol{y} \tag{5}$$

$$\sigma^{2(l)}(\boldsymbol{x}) = \kappa\left((\boldsymbol{x},l),(\boldsymbol{x},l)\right) - \kappa_N^{(l)}(\boldsymbol{x})^T \left(\boldsymbol{K} + \sigma_\epsilon \boldsymbol{I}\right)^{-1} \kappa_N^{(l)}(\boldsymbol{x}) \tag{6}$$

where $\kappa_N^{(l)}$ is defined as $\kappa_N(\boldsymbol{x}) \doteq (\kappa((\boldsymbol{x},l),(\boldsymbol{x}_1,l_1)), \cdots, \kappa((\boldsymbol{x},l),(\boldsymbol{x}_N,l_N)))$, and $\boldsymbol{K}$ is the kernel matrix formalized as follows:

$$\boldsymbol{K} = \begin{pmatrix} \kappa^{(l-1)}(\boldsymbol{x},\boldsymbol{x}')\boldsymbol{K}^{(l-1)} & \varrho\kappa^{(l-1)}(\boldsymbol{x},\boldsymbol{x}')\boldsymbol{K}^{(l-1)} \\ \varrho\kappa^{(l-1)}(\boldsymbol{x},\boldsymbol{x}')\boldsymbol{K}^{(l-1)} & \varrho^2\kappa^{(l-1)}(\boldsymbol{x},\boldsymbol{x}')\boldsymbol{K}^{(l-1)} + \kappa^{(l)}(\boldsymbol{x},\boldsymbol{x}')\boldsymbol{K}^{(l)} \end{pmatrix} \tag{7}$$

where $\boldsymbol{K}^{(l-1)}(i,j) = \kappa((\boldsymbol{x}_i,l-1),(\boldsymbol{x}_j,l-1))$ and $\boldsymbol{K}^{(l)}(i,j) = \kappa((\boldsymbol{x}_i,l),(\boldsymbol{x}_j,l))$.

The posterior mean function $\mu^{(l)}$ constitutes the prediction of the objective function at the l-th level of fidelity over the domain $\mathcal{X}$, and the posterior standard deviation $\sigma^{(l)}$ quantifies the associated level of uncertainty. The hyperparameters $(\varrho, \beta, \varpi, \varsigma)$ of the multifidelity Gaussian process surrogate model are estimated by maximizing the likelihood function [33].

3.2. Physics-aware multifidelity acquisition function

PA-MFBO incorporates the prior scientific and expert knowledge about the physical phenomena and the specific structure of the domain through the original physics-aware multifidelity acquisition function $U_{PA}(\boldsymbol{x},l)$ based on the multifidelity expected improvement [29]:

$$U_{PA}(\boldsymbol{x},l) = U_{EI}(\boldsymbol{x})\alpha_1(\boldsymbol{x},l)\alpha_2(\boldsymbol{x},l)\alpha_3(l)\alpha_4(\boldsymbol{\psi},l) \tag{8}$$

where $U_{EI}(\boldsymbol{x})$ is the expected improvement acquisition function evaluated at the highest level of fidelity [26]:

$$\begin{aligned} U_{EI}(\boldsymbol{x}) &= \mathbb{E}[\max(f^{(L)}(\hat{\boldsymbol{x}}^*) - f^{(L)}(\boldsymbol{x}), 0)] \\ &= \sigma(\boldsymbol{x})(I(\boldsymbol{x})\Phi(I(\boldsymbol{x}))) + \mathcal{N}(I(\boldsymbol{x}); 0, 1) \end{aligned} \tag{9}$$

where $I(\boldsymbol{x}) = (f^{(L)}(\hat{\boldsymbol{x}}^*) - \mu(\boldsymbol{x}))/\sigma(\boldsymbol{x})$ is the predicted improvement, $\hat{\boldsymbol{x}}^*$ is the current location of the best value of the objective sampled so far, $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution. The expectation term in Equation (9) relates only to the high-fidelity model of the objective function, and quantifies the expected gain potentially achieved adding an high-fidelity evaluation of the objective function.

The terms α_1, α_2 and α_3 are conceived to capture and balance the contributions of lower-fidelity evaluations of the objective function, and are formalized as follows:

$$\alpha_1(\boldsymbol{x},l) = corr\left[f^{(l)}(\boldsymbol{x}), f^{(L)}(\boldsymbol{x})\right] = \frac{\kappa((\boldsymbol{x},l),(\boldsymbol{x},L))}{\sqrt{\sigma^{2(l)}\sigma^{2(L)}}} \tag{10}$$

$$\alpha_2(\boldsymbol{x},l) = 1 - \frac{\sigma_\epsilon}{\sqrt{\sigma^{2(l)}(\boldsymbol{x}) + \sigma_\epsilon^2}} \tag{11}$$

$$\alpha_3(l) = \frac{\lambda^{(L)}}{\lambda^{(l)}}. \tag{12}$$

α_1 is defined as the posterior correlation coefficient between the l-th level of fidelity and the highest-fidelity available at the same location of the domain. This utility function reflects the reduction of the acquisition function when samples are evaluated with lower-fidelity models, and accounts for the decrease of the accuracy associated with a low-fidelity representation of the objective function. Accordingly, the use of a high-fidelity model is solicited when a low-fidelity estimate might produce unreliable observations of the objective function. α_2 considers the reduction of the uncertainty associated with the prediction of the multifidelity Gaussian process at the l-th level of fidelity after a new observation of the objective function with a certain level of fidelity l is added to the dataset $\mathcal{D}_N$. The objective of this term is to adjust the contribution of the high-fidelity expected gain quantified through Equation (9) considering the reduction of the optimization gain of additional evaluations at the l-th level of fidelity as the MFGP prediction becomes more accurate. Accordingly, this prevents the systematic sampling in already explored regions of the domain characterized by lower uncertainty. α_3 is formulated as the ratio between the computational cost $\lambda^{(L)}$ associated with the evaluation of the high-fidelity model and the computational cost $\lambda^{(l)}$ required to compute the l-th fidelity model. This utility function is conceived to include awareness about the computational resources required for the evaluation of the objective function adopting the l-th level of fidelity. The purpose of this term is to privilege the selection of lower-fidelity queries when similar improvements of the solution are obtained from higher-fidelity observations, and balance the computational cost and the informative contribution of different fidelity levels.

$\alpha_4(\boldsymbol{\psi},l)$ is the physics-aware utility function that embeds a source of prior knowledge represented by a set of physical variables $\boldsymbol{\psi}$ in the sampling scheme. This permits to introduce a learning bias that captures the scientific understanding and expertise underlying the physical domain of the system. Without claiming to limit the informative content that can be incorporated into our acquisition function, we identify two main sources of prior knowledge in the form of scientific and expert knowledge. Scientific knowledge refers to the body of rules formalized and validated through the scientific method such as conservation laws, physical principles or phenomenological behaviors that represent the physics of interest. Examples include the Navier-Stokes partial differential equations in fluid dynamics to model the motion of viscous fluids, and numerical methodologies to approximate the solution of the governing equations as the finite element method in structural mechanics. Expert knowledge represents the information that is held by a community of experienced specialists and validated implicitly over several years of experience in a specific field. Examples include the common knowledge within the engineering or physics community resulting from training, research and personal experience.

The proposed physics-aware multifidelity Bayesian optimization framework is illustrated and demonstrated for two cases of study, namely an aerodynamic design optimization problem (Section 4), and a structural health monitoring task (Section 5).

4. Aerodynamic design example

The design test case addresses the optimization of a transonic airfoil to improve the aerodynamic performance. Particular attention is dedicated to the cross-regime scenario where the fluid regime and the associated physical phenomena evolve during the optimization process. This defines a robust optimization procedure that potentially ensures to obtain optimal airfoil shapes for different operational conditions, without limiting the improvement of performance to a single application context [34,35]. In this case, the prior scientific knowledge about the structure of the domain relates to the transition of the fluid dynamic regimes during the optimization. Thus, the physics-aware utility function is formalized to bias the search procedure according to the evolution of the physical domain.

The Mach number M is the main physical variable that captures the evolution of the fluid domain, and constitutes a measure of the compressibility effects that modify the fluid structure. According to the fluid

mechanics theory [36,37], the flow-field around streamlined bodies is defined subsonic for values of the Mach number lower than 0.8, and represents a condition characterized by the absence of discontinuities and the fluid properties vary continuously. As the Mach number approaches the sonic condition, discontinuities in the form of local shock waves start to appear in the fluid domain with the consequent separation of the viscous boundary layer. This mixed subsonic-supersonic flow field emerges for values of the Mach number between 0.8 and 1.2, and is commonly referred as the transonic regime. The interactions between shock waves and boundary layer determine an increase of the drag force, and unsteady effects generated by a shift of the center of pressure of the aerodynamic body. Therefore, the cross-regime scenario poses significant challenges associated with the modeling of complex physics. On one hand, the subsonic regime can be represented adopting simplifications in the aerodynamic modeling due to the smooth evolution of the flow field; on the other hand, the transonic regime requires the implementation of accurate and robust modeling techniques to capture non-linear phenomena in the unsteady mixed subsonic-supersonic flow.

4.1. Optimization problem: cross-regime airfoil design

The aerodynamic design optimization problem consists in the identification of the optimal combination of design parameters that minimizes the drag coefficient C_d of a transonic airfoil, subject to a variety of aerodynamic and geometric constraints. For this demonstrative test-case, we adopt the RAE 2822 transonic airfoil that is modified through the code WG2AER developed by [38], where the original shape of the airfoil is linearly combined with weighted shape modification functions to obtain new geometries. The aerodynamic design optimization problem is formulated as follows:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \quad C_d(\boldsymbol{x}) \qquad \boldsymbol{x}=\left[w_1,\dots,w_6,M\right] \tag{13a}$$

$$s.t. \quad C_l = 0.824 \tag{13b}$$

$$-0.1 \le C_m \le -0.01 \tag{13c}$$

$$t/c = 0.1211 \tag{13d}$$

$$r \ge 0.007c \tag{13e}$$

$$\tau \ge 5^\circ \tag{13f}$$

$$t_{85}/c \ge 0.02 \tag{13g}$$

$$\mathcal{X} = I_w \times I_M \tag{13h}$$

where the design parameters $\boldsymbol{x}=\left[w_1,\dots,w_6,M\right]$ consist of six weights w_i assigned to the shape modification polynomial functions and the Mach number M. The feasibility of the design configuration is subject to obtain certain aerodynamic performances in terms of lift coefficient C_l and pitching momentum coefficient C_m, and the modified geometry must accomplish the constraints on the airfoil thickness t, chord c, trailing edge angle τ, and thickness of the airfoil at the 85% of the chord t_{85}. The search for optimal design configuration is limited to the domain $\mathcal{X}$ bounded by the move limits imposed for the weights $I_w=[-1,1]^6$ and for the Mach number $I_M=[0.6,0.99]$. This allows for the exploration of different aerodynamic configurations, and improves the robustness of the optimization procedure in presence of an evolution of the fluid domain from the low subsonic to the transonic regime.

4.2. Aerodynamic models

The fluid domain around the airfoil is modeled through the Reynolds Averaged Navier-Stokes (RANS) equations to capture the effects of turbulence that occur at high speed regimes. The differential formulation of RANS is mathematically expressed as follows:

$$\boldsymbol{R}(\boldsymbol{U}) = \frac{\partial(\boldsymbol{U})}{\partial t} + \nabla\cdot\boldsymbol{F}^c - \nabla\cdot\boldsymbol{F}^v - \boldsymbol{Q} = 0 \quad in\,\Omega, \quad t>0 \tag{14}$$

where Ω is the computational domain, $\boldsymbol{R}$ are the numerical residuals, $\boldsymbol{Q}$ is the source term, $\boldsymbol{U}=(\rho,\rho v,\rho E)$ are the conservative variables, namely the air density $\rho=\rho(h)$, the free-stream velocity v and the total energy E, and $\boldsymbol{F}^c$ and $\boldsymbol{F}^v$ are the convective and viscous fluxes, respectively:

$$\boldsymbol{F}^c = \begin{pmatrix} \rho v \\ \rho v \otimes v + \boldsymbol{I}p \\ \rho E v + pv \end{pmatrix} \tag{15}$$

$$\boldsymbol{F}^v = \begin{pmatrix} \cdot \\ \overline{\boldsymbol{\tau}} \\ \overline{\boldsymbol{\tau}} v + k\nabla T \end{pmatrix} \tag{16}$$

where $T=T(h)$ is the free-stream temperature, $p=p(h)$ is the free-stream static pressure, $k=k(h)$ is the thermal conductivity and $\overline{\boldsymbol{\tau}}$ is the tensor of viscous stresses.

We are interested in the distribution of the pressure coefficient around the airfoil which is a function of the design configuration selected at each iteration of the optimization procedure. The aerodynamic modeling approach consists in the numerical solution of the RANS equations through a Computational Fluid Dynamic (CFD) solver to obtain a finite-dimensional approximation of the pressure coefficient over the fluid domain. We use the SU2 v6.2.0 CFD code based on the finite-volume method to discretize the RANS partial differential equations considering a fully turbulent flow-field [39]. To ensure the robustness of the aerodynamic outcomes, the convergence criteria is set for computational residuals minor than 10^{-6} with a fixed maximum number of 20000 iterations. The fluid domain is discretized through a computational mesh generated using the GMSH software v4 [40] where an automated procedure embedded within the adopted computational tool [38] adapts the hybrid grid of triangles and quadrangles elements with the modified geometry of the airfoil. This permits to target the balance between accuracy and efficiency of the CFD computations.

The pressure field around the airfoil is represented through three numerical models based on the aforementioned aerodynamic modeling approach, and differ for the accuracy and related CPU time associated with their evaluation. These models return the design objective C_d and the aerodynamic constraints on lift C_l and pitching momentum coefficients C_m given the selected design configuration in terms of modified geometry of the airfoil through the assignment of the weights $\boldsymbol{w}$ and Mach number M. The fidelity of the aerodynamic simulations is determined controlling the granularity of the computational mesh through the associated element scale factor ES where the higher the value of ES the coarser the discretization of the fluid domain. Specifically, three levels of fidelity are considered for the aerodynamic modeling: we set $ES=2.5$ for the high-fidelity model corresponding to a grid of about 90000 cells, $ES=12$ for the mid-fidelity model with about 30000 cells, and $ES=20$ for the low-fidelity model consisting of a mesh with 15000 cells.

Fig. 1 illustrates the drag coefficient computed with the aerodynamic model for different element scale factors. The high-fidelity model – marker corresponding to $ES=2.5$ in Fig. 1 – achieves an accurate representation of complex aerodynamic phenomena that occurs at higher regimes of speed including discontinuities, shock-waves and unsteadiness of the flow-field. This provides a close prediction of the mixed subsonic-supersonic fluid domain that characterizes the transonic regime. The mid-fidelity model – marker corresponding to $ES=12$ in Fig. 1 – reduces the demand for CPU if compared with the high-fidelity model by decreasing the number of cells that discretize the fluid domain. This produces a reliable estimate of the aerodynamic coefficients for Mach number regimes far from the sonic condition where the unsteady phenomena have marginal effects, and a reduced accuracy for discontinuous flows that occur at the transonic regime. The low-fidelity model – marker corresponding to $ES=20$ in Fig. 1 – further

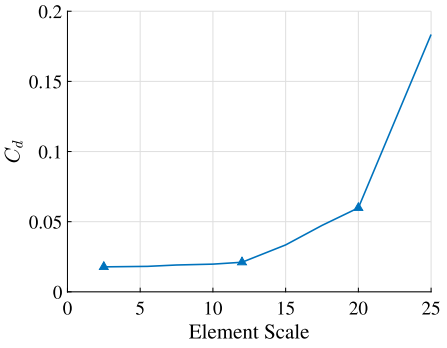


Fig. 1. Drag coefficient of the RAE-2822 airfoil varying the element scale of the CFD computational mesh. The markers indicate the drag coefficients computed adopting the element scale of the high-fidelity $ES = 2.5$, mid-fidelity $ES = 12$, and low-fidelity $ES = 20$ aerodynamic model.

decreases the number of elements adopted to discretize the fluid domain, and leads to an inaccurate representation of the compressibility effects that characterize the more turbulent flows at higher Mach values ($M > 0.65$). However, the coarser discretization reduces the computational cost required for its evaluation of the 65% with respect to the cost associated with the high-fidelity model.

For the aerodynamic representations, we define the relative computational cost $\lambda^{(l)}$ of the l-th CFD model as the ratio between the element scale factors of the high-fidelity model and the l-th level of fidelity model. Accordingly, we set $\lambda^{(L)} = 1$ for the high-fidelity aerodynamic model, $\lambda^{(2)} = 0.2$ for the mid-fidelity model, and $\lambda^{(1)} = 0.125$ for the low-fidelity model. The high-fidelity (Fig. 2), mid-fidelity (Fig. 3), and low-fidelity (Fig. 4) meshes and distributions of the pressure coefficient highlight the increasing discrepancy of the flow field representations as the level of fidelity decreases.

4.3. Physics-aware utility function for aerodynamic design

We formulate our physics-aware utility function α_4 to include a sort of physics-based reasoning that wisely quantifies the utility of querying an aerodynamic model according to the characteristics of the fluid dynamic regime. This is achieved by formalizing a bias dependent on the Mach number $\boldsymbol{\psi} = M$ as the variable representative of the physical phenomena involved in the aerodynamic domain. Accordingly, $\alpha_4(M, l)$ is formalized as follows:

$$\alpha_4(M,l) = \begin{cases} 1 & \text{if } l = 1,\dots,L-1 \\ \frac{M_s}{M_s - M} & \text{if } l = L \quad M_s = 1 \end{cases} \tag{17}$$

This physics-aware utility function encourages the query of the high-fidelity model for values of the Mach number close to the sonic condition ($M = 1$). Indeed, α_4 increases the value of the multifidelity acquisition function (Equation (8)) when an aerodynamic configuration is evaluated with the high-fidelity model in the transonic regime ($M > 0.8$). The goal is to capture large-scale separation of the fluid vein and unsteady effects that deeply influence the overall performance of the aerodynamic system. This permits to better support and improve the search for optimal designs through the a priori scientific knowledge about the aerodynamic domain structure derived directly from the governing equations.

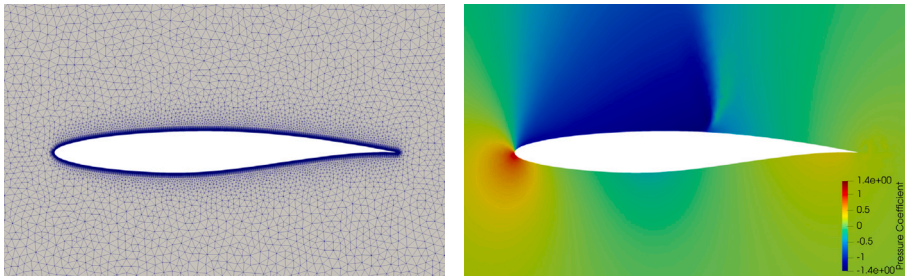
4.4. Aerodynamic design results

This section illustrates and discusses the results achieved with the physics-aware multifidelity Bayesian optimization (PA-MFBO) framework for the aerodynamic design optimization problem of the RAE 2822 transonic airfoil. The effectiveness of the PA-MFBO algorithm is compared with other existing methods commonly adopted to address black-box optimization problems, namely the single-fidelity efficient global optimization (EGO) algorithm [26], and the multifidelity Bayesian optimization based on the multifidelity expected improvement acquisition function (MFBO) [29]. For the aerodynamic design experiments, the initial set of samples is determined through a Latin hypercube strategy [41] and are used to compute the surrogate model at the first step (Section 3.1). Specifically, the multifidelity algorithms are initialized with 32 initial aerodynamic design configurations among which 20 evaluations of the objective function are obtained with the low-fidelity model, 10 are computed with the mid-fidelity model, and 2 are evaluated with the high-fidelity model, while for the single fidelity frameworks we consider 6 initial design configurations at which we compute high-fidelity observations. We select the minimum drag coefficient as the assessment metric to evaluate the capabilities of the competing algorithms, and provide a measure of the improvement of the aerodynamic performance achieved by the identified design configurations:

$$C_d^* = \min C_d(\boldsymbol{x}) \tag{18}$$

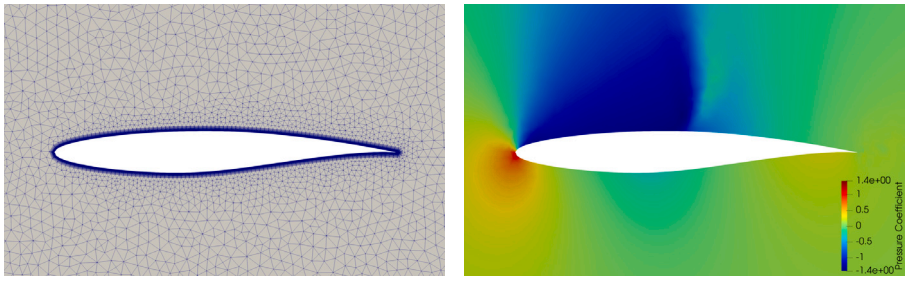
Fig. 5 reports the progression of the optimization procedure in terms of values of the minimum drag coefficient C_d^* as a function of the computational budget B. We define the computational budget $B = \sum \lambda_i^{(l)}$ as the cumulative computational cost $\lambda_i^{(l)}$ expended evaluating the l-th aerodynamic level of fidelity at each iteration of the optimization procedure. We compute 25 independent replications of the experiment for each methodology to measure and compensate the influence of the random initial sampling procedure and ensure a fair comparison of the algorithms. The outcomes of the statistics are represented through the median values of the assessment metric C_d^* together with the associated values in between the 25-th and 75-th percentiles. We adopt as the baseline solution the value of the drag coefficient $C_d^* = 0.017796$ obtained for the RAE 2822 airfoil corresponding to the design configuration $\boldsymbol{x} = [0, 0, 0, 0, 0, 0, 0.65]$. The overall convergence histories show that all the algorithms are capable to identify promising design configurations, and provide significant reductions of the drag coefficient if compared with the baseline RAE 2822 design solution. However, the PA-MFBO method leads to superior design solutions in terms of aerodynamic performance if compared with EGO and MFBO, as a result of the better management and exploitation of multiple numerical models. As can be seen, PA-MFBO reduces the drag coefficient at the beginning of the optimization procedure and identifies optimal design configurations consuming a fraction of the available computational budget. In addition, we note from the convergence of the PA-MFBO experiments that the algorithm starts the search allocating budget for the exploration of different design configurations over the domain, which corresponds to a moderate reduction of the drag coefficient. Then, the computational resources are directed towards the exploitation phase reducing the values of the design objective. After the identification of an optimal design solution, it is possible to notice that PA-MFBO queries design solutions that perform worse that the best design identified so far – the median of the minimum drag coefficient remains constant as the consumed budget increases. We observed that the algorithm in this phase mostly evaluates the low-fidelity aerodynamic model to contain the computational expense during a secondary exploration phase.

To further quantify the performance of the PA-MFBO methodology, Table 1 reports the median values of the minimum drag coefficient for incremental computational expense $B = 6$, 10, 25, 50, and 100. At the end of the initial sampling phase ($B = 6$), the identified designs of all the algorithms determine values of the drag coefficient higher than the baseline solution. However, after the collection of data from the aerodynamic models ($B = 10$), the multifidelity frameworks are capable to improve the baseline design configuration while the EGO algorithm still achieves worst designs if compared with the unmodified RAE 2822 airfoil. The PA-MFBO methodology realizes the larger design improvement of the 24.31% before consuming a Budget of $B = 50$, which is superior to the MFBO design upgrade of the 20.32% obtained adopting much more computational resources. Moreover, the EGO methodology is capable



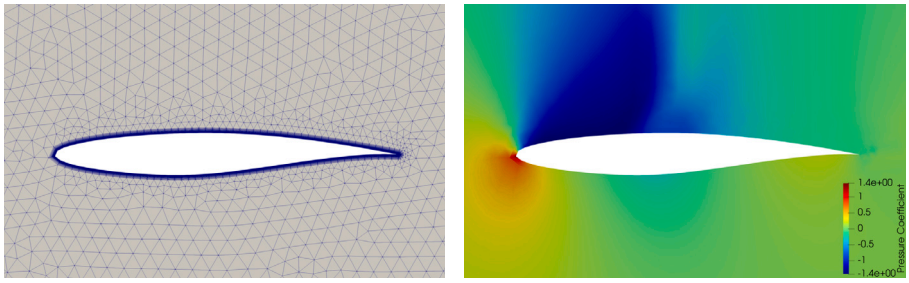
(a) (b)

Fig. 2. (a) High-fidelity discretization of the computational domain, and (b) high-fidelity pressure coefficient contours for the RAE-2822 airfoil.



(a) (b)

Fig. 3. (a) Mid-fidelity discretization of the computational domain, and (b) mid-fidelity pressure coefficient contours for the RAE-2822 airfoil.



(a) (b)

Fig. 4. (a) Low-fidelity discretization of the computational domain, and (b) low-fidelity pressure coefficient contours for the RAE-2822 airfoil.

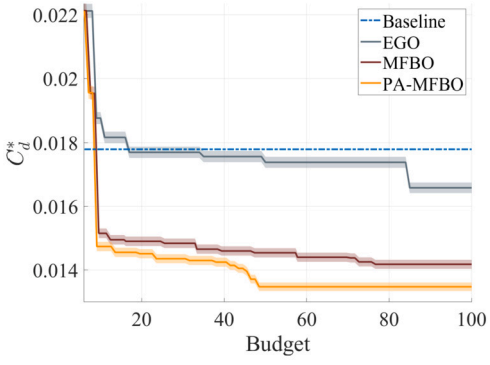


Fig. 5. Statistics over 25 runs of the minimum drag coefficient C_d^* obtained with the competing algorithms.

to deliver a design improvement of only the 6.83% using all the available computational budget. Table 2 provides details about the average evaluations of the aerodynamic models at different levels of fidelity for each competing algorithm over the statistics of 25 runs. It is possible to observe that the multifidelity methods (MFBO and PA-MFBO) drastically reduce the acquisition of high-fidelity data with respect to the single-fidelity EGO algorithm, and use lower levels of fidelity to efficiently explore the design space. However, the proposed PA-MFBO permits to better direct the computational resources towards optimal design solutions through the wise evaluation of the costly high-fidelity model guided by the physics awareness about the evolution of the aerodynamic domain. This guarantees the identification of superior design solutions with contained computational cost if compared with the standard EGO and MFBO.

Table 1
Median values of the minimum of the drag coefficient C_d^* and corresponding design improvement (·) obtained with the competing algorithms.

B	C_d^* EGO	C_d^* MFBO	C_d^* PA-MFBO
6	0.02212 (-24.30%)	0.02212 (-24.30%)	0.02212 (-24.30%)
10	0.01887 (-6.055%)	0.01515 (14.87%)	0.01455 (18.24%)
25	0.01770 (0.5394%)	0.01484 (16.61%)	0.01435 (19.36%)
50	0.01738 (2.337%)	0.01454 (18.30%)	0.01347 (24.31%)
100	0.01658 (6.833%)	0.01418 (20.32%)	0.01348 (24.31%)

To clarify and interpret the results obtained, Fig. 6 illustrates the aerodynamic performance of the optimal designs determined by all the algorithm. In particular, we report the optimal airfoil shapes corresponding to the best aerodynamic design (Fig. 6(a)), and the related pressure coefficient distribution for the PA-MFBO (Fig. 6(b)), MFBO

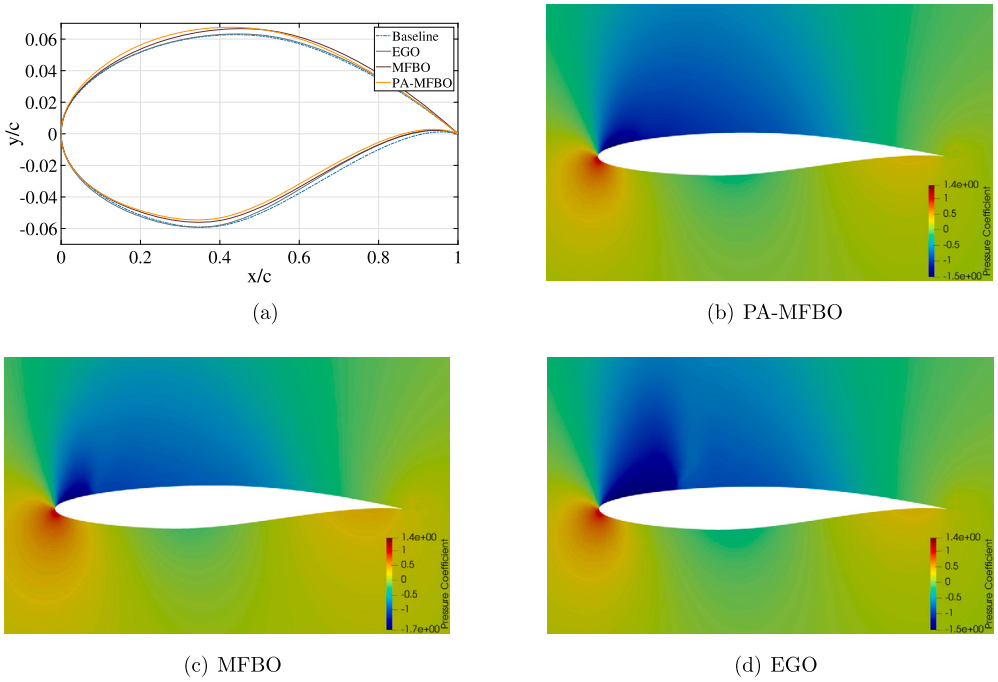


Fig. 6. (a) Optimal airfoils geometry and associated pressure coefficient contours obtained with (b) the PA-MFBO, (c) MFBO, and (d) EGO algorithm.

Table 2
Average evaluations of the high-fidelity (HF), mid-fidelity (MF), and low-fidelity (LF) aerodynamic model over the 25 runs of the aerodynamic design optimization problem.

Method	HF evaluations	MF evaluations	LF evaluations
EGO	100	-	-
MFBO	11.6	127.3	503.5
PA-MFBO	12.3	99.5	542.4

(Fig. 6(c)), and EGO (Fig. 6(d)) design solutions. It can be noticed that the superior performance of the design configuration identified with the PA-MFBO can be explained with the efficient expansion of the fluid at the upper surface leading edge that induces low-intensity shock waves if compared with the other modified airfoils. This results from the increase of the leading edge radius and aft camber that produces a reduction of the adverse pressure gradient, and permits a smooth evolution of the pressure coefficient in the supersonic bubble. These features of the aerodynamic domain determine the substantial decrease of the drag coefficient and enhance the overall efficiency of the modified airfoil. In this design test case, the remarkable performance of the PA-MFBO framework is related to the physical bias introduced in the sampling scheme. This enables the capitalization from the prior scientific knowledge about the fluid dynamic regime, and permits to accelerate and improve the optimization search through the wise selection of the aerodynamic model to query with a continuous balance between computational cost and accuracy of the solution.

5. Structural health monitoring example

The structural health monitoring problem requires the assessment of the health status of a composite skin plate of an aircraft wing. Particular attention is dedicated to the incipient fracture of the carbon fiber: this represents one of the most critical failure for laminates since involves the degradation of the mechanical properties of the material and cannot be easily detected by standard non-destructive health monitoring techniques [42]. For this application, the expert knowledge about the physics relates to specific structures of the domain characterized by damage conditions that might be misinterpreted by simplified modeling approaches. This knowledge is included in the search procedure through a physics-aware utility function that biases the query of numerical models to accurately distinguish the actual fault condition affecting the plate.

5.1. Optimization problem: composite plate health monitoring

The structural health monitoring problem demands for the identification of the damage parameters affecting a composite plate subject to a cut in the fibers. The composite plate is constituted of four layers of plain weave fabric of carbon prepreg (IM7/8552 AS4) laminated with a stacking sequence $[45°/0°/0°/45°]$, and with dimension of 102 mm transversal length, 456 mm longitudinal length, and 0.76 mm thickness of each ply. The material properties for the IM7/8552 AS4 considered in this application are reported in the data sheet published from the national center for advanced materials [43]. To reproduce an operational condition, we consider a load applied along the major dimension of the plate which represents a simplified load condition of a wing panel during the flight. The damage consists in a cut of the fibers along the transversal direction in the third layer, and is selected to simulate a critical condition where the fault involves the layer with 0° orientation that mainly contributes to support the load.

The health status of the system is represented through different damage parameters $\boldsymbol{x} = [q_1, q_2, q_3, q_4]$, including the transversal q_1 and longitudinal q_2 position of the cut, the extension of the cut q_3 along the transversal direction, and the load q_4 acting on the structure. Accordingly, the health monitoring task aims at identify the health status of the composite plate minimizing the discrepancy γ between a real-world signal measured from the real system and the same signal computed evaluating a structural numerical model. For this procedure, we adopt the strain field S as the output signal to determine the health status of the structure: this signal is sensitive to failures in the fibers and can be easily measured in real-world applications and in laboratory. Formally, this health monitoring task is formulated as an inverse optimization problem:

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in\mathcal{X}} \gamma(\boldsymbol{x}) \tag{19}$$

where the discrepancy function $\gamma(\boldsymbol{x})$ is computed as follows:

$$\gamma(\boldsymbol{x}) = RMSE\left(S_{ref}(\boldsymbol{x}), S^{(l)}_{mon}(\boldsymbol{x})\right) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\frac{(S^{j}_{ref}(\boldsymbol{x}) - S^{(l)j}_{mon}(\boldsymbol{x}))^2}{S^{j}_{ref}(\boldsymbol{x})}} \tag{20}$$

where $S_{ref}(\boldsymbol{x})$ is the reference strain field measured from the real system, $S^{(l)}_{mon}(\boldsymbol{x})$ is the strain field computed with the l-th level of fidelity numerical model, and N is the number of elements of the l-th level of fidelity field. The domain of the fault parameters $\mathcal{X} = I_{q_1} \times I_{q_2} \times I_{q_3} \times I_{q_4}$ bounds the transversal $I_{q_1} = [0, 102]$ mm and longitudinal $I_{q_2} = [0, 456]$ mm position of the cut according to the maximum dimensions of the plate, while the intervals for the length of the cut $I_{q_3} = [0, 30]$ mm and the load $I_{q_4} = [0, 20]$ N are imposed from the expert knowledge about the specific structural health monitoring problem.

5.2. Structural models

The strain field S of the composite plate is modeled through the Reissner-Mindlin plate equations [44] and numerically solved adopting the Finite Element Method (FEM). The structural modeling approach represents the composite material of the undamaged structure as an orthotropic material assigning the properties of the carbon prepreg IM7/8552 AS4, and the cut in the fiber is modeled as an homogeneous material with the mechanical properties of the matrix. The boundary conditions impose a clamp in the lower section and a displacement in the upper portion of the plate with a region extended for the 10% of the total longitudinal length. This represents a simplification of the aerodynamic load acting on a composite panel adopted for the skin of an aircraft wing.

We use the software MSC Patran and MSC Nastran to develop two FEM models and compute the strain field of the damaged composite plate at different levels of fidelity. The high-fidelity model consists of a three-dimensional representation of the structure discretized through an adaptive grid of HEXA8 3D elements characterized by a dimension of 1 mm in both the longitudinal and transversal direction near the cut region thought the border, and an increasingly coarse discretization away from the cut. This permits to capture the variation of the strain field that occurs in a small region near the cut with an high level of accuracy, while containing the overall computational cost reducing the number of elements far from the damaged location. The thickness of the plate is modeled inserting three HEXA8 elements for each of the four layers along the thickness direction to further enhance the accurate representation of the strain field. The cut is represented as a rectangular parallelepiped in the third layer characterized by a transversal extension discretized with HEXA8 elements, longitudinal extension equal to one element and thickness of three elements. The low-fidelity model approximates the composite plate through a two-dimensional representation discretized using an adaptive mesh of QUAD4 elements with transversal dimension of 2 mm and longitudinal dimension of 4 mm around the cut, and progressively increases the coarseness towards the boundaries of the plate. The cut is modeled through the same methodology of the accurate numerical model.

The high-fidelity model provides a reliable representation of the strain field as a result of the refined computational grid near the damage. This guarantees an high sensitivity to small incipient faults for which the variation of the strain field occurs in a contained region around the cut. In addition, this model allows to distinguish variations in the strain field caused by the application of intense loads in presence of a small cut in fiber – which leads to a significant variation of the strain field even in regions far from the damage – from an extended cut of the fiber – which produces large strains in an extended region due to the size of the damage. We consider the high-fidelity structural model as an emulator of the real-world composite plate that is adopted to compute the reference strain field $S_{ref}(\boldsymbol{x})$, and is used as the highest level of fidelity available to evaluate the monitoring signal $S^{(L=2)}_{mon}(\boldsymbol{x})$. The low-fidelity representation reduces the computational burden if compared with the high-fidelity model, and achieves a satisfactory accuracy of the strain field prediction for damages characterized by an extended cut in the fiber. However, the coarse discretization entails an inaccurate evaluation of the strain field for small incipient damages of the composite plate, and fails in differentiating the increase of the strain associated with small cuts in presence of significant loads from extended damages in the fiber. This results in an approximated representation of the monitoring strain signal $S^{(l=1)}_{mon}(\boldsymbol{x})$.

For this set of structural models, the relative computational cost $\lambda^{(l)}$ is measured as the average CPU time required to solve the structural model at l-th level of fidelity referred to the high-fidelity CPU time. From our preliminary experiments, we observe that the high-fidelity model requires 40 minutes to achieve convergence while the low-fidelity representation takes 8 minutes on a single core of a desktop PC with Intel Core i7-8700 (3.2 GHz) and 32 GB of RAM. Accordingly, we set $\lambda^{(L)} = 1$ for the high-fidelity structural model and $\lambda^{(1)} = 0.2$ for the low-fidelity structural model.

Fig. 7 and Fig. 8 illustrate the computational mesh and strain field over the four layers of the plate computed with the high-fidelity and low-fidelity structural model, respectively. These results are achieved for a cut in the fiber of the third layer located horizontally at 40 mm and vertically at 250 mm considering a cut length of 10 mm and load equal to 5 N.

5.3. Structural physics-aware utility function

The physics-aware utility function is conceived to incorporate expert knowledge about the appropriate structural model to be evaluated in presence of a small incipient cut concurrently with a significant load condition, or an extended damage in the fiber of the composite plate. This is realized through a bias in the search procedure $\boldsymbol{\psi} = \left[q_3, q_4\right]$ that encodes the specific structure of the domain, and is induced by the length of the fiber cut q_3 and the load applied on the plate q_4. Thus, we formalize $\alpha_4(q_3, q_4, l)$ for the health monitoring problem as follows:

$$\alpha_4(q_3, q_4, l) = \begin{cases} 1 & \text{if } l = 1, ..., L-1 \\ 0.5\frac{q_{3max}}{q_3} + 0.5\frac{1}{q_{4max} - q_4} & \text{if } l = L \end{cases} \tag{21}$$

where q_{3max} is the maximum length of the cut in the fiber and q_{4max} is the maximum load applied on the plate. This physics-aware utility function realizes a sort of expert reasoning and privileges the evaluation of the high-fidelity structural model for small incipient damages and high load values. Indeed, $\alpha_4(q_3, q_4, l)$ increases the value of the multifidelity acquisition function (Equation (8)) when the health status of the plate affected by an incipient cut and high load condition is evaluated with the high-fidelity numerical model. This ensures an accurate estimate of the strain field and permits to distinguish a narrow cut that generates large variations of the strain field amplified by high loading conditions form the magnification of strains generated by an extended cut in the fiber.

5.4. Structural health monitoring results

This section reports and discusses the results achieved with the PA-MFBO framework for the structural health monitoring example to evaluate the capabilities of the algorithm against a damage identification problem. The outcomes of the PA-MFBO are compared against the efficient global optimization (EGO) algorithm [26] and the multifidelity Bayesian optimization implementing the multifidelity expected improvement (MFBO) [29]. To assess the performance of the optimization algorithms on this test case, we compute the following assessment metrics:

$$e(q_i) = \frac{|q_i^* - q_i|}{q_i^*} \cdot 100 \tag{22}$$

$$\gamma^* = \min(\gamma(\boldsymbol{x})) \tag{23}$$

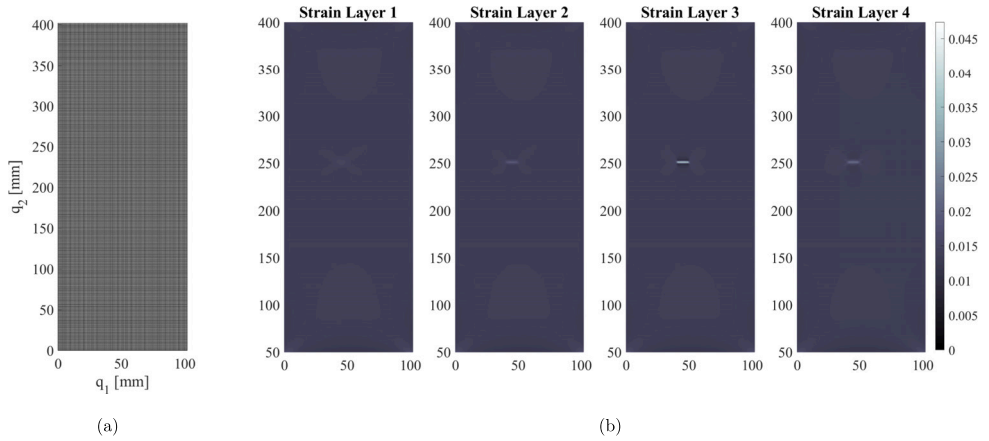


Fig. 7. (a) High-fidelity discretization of the computational domain, and (b) high-fidelity strain distribution for the four layers of the damaged composite plate.

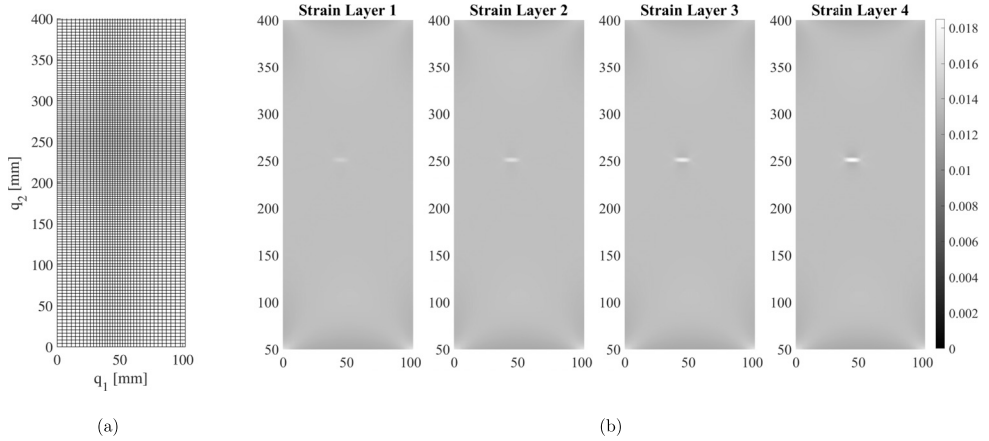


Fig. 8. (a) Low-fidelity discretization of the computational domain, and (b) low-fidelity strain distribution for the four layers of the damaged composite plate.

Table 3
Convergence results of the percentage relative error of the inference of the damage parameters $e(q_i)$, and minimum discrepancy value γ^* obtained with the competing algorithms.

Method	$e(q_1)$	$e(q_2)$	$e(q_3)$	$e(q_4)$	γ^*
PA-MFBO	0.00%	0.00%	0.00%	0.00%	0.00
MFBO	8.46%	11.5%	4.26%	0.19%	0.0261
EGO	34.1%	55.5%	10.7%	2.47%	0.0474

Table 4
Average evaluations of the high-fidelity (HF) and low-fidelity (LF) structural model over the 25 runs of the structural health monitoring problem.

Method	HF evaluations	LF evaluations
EGO	50	-
MFBO	7.7	211.5
PA-MFBO	8.1	209.5

where q_i^* is the actual level of damage that affects the composite plate, q_i is the level of damage inferred by the algorithm considering the i-th fault parameter, and $\gamma(\boldsymbol{x})$ is the value of the discrepancy between the reference strain signal and the strain field computed with the high-fidelity model. The percentage relative error $e(q_i)$ quantifies the accuracy related to the identification of the faults parameters, and γ^* represents the minimum value of the discrepancy computed by the algorithms and provides a measure of the improvement in the solution of the optimization procedure.

Fig. 9 reports the outcomes in terms of median and interval between the 25-th and 75-th percentiles for both the assessment metrics, and Table 3 summarizes the convergence results both as functions of the computational budget $B = \sum \lambda_i^{(l)}$ measured as the cumulative computational cost $\lambda_i^{(l)}$ used at each iteration i to evaluate the l-th structural level of fidelity. Overall, the multifidelity algorithms – PA-MFBO and MFBO – achieve lower values of the identification error rather than the single fidelity strategy implementing the high-fidelity structural model – EGO. However, it can be noticed that the proposed PA-MFBO is the only optimization method capable to infer the exact health status of the composite plate ($e(q_i) = 0\%$) with a computational budget of just $B = 22.8$, which corresponds to less than half of the budget consumed by EGO and MFBO ($B = 50$) to converge to suboptimal values of the identification error $e(q_i) > 0\%$. Table 4 illustrates the average number of aerodynamic models queries at different levels of fidelity for EGO, MFBO and PA-MFBO over the performed 25 tests. We notice that the multifidelity algorithms MFBO and PA-MFBO massively query the low-fidelity structural model to accelerate the exploration of different damage configurations and contain the evaluation of expensive high-fidelity analysis. However, PA-MFBO effectively evaluates the high-fidelity model to progressively reduce the inference error of damages, and accurately identify the health status of the composite plate adopting almost the same number of high-fidelity evaluation as MFBO which leads to non-negligible errors. These results suggest that the introduction of prior expert knowledge about the health monitoring problem enhances the accuracy of the damage identification procedure. A remarkable outcome is that the PA-MFBO algorithm is the only

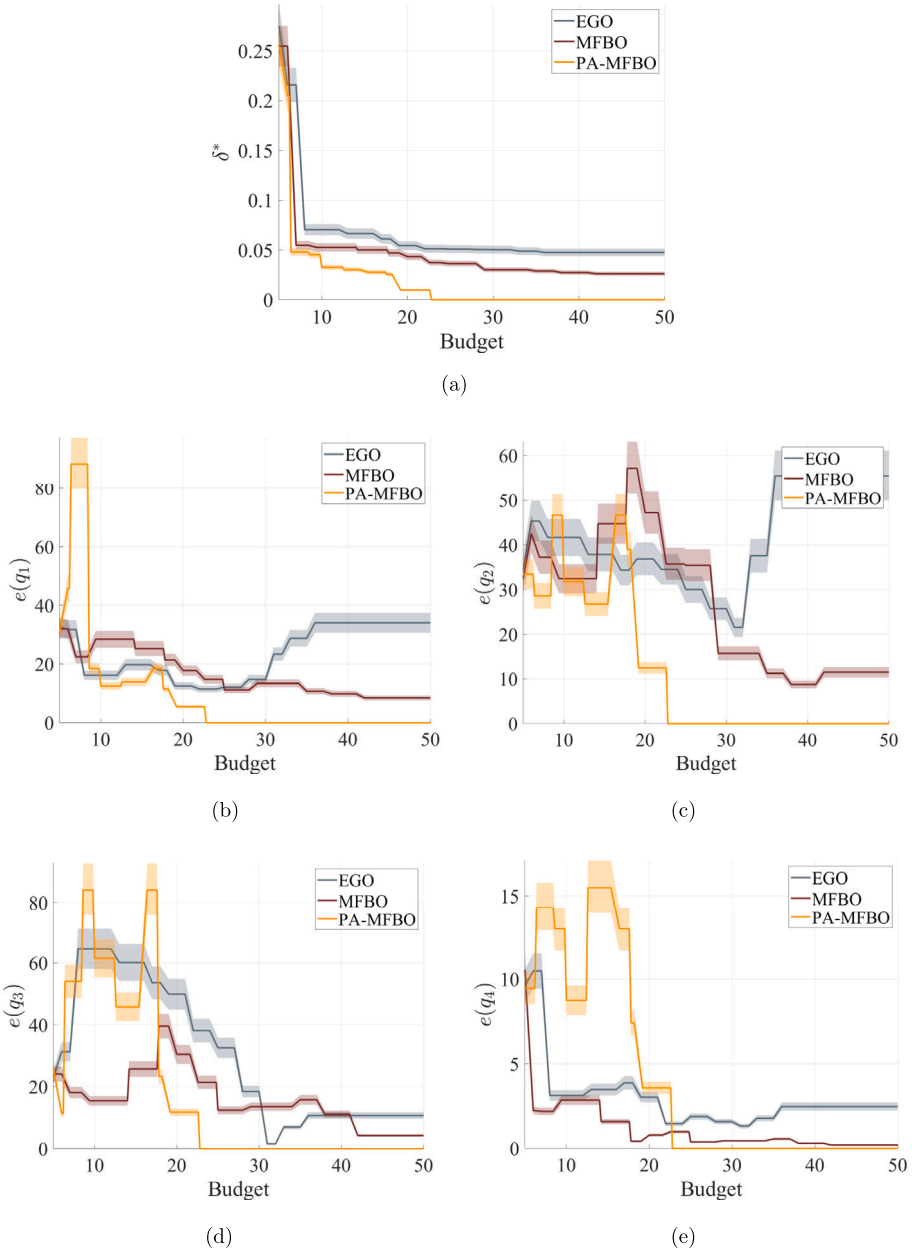


Fig. 9. Statistics over 25 runs of the minimum discrepancy γ^* and percentage relative error of the inference of the damage parameters $e(q_i)$ obtained with the competing algorithms.

optimization framework capable to accurately identify the health status of the composite plate within the allocated budget. This outcome suggests that the inclusion of the expert knowledge about the structures of the domain and the behavior of the numerical models over those structures allow to obtain a fast and robust inference performance.

We consider a statistics over 25 different combinations of fault parameters determined through the scaled Latin hypercube sampling process proposed by Berri et al. [45]. This design of experiments permits to increase the distribution of the fiber cut length located in proximity of the undamaged condition, and improves the amount of incipient damages evaluated during the experiments. In particular, the multifidelity methods start the health monitoring procedure with 17 damage configurations among which 15 faults are evaluated with the low-fidelity structural model and 2 with the high-fidelity representation, while the single-fidelity algorithm is initialized with 5 damages evaluated with the high-fidelity model.

6. Concluding remarks

This paper recognizes that domain knowledge is commonly available in science and engineering, and can be used to accelerate and improve the multifidelity optimization process. We propose a Physics-Aware Multifidelity Bayesian Optimization – PA-MFBO – framework that incorporates forms of prior scientific and expert knowledge about the physical domain during the search procedure. This is achieved introducing a learning bias formalized as a physics-aware multifidelity acquisition function that leverages the knowledge about the structure of the domain to enhance the accuracy of the solution and alleviate the computational cost for optimization.

The results achieved with the PA-MFBO are observed and discussed for an aerodynamic design problem and a structural health monitoring problem. In the design test-case, the PA-MFBO introduces a bias to pursue the awareness about the transition of fluid regimes through the Mach number. In the health assessment task, PA-MFBO incorpo-

rates a bias based on the expert knowledge about the features of the domain for specific combinations of load and extension of the damage. We note that for the two optimization problems our methodology outperforms standard single-fidelity and multifidelity Bayesian algorithms in terms of accuracy and acceleration of the search. In particular, the PA-MFBO identifies aerodynamic design solutions capable to deliver a performance improvement of the 24.31% in less than half the computational time required by competing algorithms to search suboptimal designs. Moreover, PA-MFBO is the only algorithm that permits the robust identification of damages, which otherwise would have required more computational resources or might have led to an inaccurate health assessment.

Overall, the results show the importance of embedding forms of prior knowledge in design and health monitoring optimization procedures as an enabling technique to satisfy the ever-increasing demand for performance and reliability to meet sustainability goals.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by project Multisource Frameworks to Support Real-time Structural Assessment and Autonomous Decision Making under the Visiting Professor program of Politecnico di Torino, and by the University's Doctoral Scholarship. Additional acknowledgment to Prof. Paolo Maggiore and the LABINF HPC Center of Politecnico di Torino for the support and the access to the computing resources. The authors thank Mr. Mirko Ermacora and Mr. Matteo Fenoglio for their assistance with the structural models of the composite plate.

References

[1] Martins JR, Ning A. Engineering design optimization. Cambridge University Press; 2021.
[2] Forrester AI, Sóbester A, Keane AJ. Multi-fidelity optimization via surrogate modelling. Proc R Soc A, Math Phys Eng Sci 2007;463:3251–69. https://doi.org/10.1098/rspa.2007.1900.
[3] Park C, Haftka RT, Kim NH. Remarks on multi-fidelity surrogates. Struct Multidiscip Optim 2017;55:1029–50. https://doi.org/10.1007/s00158-016-1550-y.
[4] Peherstorfer B, Willcox K, Gunzburger M. Survey of multifidelity methods in uncertainty propagation, inference, and optimization. SIAM Rev 2018;60:550–91. https://doi.org/10.1137/16m1082469.
[5] Beran PS, Bryson D, Thelen AS, Diez M, Serani A. Comparison of multi-fidelity approaches for military vehicle design. In: AIAA aviation 2020 forum; 2020. p. 3158.
[6] Thelen AS, Bryson DE, Stanford BK, Beran PS. Multi-fidelity gradient-based optimization for high-dimensional aeroelastic configurations. Algorithms 2022;15:131. https://doi.org/10.3390/a15040131.
[7] Tezzele M, Fabris L, Sidari M, Sicchiero M, Rozza G. A multifidelity approach coupling parameter space reduction and nonintrusive pod with application to structural optimization of passenger ship hulls. Int J Numer Methods Eng 2023;124:1193–210. https://doi.org/10.1002/nme.7159.
[8] Anselma P, Niutta CB, Mainini L, Belingardi G. Multidisciplinary design optimization for hybrid electric vehicles: component sizing and multi-fidelity frontal crashworthiness. Struct Multidiscip Optim 2020;62:2149–66. https://doi.org/10.1007/s00158-020-02603-6.
[9] Lai X, He X, Wang S, Wang X, Sun W, Song X. Building a lightweight digital twin of a crane boom for structural safety monitoring based on a multifidelity surrogate model. J Mech Des 2022;144:064502. https://doi.org/10.1115/1.4053606.
[10] Makkar G, Smith C, Drakoulas G, Kopsaftopoulos F, Gandhi F. A machine learning framework for physics-based multi-fidelity modeling and health monitoring for a composite wing. ASME international mechanical engineering congress and exposition, vol. 86625. American Society of Mechanical Engineers; 2022. V001T01A008.
[11] Viana FA, Simpson TW, Balabanov V, Toropov V. Special section on multidisciplinary design optimization: metamodeling in multidisciplinary design optimization: how far have we really come? AIAA J 2014;52:670–90. https://doi.org/10.2514/1.j052375.
[12] Guo Z, Song L, Park C, Li J, Haftka RT. Analysis of dataset selection for multi-fidelity surrogates for a turbine problem. Struct Multidiscip Optim 2018;57:2127–42. https://doi.org/10.1007/s00158-018-2001-8.
[13] Song J, Chen Y, Yue Y. A general framework for multi-fidelity Bayesian optimization with Gaussian processes. In: The 22nd international conference on artificial intelligence and statistics. PMLR; 2019. p. 3158–67.
[14] Močkus J. On Bayesian methods for seeking the extremum. In: Optimization techniques IFIP technical conference. Springer; 1975. p. 400–4.
[15] Snoek J, Larochelle H, Adams RP. Practical Bayesian optimization of machine learning algorithms. Adv Neural Inf Process Syst 2012;25.
[16] Frazier PI. Bayesian optimization. In: Recent advances in optimization and modeling of contemporary problems. Informs; 2018. p. 255–78.
[17] Meliani M, Bartoli N, Lefebvre T, Bouhlel M-A, Martins JR, Morlier J. Multi-fidelity efficient global optimization: methodology and application to airfoil shape design. In: AIAA aviation 2019 forum; 2019. p. 3236.
[18] Tran A, Tranchida J, Wildey T, Thompson AP. Multi-fidelity machine-learning with uncertainty quantification and Bayesian optimization for materials design: application to ternary random alloys. J Chem Phys 2020;153:074705. https://doi.org/10.1063/5.0015672.
[19] Serani A, Pellegrini R, Wackers J, Jeanson C-E, Queutey P, Visonneau M, et al. Adaptive multi-fidelity sampling for cfd-based optimisation via radial basis function metamodels. Int J Comput Fluid Dyn 2019;33:237–55. https://doi.org/10.1080/10618562.2019.1683164.
[20] Perdikaris P, Karniadakis GE. Model inversion via multi-fidelity Bayesian optimization: a new paradigm for parameter estimation in haemodynamics, and beyond. J R Soc Interface 2016;13:20151107. https://doi.org/10.1098/rsif.2015.1107.
[21] Di Fiore F, Maggiore P, Mainini L. Multifidelity domain-aware learning for the design of re-entry vehicles. Struct Multidiscip Optim 2021;64:3017–35. https://doi.org/10.1007/s00158-021-03037-4.
[22] Di Fiore F, Mainini L. Nm-mf: non-myopic multifidelity framework for constrained multi-regime aerodynamic optimization. AIAA J 2023;61:1270–80. https://doi.org/10.2514/1.j062219. 9.
[23] Williams C, Rasmussen C. Gaussian processes for regression. Adv Neural Inf Process Syst 1995;8.
[24] Schulz E, Speekenbrink M, Krause A. A tutorial on Gaussian process regression: modelling, exploring, and exploiting functions. J Math Psychol 2018;85:1–16. https://doi.org/10.1101/095190.
[25] Kennedy MC, O'Hagan A. Predicting the output from a complex computer code when fast approximations are available. Biometrika 2000;87:1–13. https://doi.org/10.1093/biomet/87.1.1.
[26] Jones DR, Schonlau M, Welch WJ. Efficient global optimization of expensive black-box functions. J Glob Optim 1998;13:455. https://doi.org/10.1023/A:1008306431147.
[27] Kushner HJ. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. J Basic Eng 1964;86:97–106. https://doi.org/10.1115/1.3653121.
[28] Hennig P, Schuler CJ. Entropy search for information-efficient global optimization. J Mach Learn Res 2012;13.
[29] Huang D, Allen TT, Notz WI, Miller RA. Sequential kriging optimization using multiple-fidelity evaluations. Struct Multidiscip Optim 2006;32:369–82. https://doi.org/10.1007/s00158-005-0587-0.
[30] Ruan X, Jiang P, Zhou Q, Hu J, Shu L. Variable-fidelity probability of improvement method for efficient global optimization of expensive black-box problems. Struct Multidiscip Optim 2020;62:3021–52. https://doi.org/10.1007/s00158-020-02646-9.
[31] Zhang Y, Hoang TN, Low BKH, Kankanhalli M. Information-based multi-fidelity Bayesian optimization. In: NIPS workshop on Bayesian optimization; 2017.
[32] Takeno S, Fukuoka H, Tsukada Y, Koyama T, Shiga M, Takeuchi I, et al. Multi-fidelity Bayesian optimization with max-value entropy search and its parallelization. In: International conference on machine learning. PMLR; 2020. p. 9334–45.
[33] Forrester A, Sobester A, Keane A. Engineering design via surrogate modelling: a practical guide. John Wiley & Sons; 2008.
[34] Drela M. Pros and cons of airfoil optimization. Frontiers Comput Fluid Dyn 1998:363–81. https://doi.org/10.1142/9789812815774_0019.
[35] Li W, Huyse L, Padula S. Robust airfoil optimization to achieve drag reduction over a range of Mach numbers. Struct Multidiscip Optim 2002;24:38–50. https://doi.org/10.1007/s00158-002-0212-4.
[36] Elger DF, LeBret BA, Crowe CT, Roberson JA. Engineering fluid mechanics. John Wiley & Sons; 2020.
[37] Young DF, Munson BR, Okiishi TH, Huebsch WW. A brief introduction to fluid mechanics. John Wiley & Sons; 2010.
[38] Quagliarella D, Diez M. An open-source aerodynamic framework for benchmarking multi-fidelity methods. In: AIAA aviation 2020 forum; 2020. p. 3179.
[39] Economon TD, Palacios F, Copeland SR, Lukaczyk TW, Alonso JJ. Su2: an open-source suite for multiphysics simulation and design. AIAA J 2016;54:828–46. https://doi.org/10.2514/1.j053813.

[40] Geuzaine C, Remacle Gmsh J-F. A 3-d finite element mesh generator with built-in pre- and post-processing facilities. Int J Numer Methods Eng 2009;79:1309–31. https://doi.org/10.1002/nme.2579.
[41] Viana FA. A tutorial on Latin hypercube design of experiments. Qual Reliab Eng Int 2016;32:1975–85. https://doi.org/10.1002/qre.1924.
[42] Giurgiutiu V. Structural health monitoring of aerospace composites; 2015. p. 449–507.
[43] H. Corporation. Hexply® 8552 product datasheet; 2016.
[44] Mindlin R. Influence of rotatory inertia and shear on flexural motions of isotropic, elastic plates. J Appl Mech 1951;18:31–8. https://doi.org/10.1115/1.4010217.
[45] Berri PC, Dalla Vedova MDL, Mainini L. Real-time fault detection and prognostics for aircraft actuation systems. In: AIAA Scitech 2019 forum; 2019. p. 2210.