

Bayesian Optimization of Multiobjective Functions Using Multiple Information Sources

Danial Khatamsaz,^{*} Lalith Peddareddygar,† Samuel Friedman,‡ and Douglas Allaire§
Texas A&M University, College Station, Texas 77843

<https://doi.org/10.2514/1.J059803>

Multiobjective optimization is often a difficult task owing to the need to balance competing objectives. A typical approach to handling this is to estimate a Pareto frontier in objective space by identifying nondominated points. This task is typically computationally demanding owing to the need to incorporate information of high enough fidelity to be trusted in design and decision-making processes. In this work, we present a multi-information source framework for enabling efficient multiobjective optimization. The framework allows for the exploitation of all available information and considers both potential improvement and cost. The framework includes ingredients of model fusion, expected hypervolume improvement, and intermediate Gaussian process surrogates. The approach is demonstrated on a test problem and an aerostructural wing design problem.

Nomenclature

d	=	number of input space dimensions
e	=	column vector of ones
f_i	=	function model of information source i
$\mathcal{H}$	=	hypervolume
$\mathcal{H}_T$	=	hypervolume improvement
k	=	kernel function
l_h	=	length scale for dimension h
N_i	=	number of evaluations from information source i
$\mathcal{S}$	=	set of known nondominated solutions
X	=	set of randomly sampled points from input space
$\mathbf{x}$	=	set of evaluated design points from input space
$\mathbf{y}$	=	objective vector
$\tilde{\Sigma}$	=	covariance matrix
μ_i	=	mean function of Gaussian process for information source i
ρ_{ij}	=	correlation coefficient between information sources i and j
σ_s	=	signal noise
$\sigma_{f,i}^2$	=	variance related to the fidelity of information source i
$\sigma_{GP,i}^2$	=	variance related to the Gaussian process of information source i
σ_i^2	=	total variance of Gaussian process for information source i
$\mathcal{X}$	=	feasible input space

Subscript

GP	=	Gaussian process
----	---	------------------

I. Introduction

WHEN estimating a ground truth quantity of interest (for example, fuel burn for an aircraft or a particular material property), we can often consider different mathematical formulations of the analysis or prediction. This, in addition to experimental data and expert opinion, can give rise to the ability to use multiple different sources of information for the estimation task at hand. The different assumptions made lead to differing levels of fidelity among the sources, as well as different costs, in terms of both time and monetarily. In the presence of multiple sources of information, we seek analysis and design processes that exploit the extra information that would not be present if only a single source were available. The opportunity is the efficient selection of which information source to query and where to query it on the bases of cost and potential for improvement in the estimation of a quantity of interest. To do so, we employ a Bayesian optimization framework well suited to the optimization of black-box functions. These approaches generally use an acquisition function to search the design space effectively and efficiently through the tradeoff of exploration and exploitation. The challenge is ensuring proper fusion of information as it becomes available and a need for a rapid capability for moving from prior predictive information to posterior predictive information without necessarily executing a true information source. The standard practice of using Gaussian processes (GPs) within Bayesian optimization frameworks as updatable surrogates provides an avenue for efficiently incorporating information source fusion within the search process. We exploit this here. We also note here that our goal is to estimate and optimize properties or other performance metrics of real systems. There is, therefore, a notion of ground truth, which is the true quantity being estimated, which is likely only observable with noise. Often, this ground truth is represented as the information that can be acquired from the highest-fidelity information source. This assumption may be reasonable in some circumstances, particularly if the information source is an experiment with the realized system. Here, we keep the term ground truth to ensure that the overall goal is clear; and we use our highest-fidelity information source as a proxy for that ground truth. In prior work by Ghoreishi et al. (see, e.g., Ref. [1]), the case where ground truth is measured with noise is handled. Here, we do not include the noise in the ground truth for clarity, but the framework we present can incorporate this if known.

Whereas a multi-information source capability can be applicable to a wide variety of contexts, our focus here is on multiobjective optimization. Previous work in this area, particularly with emphasis on multifidelity methods, includes (for example) an efficient global optimization (EGO) approach based on the use of a hypervolume indicator technique and surrogate model creation for every objective [2]. In Ref. [3], a point-by-point approach is employed that considers the ends of the estimated Pareto front in an effort to obtain better solutions via single objective optimizations. Reference [4] encour-

Presented as Paper 2020-2127 at AIAA SciTech 2020 Forum, Special Session: Managing Multiple Information Sources of Multi-Physics Systems, Orlando, FL, January 6–10, 2020; received 24 May 2020; revision received 29 October 2020; accepted for publication 10 December 2020; published online 25 March 2021. Copyright © 2021 by the American Institute of Aeronautics and Astronautics, Inc. All rights reserved. All requests for copying and permission to reprint should be submitted to CCC at www.copyright.com; employ the eISSN 1533-385X to initiate your request. See also AIAA Rights and Permissions www.aiaa.org/randp.

^{*}Ph.D. Candidate, J. Mike Walker '66 Department of Mechanical Engineering; danialkh26@tamu.edu. Student Member AIAA.

[†]Graduate Student, J. Mike Walker '66 Department of Mechanical Engineering; plalith200@tamu.edu. Student Member AIAA.

[‡]Ph.D. Candidate, J. Mike Walker '66 Department of Mechanical Engineering; samfriedman@tamu.edu. Student Member AIAA.

[§]Assistant Professor, J. Mike Walker '66 Department of Mechanical Engineering; dallaire@tamu.edu. Senior Member AIAA.

ages the use of standard multiobjective evolutionary algorithm introduced in Refs. [5,6] to apply on the lower-fidelity information source to build a surrogate model to search and obtain a Pareto front. Then, the high-fidelity information source is evaluated at those nondominated designs to correct the surrogate model associated to the lower-fidelity information source. A similar approach is suggested in Ref. [7] using a surrogate model built with samples from a low-fidelity information source to search the design space for potential nondominated designs. A high-fidelity information source is then evaluated at those design points, and the approximation of the Pareto front is obtained by optimizing a cokriging model constructed with these new evaluations.

In this work, we present a novel framework for exploiting available information sources to identify nondominated points in objective space to estimate the true Pareto front of a given problem. To handle fusion of information sources, we incorporate model reification introduced in Ref. [8], which builds off Refs. [9,10]. Model reification is a fusion technique that learns correlations among information sources and guards against overconfidence that can occur when nearly identical sources are used. This fusion process is also nonhierarchical and allows fidelity to vary over a design space. To enable rapid assessment of posterior predictive information, we use GPs as intermediate surrogate models that may be temporarily updated with candidate query points [8,11–13]. To drive candidate query points toward the Pareto front, we use the expected hypervolume improvement metric presented in Ref. [14]. The approach presented in Ref. [14] provides an exact means of calculating the expected hypervolume improvement. This leads to an efficient computational process since a closed-form expression can be used to find the expected hypervolume improvement (EHVI). Overall, our novel Bayesian multi-information source multiobjective optimization framework can exploit multiple nonhierarchical information sources in an efficient manner that produces higher-quality Pareto fronts at less computational expense than current available methods. This is achieved via combined use of model reification-based information fusion within a Bayesian optimization paradigm over a set of available information sources where querying is directed by an easily computable closed-form acquisition function based on the EHVI. We have chosen a Bayesian optimization paradigm here because the problems we seek to address involve data-driven optimization. That is, our objective function estimates, as computed by available information sources, are analytically unknown and must be learned during search. Although there are other optimization strategies, such as the model management approaches of Refs. [15,16] and model fusion approaches of Refs. [17,18], Bayesian optimization is viewed as a superior computational strategy when tasks of exploration and exploitation must be traded off as discussed in Ref. [19]. Our approach is demonstrated on a test problem with a two-dimensional input space. We then demonstrate our approach on an aerostuctural wing design problem involving a 17-dimensional input space. These input space dimensions stress the limits of typical Gaussian process regression modeling, and our approach is still shown to perform well. In each demonstration, we consider two objectives; however, this is not a limitation of the work.

The rest of the paper is organized as follows. In Sec. II, a background on surrogate modeling, fusion, and multiobjective problems is provided. Then, in Sec. III, the approach to build the optimization algorithm is addressed, supported by pseudocode and a flowchart. Section IV then presents the results on the test function and the aerostuctural design optimization problem. Conclusions are then drawn in Sec. V.

II. Background

Our multi-information source optimization approach for multiple objectives employs GPs as intermediate surrogate models and fuses information using the process of model reification. We describe each of these ingredients in turn in this section. We then conclude this section with background on a general multiobjective optimization formulation based on the Pareto frontier, which is how we approach such problems here. Once we have established the necessary ingredients of our approach, we move to a description of our formal

hypervolume indicator-based framework for multi-information source multiobjective optimization in Sec. III.

A. Gaussian Process Regression Surrogates

For the estimation of a quantity of interest (e.g., an objective or constraint), there are often available several information sources that may be employed. These encompass computational models, experiments, expert opinions, etc., and have varying fidelity (that also varies over the input space) and varying costs. Our approach here seeks to optimally exploit all available information sources by balancing the cost and fidelity of each information source when determining which source and where in the input space to query. Following Refs. [1,20], we assume we have some set of information sources $f_i(\mathbf{x})$ available, where $i \in \{1, 2, \dots, S\}$, that can be used to estimate the quantity of interest $f(\mathbf{x})$ at design point $\mathbf{x}$. We note here that in the multiobjective context we address in this work, the output of the information sources can be vector valued. To predict the output of each information source at input configurations that have not yet been executed, a surrogate model is constructed for each information source using Gaussian process regression [11]. Easy manipulation and implementation of Gaussian process models make them a powerful tool in the Bayesian context for probabilistic modeling purposes. A Gaussian process model provides a normal distribution in the objective space defined by mean and covariance functions for any sample in the design space. Additionally, Gaussian process surrogates can be constructed using different kernel functions defining the correlations between the data points. This leads to more flexibility in fitting a distribution to the training data. The use of GPs with Bayesian optimization is standard practice, and we follow that practice here. These surrogates are denoted by $f_{GP,i}(\mathbf{x})$. Assuming we have available N_i evaluations of information source i denoted by $\{X_{N_i}, y_{N_i}\}$, where $X_{N_i} = (\mathbf{x}_{1,i}, \dots, \mathbf{x}_{N_i,i})$ represents the N_i input samples to information source i and $y_{N_i} = (f_i(\mathbf{x}_{1,i}), \dots, f_i(\mathbf{x}_{N_i,i}))$ represents the corresponding outputs from information source i , the posterior distribution of information source i at design point $\mathbf{x}$ is given as

$$f_{GP,i}(\mathbf{x})|X_{N_i}, y_{N_i} \sim \mathcal{N}(\mu_i(\mathbf{x}), \sigma_{GP,i}^2(\mathbf{x})) \quad (1)$$

where

$$\begin{aligned} \mu_i(\mathbf{x}) &= K_i(X_{N_i}, \mathbf{x})^T [K_i(X_{N_i}, X_{N_i}) + \sigma_{n,i}^2 I]^{-1} y_{N_i} \\ \sigma_{GP,i}^2(\mathbf{x}) &= k_i(\mathbf{x}, \mathbf{x}) - K_i(X_{N_i}, \mathbf{x})^T \\ &\quad [K_i(X_{N_i}, X_{N_i}) + \sigma_{n,i}^2 I]^{-1} K_i(X_{N_i}, \mathbf{x}) \end{aligned} \quad (2)$$

where k_i is a real-valued kernel function over the input space, $K_i(X_{N_i}, X_{N_i})$ is the $N_i \times N_i$ matrix whose m, n entry is $k_i(\mathbf{x}_{m,i}, \mathbf{x}_{n,i})$, and $K_i(X_{N_i}, \mathbf{x})$ is the $N_i \times 1$ vector whose m th entry is $k_i(\mathbf{x}_{m,i}, \mathbf{x})$ for information source i . Here we have included the term $\sigma_{n,i}^2$, which can be used to model the observation error for information sources based on experiments. For the kernel function, without loss of generality, we employ the squared exponential covariance function. This is specified as

$$k_i(\mathbf{x}, \mathbf{x}') = \sigma_s^2 \exp\left(-\sum_{h=1}^d \frac{(x_h - x'_h)^2}{2l_h^2}\right) \quad (3)$$

where d is the dimension of the input space, σ_s^2 is the signal variance, and l_h (where $h = 1, 2, \dots, d$) is the characteristic length scale that indicates the correlation between the points within dimension h . The parameters σ_s^2 and l_h associated with each information source can be estimated by maximizing the log marginal likelihood.

For each information source surrogate, we further quantify the uncertainty, or discrepancy, with respect to the ground truth quantity of interest by adding a term associated with the fidelity of the information source at a given location in the input space. Specifically, we quantify the total variance, which captures both the variance associated with the Gaussian process representation and the quantified

variance associated with the fidelity of the information source over the input space as

$$\sigma_i^2(\mathbf{x}) = \sigma_{\text{GP},i}^2(\mathbf{x}) + \sigma_{f,i}^2(\mathbf{x}) \quad (4)$$

where $\sigma_{f,i}^2(\mathbf{x})$ is the variance related to the fidelity of information source i that has been estimated from, for example, expert opinion or available real-world data. Here, we estimate this variance by computing the absolute difference between the available data from the true quantity of interest and the information source. A GP is then performed using the square of these error values as training points to estimate the discrepancy variance over the input space as the mean of the GP. This is described in more detail in Ref. [21].

B. Fusion of Information from Multiple Sources

Following Refs. [1,8,21,22], we assume that every information source contains potentially useful information regarding a given quantity of interest. Given that this information may be correlated across information sources, we aim to accurately fuse available new information from each source query so as to ensure we use our resources as efficiently as possible. Our approach, unlike most traditional multi-fidelity approaches [12,15,23–28], does not assume a hierarchy of information sources; and our goal is not optimization with the highest-fidelity source but optimization with respect to ground truth.

Several approaches exist for fusing multiple sources of information. Among these are approaches such as Bayesian modeling averaging [29–34], the use of adjustment factors [35–38], covariance intersection methods [39], and fusion under known correlation [9,40,41]. As noted previously, we assume that every information source contains useful information regarding the ground truth quantity of interest. Thus, as more information sources are incorporated into a fusion process, we expect the variance of the quantity of interest estimates to decrease. This is not necessarily the case for all of the aforementioned fusion techniques, with the exception of fusion under known correlation. Thus, there is significant value in determining correlations before fusion.

Following the work of Refs. [1,8,21,22], we note that since our information sources are represented by intermediate Gaussian processes, their fusion follows that of normally distributed information. Under the case of known correlations between the discrepancies of information sources, the fused mean and variance are shown to be [9]

$$\mathbb{E}[\hat{f}(\mathbf{x})] = \frac{\mathbf{e}^T \tilde{\Sigma}(\mathbf{x})^{-1} \boldsymbol{\mu}(\mathbf{x})}{\mathbf{e}^T \tilde{\Sigma}(\mathbf{x})^{-1} \mathbf{e}} \quad (5)$$

$$\text{Var}(\hat{f}(\mathbf{x})) = \frac{1}{\mathbf{e}^T \tilde{\Sigma}(\mathbf{x})^{-1} \mathbf{e}} \quad (6)$$

where $\mathbf{e} = [1, \dots, 1]^T$, $\boldsymbol{\mu}(\mathbf{x}) = [\mu_1(\mathbf{x}), \dots, \mu_S(\mathbf{x})]^T$ given S models, and $\tilde{\Sigma}(\mathbf{x})^{-1}$ is the inverse of the covariance matrix between the information sources.

To estimate the correlation coefficients between information sources over the domain, we use the reification process defined in Refs. [8,22]. In this process, to estimate the correlation coefficients between the deviations of information sources i and j , each of the information sources i and j , one at a time, is reified or treated as ground truth. Assuming that information source i is reified, the correlation coefficients between the information sources i and j , for $j = 1, \dots, i-1, i+1, \dots, S$, are given as

$$\rho_{ij}(\mathbf{x}) = \frac{\sigma_i^2(\mathbf{x})}{\sigma_i(\mathbf{x})\sigma_j(\mathbf{x})} = \frac{\sigma_i(\mathbf{x})}{\sqrt{(\mu_i(\mathbf{x}) - \mu_j(\mathbf{x}))^2 + \sigma_j^2(\mathbf{x})}} \quad (7)$$

where $\mu_i(\mathbf{x})$ and $\mu_j(\mathbf{x})$ are the mean values of the GPs of information sources i and j , respectively; and $\sigma_i^2(\mathbf{x})$ and $\sigma_j^2(\mathbf{x})$ are the total variances at $\mathbf{x}$. Next, information source j is reified to estimate $\rho_{ji}(\mathbf{x})$. The variance weighted average of the two estimated correlation coefficients can then be used as the estimate of the correlation between the errors as

$$\bar{\rho}_{ij}(\mathbf{x}) = \frac{\sigma_j^2(\mathbf{x})}{\sigma_i^2(\mathbf{x}) + \sigma_j^2(\mathbf{x})} \rho_{ij}(\mathbf{x}) + \frac{\sigma_i^2(\mathbf{x})}{\sigma_i^2(\mathbf{x}) + \sigma_j^2(\mathbf{x})} \rho_{ji}(\mathbf{x}) \quad (8)$$

These average correlations are then used to calculate the covariance matrix to estimate the fused mean and variance in Eqs. (5) and (6).

C. Multiobjective Optimization

A multiobjective optimization problem can be defined as

$$\text{minimize}\{f_1(\mathbf{x}), \dots, f_n(\mathbf{x})\}, \mathbf{x} \in \mathcal{X} \quad (9)$$

where $f_1(\mathbf{x}), \dots, f_n(\mathbf{x})$ are the objectives, and $\mathcal{X}$ is the feasible design space. Throughout this work, we develop unconstrained approaches; however, the inclusion of penalty terms could be considered for constraint handling. Another possibility for constraint handling could involve the construction of a Lagrangian, where the objective is the EHVI and constraints are incorporated in the usual fashion. This could provide a means for evaluating the Karush–Kuhn–Tucker conditions within the Bayesian optimization framework and an open avenue for exciting future work in algorithmic development aimed at pursuing the satisfaction of these conditions. For problems such as Eq. (9), it is usually the case that no single point optimizes each individual objective simultaneously. To deal with this, approaches based on the creation of a scalar objective using utility theory are common, as well as approaches based on finding non-dominated solutions approaching the Pareto frontier. We focus on the latter here. For this case, optimal solutions $\mathbf{y}$ to a multiobjective problem with n objectives are denoted as $\mathbf{y} < \mathbf{y}'$, and they are defined as

$$\begin{aligned} \{\mathbf{y} : \mathbf{y} = (y_1, y_2, \dots, y_n), y_i \leq y'_i \forall i \in \{1, 2, \dots, n\}, \\ \exists j \in \{1, 2, \dots, n\} : y_j < y'_j\} \end{aligned} \quad (10)$$

where $\mathbf{y}' = (y'_1, y'_2, \dots, y'_n)$ denotes any possible objective output. The set of $\mathbf{y} \in \mathcal{Y}$, where $\mathcal{Y}$ is the objective space, is the Pareto front of the problem. This is shown conceptually for a biobjective problem in Fig. 1. All points on the Pareto front are nondominated. Our approach, which is common in the literature, is thus to find the Pareto front as efficiently as possible for a given multiobjective optimization problem.

There are many techniques in use for approximating Pareto frontiers for multiobjective optimization problems. Among these are the weighted sum approach [42], the adaptive weighted sum approach [43], normal boundary intersection methods [44], hypervolume indicator methods [45–51], and others. The hypervolume indicator approach is well suited to expected improvement-based algorithms, which have been shown to work well in a multiple information source setting (see, e.g., Refs. [1,13,17,20]). Thus, our approach proposes the incorporation of hypervolume indicator improvement within a multi-information source querying framework. Hypervolume indicator approaches are based on the concept of a hypervolume in objective space. These hypervolumes are measured relative to a fixed

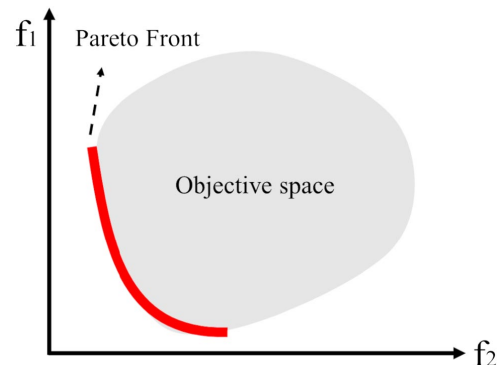


Fig. 1 All points on the red line are nondominated and constitute the solution set.

reference point, and the enclosed volume between the approximated set of Pareto points and the reference point is computed. The concept is shown notionally in Fig. 2. Here, the shaded area is the hypervolume to be computed. In general, if a given set of points has a higher hypervolume than another set, then the given set is a better estimator of the Pareto front. Hypervolume indicator algorithms seek to maximize the hypervolume in objective space so as to best approximate the Pareto front. Thus, the value of new query points can be estimated (using prior predictive distributions if using GPs) by measuring the expected improvement in the hypervolume that would occur, given that the query takes place. In Fig. 2, the shaded area in blue shows the amount of increase in hypervolume when a new nondominated point is found and added to the solution set.

III. Approach

Bayesian optimization is an optimization technique aimed at learning what is needed about an underlying black-box function to efficiently optimize it (see, e.g., Refs. [52–55]). As such, Bayesian optimization methods seek to trade off the tasks of exploration and exploitation. These methods traditionally employ Gaussian process surrogate models that can be temporarily updated to assess the quality of a candidate query point. This quality is measured by an acquisition function, such as expected improvement, probability of improvement, the knowledge gradient, and others. Our multiobjective optimization approach here treats information sources as black boxes and uses the EHVI acquisition function. Thus, our proposed method is one of Bayesian optimization.

Generally, our approach is based on determining, with available prior information, where to query and what source to query to maximize the hypervolume indicator while being budget aware. To achieve this, we make use of the updatable Gaussian process surrogates described previously for each information source. These surrogates can be used as prior predictive distributions that can be temporarily updated with potential query locations that result in potential changes to the hypervolume indicator. By searching over the space of potential query locations and potential information sources with these prior predictive surrogates, we are able to efficiently identify the next best query to execute. Once this query is executed, all surrogates (including correlation information) may be updated; and then they can serve as prior predictive distributions for the next iteration. In this section, we describe in detail our approach to achieving this. We begin with the necessary preliminaries regarding the fast calculation of the expected hypervolume improvement [56] within a multi-information source framework. This discussion follows largely from Ref. [14], where more details can be found if desired. We then describe our algorithm for multi-information source multiobjective optimization. In Sec. IV, we demonstrate the use of this framework on a test problem and an aircraft wing design problem.

A. Preliminaries

Following Ref. [14] for the development of the fast computation of EHVI, we present here our implementation within a multi-information source setting. We begin by considering a current solution set $\mathcal{S}$

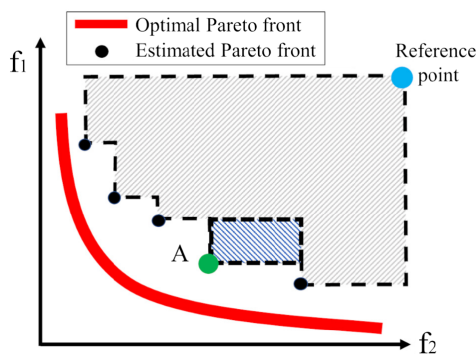


Fig. 2 The blue shaded region corresponds to the hypervolume improvement due to adding point A to the solution set.

of nondominated points in objective space at some point during a multiobjective optimization process. The dominated hypervolume, denoted as $\mathcal{H}(\mathcal{S})$, can then be computed given $\mathcal{S}$ and a reference point. Improvement to the hypervolume due to adding a new solution vector $\mathbf{y}$ is then defined as

$$\mathcal{H}_{\mathcal{I}}(\mathbf{y}, \mathcal{S}) = \mathcal{H}(\mathcal{S} \cup \mathbf{y}) - \mathcal{H}(\mathcal{S}) \quad (11)$$

If $\mathcal{H}_{\mathcal{I}}(\mathbf{y}, \mathcal{S}) > 0$, then $\mathbf{y}$ is in the nondominated region of $\mathcal{S}$ and can be used to update the solution set. Otherwise, there is no improvement over $\mathcal{H}(\mathcal{S})$ by adding $\mathbf{y}$ and the query adds no value. In the context of Bayesian optimization, $\mathbf{y}$ is a random output of a probabilistic model related to a potential solution in the design space. Hence, $\mathcal{H}_{\mathcal{I}}(\mathbf{y}, \mathcal{S})$ is also a random variable. Therefore, it is possible to calculate its expected value, which is the EHVI. Comparing EHVI values for different potential solutions in the design space and finding the maximum EHVI leads to an information-economic querying policy that ensures maximum gains are achieved in each successive query. In a multi-information source context, however, the different costs of querying each source should also be taken into account.

The formula for calculating EHVI as outlined in Ref. [57] is given as

$$\mathbb{E}[\mathcal{H}_{\mathcal{I}}(\mathbf{y})] = \int_U \mathbb{P}(\mathbf{y} < \mathbf{y}') d\mathbf{y}' \quad (12)$$

where $\mathbb{P}(\mathbf{y} < \mathbf{y}')$ is the probability that $\mathbf{y}'$ is dominating $\mathbf{y}$, and U is the dominated hypervolume. In our context, this can be computed in closed form, as will be shown in the following. Given that we have independent Gaussian process models for every objective for each information source, the posterior predictive output of each model given the data is a random variable identified as $y_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, where $i \leq m$ and μ_i , and σ_i^2 are the mean and variance of the i th objective, accordingly (note that we have not included information source specific indices here for notational clarity). For a new potential solution in the design space, we have the following equation:

$$\mathbb{P}(\mathbf{y} < \mathbf{y}') = \prod_{i=1}^m \Phi\left(\frac{y'_i - \mu_i}{\sigma_i}\right) \quad (13)$$

where Φ is the cumulative distribution function of the standard normal random variable. Details regarding the closed-form expression of Eq. (12) along with a fast approach to compute the hypervolume associated with a solution set can be found in Refs. [14,57–60].

B. Multi-Information Source Multiobjective Optimization Framework

Using the GP as the surrogate model for each objective and EHVI as the acquisition function, we can perform Bayesian optimization to approximate a solution set for a multiobjective optimization problem. It is necessary to notice that the model discrepancies are changing whenever new information is found about the ground truth by querying the information sources. The model discrepancy is defined as the difference between the predicted value of the model built with data from an information source and the model built with the available data from ground truth for a specific design space point. Therefore, model discrepancies should be updated regularly. However, querying the ground truth to update its surrogate model is costly. Thus, we need to define a condition for when to query the ground truth. Such a condition can be, for example, when a certain number of updates have been made to available information sources, or also when a specific amount of the total allotted budget is spent. This method allows the decision maker to query a cheap information source more between ground truth queries if it finds the cheap information source is still providing useful knowledge about the ground truth. This is in line with expected intuition regarding the exploitation of cheap information sources, given their nearly negligible cost in comparison to expensive sources and ground truth itself.

Algorithm 1 presents our overall framework. Our procedure to optimize a multiobjective function is established by assuming

Algorithm 1: Multiobjective Bayesian optimization

```

1: construct  $GP_1$  to  $GP_m$  given available data from the ground truth
2: for  $i$  from 1 to  $n$ , do
3:   for  $j$  from 1 to  $m$ , do
4:     construct  $GP_{j,i}$  for objective ( $j$ ) of the information source ( $i$ ) given the data
5:   end for
6: end for
7: fuse models and construct the initial Pareto front
8: while available budget  $> 0$ , do
9:   X-sample set  $\leftarrow$  Latin hypercube sampling
10:  for  $k$  from 1 to  $n$ , do
11:    for  $s$  from 1 to size (X-sample set), do
12:      Y sample  $\leftarrow$  query X samples from  $GP_{1k}$  to  $GP_{mk}$ 
13:      construct temporary GPs by updating  $GP_{1k}$  to  $GP_{mk}$  using sample  $s$ 
14:      updated_fused_values  $\leftarrow$  fuse other models with the updated one
15:      generate test_samples using fused_means and fused_variances
16:      improvement( $s,k$ )  $\leftarrow$  EHVI(test_samples, updated_fused_values, Pareto front)
17:    end for
18:  end for
19:  X = sample to be queried, V = information source (IS) to query from  $\leftarrow$  Max(improvement)
20:  Y = ( $y_1, \dots, y_m$ ) =  $IS_{(V)}(X)$ 
21:  update  $GP_{1,V}$  to  $GP_{m,V}$  using X and Y
22:  fuse models
23:  U  $\leftarrow$  query a randomly generated set of design points from fused model
24:  find nondominated vectors in U to update Pareto front
25:  if requirements to query ground truth is met, then
26:    G  $\leftarrow$  a set of design points with arbitrary size distributed along the Pareto front
27:     $Y_G \leftarrow$  query set G from the ground truth
28:    update  $GP_1$  to  $GP_m$ 
29:    update model discrepancies
30:    fuse models and update the Pareto front
31:  end if
32: end while
33: fuse models to construct fused  $GP_1$  to fused  $GP_m$ 
34:  $\mathcal{U} \leftarrow$  query a randomly generated set of design points from fused models
35:  $\mathcal{S} \leftarrow$  find nondominated vectors in  $\mathcal{U}$  to update the Pareto front
36:  $\mathcal{X} \leftarrow$  the design space corresponding to nondominated set  $\mathcal{S}$ 

```

the function has m objectives, and there are n information sources of differing fidelity available to provide information about the ground truth. Here, the ground truth is the highest-fidelity information source for estimating a quantity of interest. This could, for example, be a real-world experiment on a realized system or a validated high-fidelity simulation model (often with associated uncertainty, which does not incorporate here without loss of generality). We assume that the querying of ground truth is the most expensive means of acquiring information about it. Expense here could mean runtime, cost, or other resources. Although it is possible that ground truth may not be the most expensive information source to query, we do not consider that scenario here. The first step involves the construction of Gaussian processes and creation of the initial Pareto front. This can be established by finding nondominated design points of initial data available from the ground truth. Next, the fusion step takes place, which involves the previously described model reification process. This is followed by the generation of candidate query points, which are tested for EHVI potential given the current set of GPs. The best candidate (query point and information source) is selected and executed. This is followed by another fusion step, given the new information. The budget condition is then checked, which would lead to a ground truth query or a check on whether the budget is exhausted. The term budget refers to the resources available to run a new experiment and limits the total number of evaluations. The budget can be defined, for example, as the computational time in simulations or the total money granted to design experiments in a laboratory. If the budget is exhausted, the process terminates with a final analysis of the estimated Pareto front from the fused GP, which leads to subsequent evaluations of best points from the ground truth. If the budget is not exhausted,

the process resamples candidate points and repeats. A ground truth evaluation is triggered after spending a specified amount of budget on evaluating lower-fidelity information sources. When this occurs, the ground truth is queried, its GP is updated, and then all other GPs are updated (owing to a change in the discrepancies and correlations given new ground truth information). The budget exhaustion condition is then checked, and the process proceeds as previously described from this point. A complete flowchart of this process is provided in Fig. 3.

When the decision to query the ground truth is made based on the budget condition set by expert opinion, a certain number of points N are considered as potential queries. Although choosing larger values of N results in more information gain and higher accuracy to estimate the model discrepancies, it is not necessarily desirable since the ground truth is an expensive-to-evaluate function or experiment. Hence, a tradeoff should be considered in assigning a value to N . For the purposes of the demonstration cases that follow in Sec. IV, we have set $N = 10$ for the test function and $N = 4$ for the OpenAeroStruct demonstration; however, the study of this parameter is a topic of future work. Algorithm 2 presents our ground truth querying strategy.

Algorithm 2: Querying the ground truth

```

1: divide the most updated Pareto front into  $N$  slices
2: construct a smooth GP for each slice, given data points in the slice
3:  $P \leftarrow$  choose the closest point to the GP mean in each slice
4:  $P_Y \leftarrow$  query the design points corresponding to  $P$  from the ground truth

```

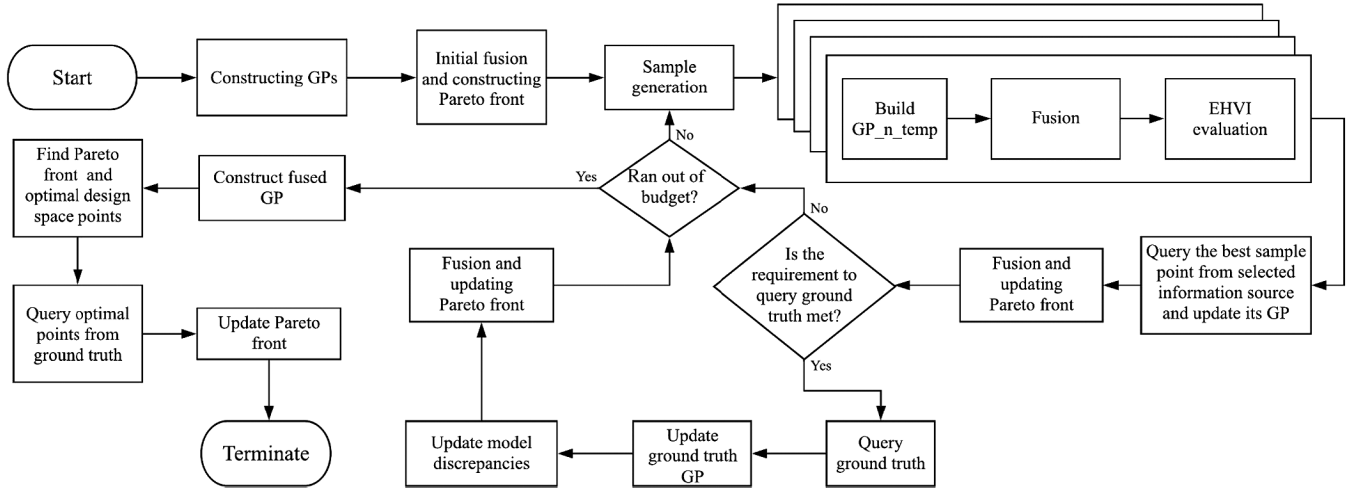


Fig. 3 Procedure flow of the proposed framework. The ground truth query requirement can meet a certain number of iterations or spent budget.

IV. Application and Results

To evaluate the performance of our proposed algorithm, we have applied it on a test function from Ref. [61]. This is referred to as Poloni's two-objective test function, which maps points from a two-dimensional design space to a two-dimensional objective space. A comparison between the optimal Pareto front associated with the problem and the approximated Pareto front is made to show the effectiveness of the algorithm. To apply the concept of a multi-fidelity approach in optimizing the test function, we have constructed two other functions close to the test function by changing the coefficients and constants. The test function itself is considered ground truth. We follow the demonstration of our framework on the test function with its application to an aircraft wing design problem using OpenAeroStruct [62]. We describe the software and the problem for this application in Sec. IV.B.

We note here that our proposed optimization approach is stochastic in nature, and thus involves uncertainty from a few different sources. The result is that the results are also stochastic in nature. The sources of uncertainty include the use of different training sets to build the initial GPs, the random nature of how we generate candidate design points to be tested, and the sample-based nature of the fusion process employed. To account for these uncertainties, we present the results for several different simulations using different initializations and candidate point locations. We show this uncertainty in the form of 95% empirical confidence intervals in the relevant figures.

A. Poloni's Test Function

Poloni's two-objective function is a two-dimensional input test function defined as

$$\text{minimize: } f_1(x_1, x_2) = 1 + (A_1 - B_1(x_1, x_2))^2 + (A_2 - B_2(x_1, x_2))^2$$

$$\text{minimize: } f_2(x_1, x_2) = (x_1 + 3)^2 + (x_2 + 1)^2$$

where

$$-\pi \leq x_1, x_2 \leq \pi$$

$$A_1 = 0.5 \sin(1) - 2 \cos(1) + \sin(2) - 1.5 \cos(2)$$

$$A_2 = 1.5 \sin(1) - \cos(1) + 2 \sin(2) - 0.5 \cos(2)$$

$$B_1(x_1, x_2) = 0.5 \sin(x_1) - 2 \cos(x_1) + \sin(x_2) - 1.5 \cos(x_2)$$

$$B_2(x_1, x_2) = 1.5 \sin(x_1) - \cos(x_1) + 2 \sin(x_2) - 0.5 \cos(x_2)$$

Figure 4 shows the optimal versus final estimation of the Pareto front and the hypervolume. Since there is no closed-form solution for Poloni's test problem, the optimal Pareto front and its hypervolume are found by exhaustive search. The optimal Pareto front here matches those reported in Refs. [63,64]. The estimated Pareto front is found using the knowledge of lower-fidelity models about the ground truth. Looking at the hypervolume values, it is showing the hypervolume is increasing as a result of improved estimation of the Pareto front, admitting that the budget is spent effectively. The budget here is set

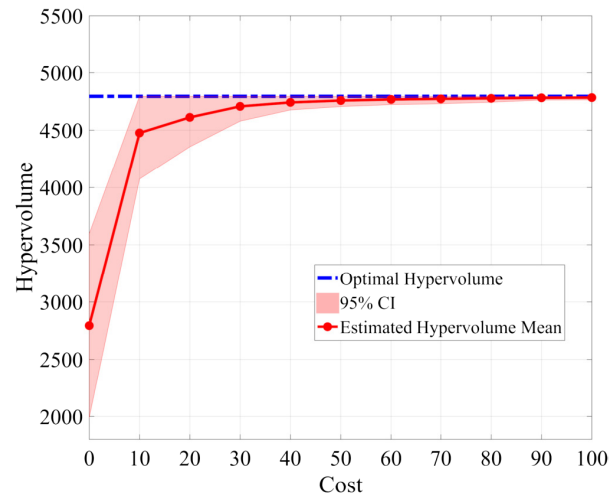
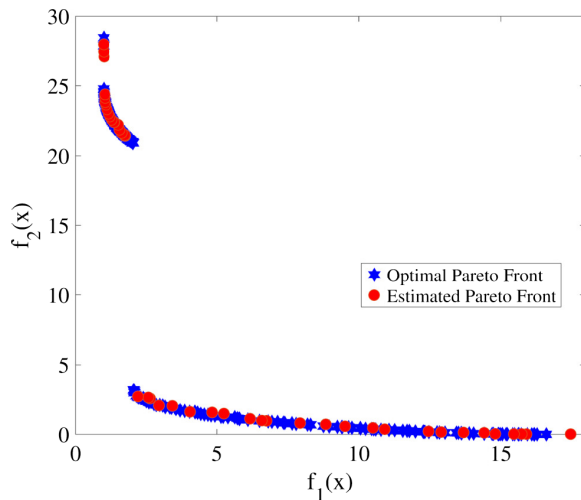


Fig. 4 The optimal and estimated Pareto fronts and hypervolumes averaged over 30 replications with 95% CI (Confidence Interval). The reference point is (70, 70).

to 100 to limit the total number of information source evaluations and is used to inform the system of when to trigger the ground truth evaluation if needed. The costs of querying the low-fidelity and medium-fidelity models are set to one and two units of cost, respectively. These values were chosen to ensure adequate use of the information sources. In a practical setting, these values would be computed via resource usage (such as actual runtime). The hypervolume is initially computed using randomly generated data points with a training set size of 20 for every simulation.

A significant improvement is achieved before spending the first 10% of the budget. This was achieved largely through the exploitation of the cheaper lower-fidelity model. Note that the ground truth is queried each time 10 units of cost is spent to update model discrepancy and hypervolume on a regular basis. We present a similar result and the information source query history for the OpenAeroStruct demonstration in Sec. IV.B. The diminishing returns in hypervolume are expected, given that finding new nondominated points becomes more difficult as more points are found.

Many previous works have used Poloni's test function to measure the performance of their proposed approaches. For example, Ref. [64] proposes a method using differential evolution. The results show they have found the optimal Pareto front after 600 function evaluations. Also, in Ref. [63], it is reported that 2500 function evaluations are used to cover the estimated Pareto front close to the optimal Pareto front found using an exhaustive search. In Ref. [65], a genetic algorithm approach is taken using populations of more than 500 for 250 generations to obtain the optimal Pareto front. Our proposed method is generally outperforming each of these prior approaches. Here, our method uses a total of 100 function evaluations with an additional set of less than 100 evaluations from lower-fidelity information sources (which are considered much cheaper than the ground truth; although, in general, this would be problem specific). This difference in the necessary number of evaluations emphasizes the efficiency gains achieved by our method while maintaining high-quality Pareto front estimates.

B. OpenAeroStruct Demonstration

OpenAeroStruct is an open-source software developed in NASA's OpenMDAO framework [66], which can be used for fast tightly coupled aerostructural design optimization. The framework implements the coupled adjoint method to compute the aerostructural derivatives used for efficient gradient-based optimization. As noted in Ref. [62], OpenAeroStruct combines a vortex lattice method (VLM) and a one-dimensional finite element analysis using six-degree-of-freedom three-dimensional spatial beam elements to model lifting surfaces [62,67]. A common aerostructural single objective optimization problem is the fuel burn minimization problem using the Breguet range equation. Structural mass minimization of the wing is also frequently considered, and thus is used as a second objective in demonstrating our proposed multi-information source multiobjective optimization framework.

The OpenAeroStruct application, as described in Ref. [62], uses the Breguet range equation to compute the fuel burn as a function of structural weight and aerodynamic performance. Design variables consist of twist distributions, spar thickness distributions, and planform variables such as skin thickness, thickness over cord ratio, and angle of attack. The first four variables are four-dimensional because four control surfaces were considered for the wing. Hence, the problem has a 17-dimensional design space. Constraints in the standard problem ensure lift equals weight, and that structural failure does not occur.

The mesh in OpenAeroStruct is defined by the number of the spanwise and chordwise points as shown in Fig. 5 [68]. The fidelity of each model depends on the number of points used to define the lifting surface. A model with a finer mesh is considered to have higher fidelity compared to a model with a coarser mesh. We use three different mesh resolutions in this demonstration to serve as three different multifidelity information sources.

The different mesh sizes and costs are shown in Table 1, where Num_y is the number of spanwise points and Num_x is the number of

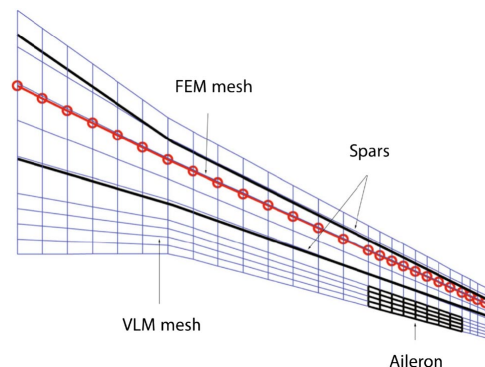


Fig. 5 A wing with the aerodynamic and structural Meshes [68].

Table 1 Mesh sizes and costs for different fidelity models

Fidelity level	Num_y	Num_x	Cost, s
Low	15	3	1.9
Medium	35	11	45.1
High (ground truth)	55	19	283.9

chordwise points. The low-fidelity mesh was chosen to ensure meaningful results, and the high-fidelity mesh was chosen through a mesh refinement analysis that ensured adequate convergence. In Fig. 6, the three different meshes are shown. The cost of evaluating each model is based on the computational runtime of a single query.

We applied our approach of multi-information source multiobjective optimization on this two-objective OpenAeroStruct problem with three information sources taking the highest-fidelity one as the ground truth. As mentioned earlier, our objective here is to minimize both fuel burn and wing mass by controlling 17 design variables. We assumed a budget of 2000 s of computational time on lower-fidelity information sources for this demonstration. The results are shown in Fig. 7, where random points are shown in blue to show the objective space (these are not part of the algorithm and are for visualization only), the green points are those points selected by our approach with the fused GP, and the red points are the nondominated green points that have been evaluated with the ground truth (that is, the final step of our algorithm). The figure reveals that our approach has done well in identifying the nondominated region in the objective space for this 17-dimensional problem. Note that the Pareto front for this problem is not spread along a large region of the objective space. This is expected based on the shape of the objective space as shown by the randomly queried point.

In Fig. 8, the hypervolume is updated during the optimization process each time 100 s of computational time are spent on evaluating the lower-fidelity information sources. This choice results in regular updating of the discrepancy terms. A careful study of the optimal allocation to lower-fidelity information sources and ground truth estimates is a topic of future work. We note here that this is not always a clear resource tradeoff since, often, computation is measured in runtime and physical experiments may need to be measured in monetary units as well as time. The objectives are normalized using the upper limit known for each objective. Since both objectives have large values, differences in hypervolumes might not be sensible and normalization helps to see the changes clearly. The reference point should be dominated by all points in the objective space, and here it is fixed as (1.1, 1.1). The most significant changes in the hypervolume quantity are made when spending the first 10% of the budget (as was seen in the previous test case as well). Beyond that, there are improvements in the hypervolume, but the returns are diminishing as expected.

Figure 9 reveals the cumulative sum of queries from any information source plotted against the overall iteration, where the iteration is defined as it was previously. As depicted in the figure, initially, the low-fidelity information source has been queried much more than the medium-fidelity information source. After some number of iterations (about 120 in this case), the value of the low-fidelity source has

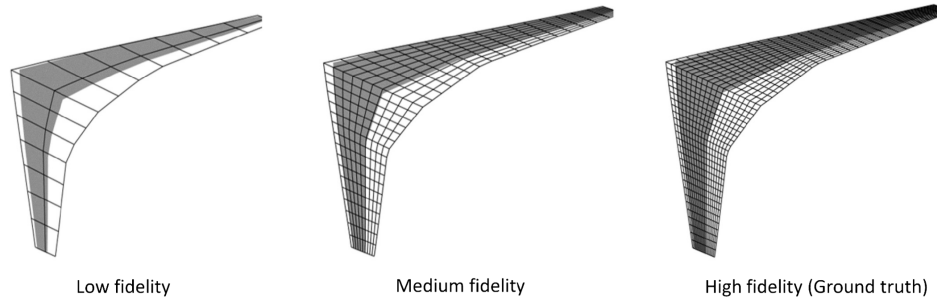


Fig. 6 Illustration of the meshed wing with different fidelity models. The number of meshes in each model is presented in Table 1.

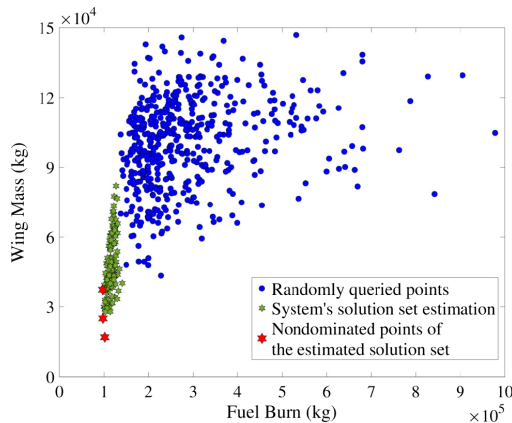


Fig. 7 Final estimation of the Pareto front from the fused model nondominated designs shown in red and green stars, respectively.

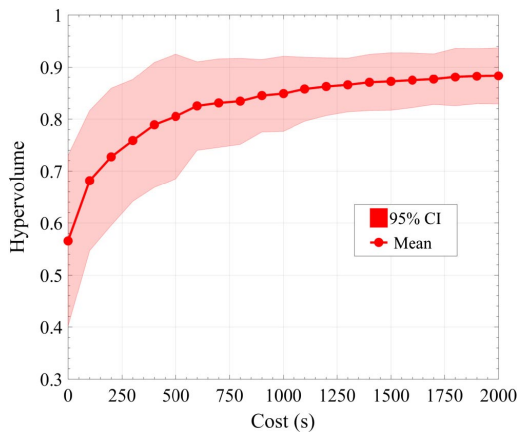


Fig. 8 Estimated hypervolume with respect to the cost averaged over 30 independent simulations with different starting points.

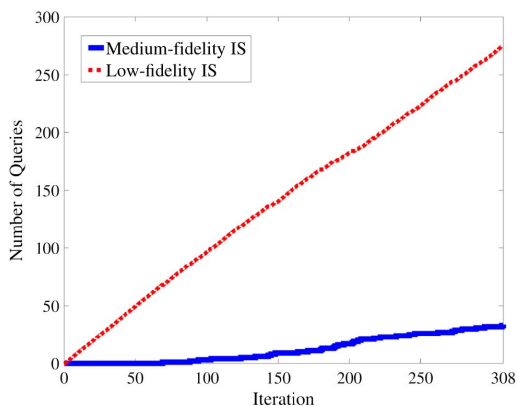


Fig. 9 The cumulative sum of queries from each information source. We query one information source at each iteration.

diminished enough that some queries to the higher-fidelity source are now necessary. This continues until the budget is exhausted. We see in this case that the low-fidelity source continues to be queried as well. This is due to the fact that as higher-fidelity information is obtained, the correlations between the low- and higher-fidelity sources are updated, which results in a possible renewed value in lower-fidelity information. This was the case here.

For the aforementioned analysis, a possible question is whether there was only one information source available to estimate the quantities of interest in this design problem. To address this question, another experiment is designed to compare the results between the multifidelity approach and the single fidelity approach. We have considered the single fidelity optimization task using the medium-fidelity information source. The GP built for the information source is taken as the predictor model, and there is no fusion of information and a fused model in the single fidelity optimization case. Figure 10 shows the hypervolume averaged over 30 simulations for each case in a normalized objective space with different starting points. The multifidelity approach outperforms the single fidelity approach since it has access to more information about the ground truth. This result shows the low-fidelity information source contribution to provide useful information about the quantity of interest in the multifidelity configuration.

The next step is to compare the effectiveness of different algorithms proposed to do multiobjective optimization task. Here, we have compared the results of NSGA-II [69–71], ParEGO [72,73], and EHV1 methods in optimizing the OpenAeroStruct design problem. We will not present the whole algorithms here but interested readers can find details regarding these approaches and implementations in the aforementioned references.

In general, the ParEGO algorithm is an extended version of the efficient global optimization algorithm initially introduced in Ref. [53] for global optimization of single objective expensive black-box functions. The EGO algorithm is a surrogate-based method and searches for new solutions using the expected improvement criterion. At each iteration, a set of random solutions is generated

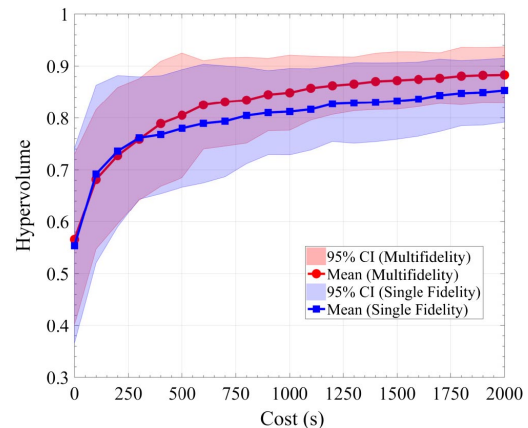


Fig. 10 Comparing the estimated hypervolume in single fidelity and multifidelity approaches averaged over 30 independent replications of simulations.

in a Latin hypercube or any other space filling design, and the solution that maximizes the expected improvement will be queried from the expensive function to update the surrogate model. To extend the method for optimizing multiobjective functions, one approach is to combine all objectives into a single objective using parameterized scalarizing weight vectors [73]. In NSGA-II, a nondominated sorting is done over the available data or in terms of a genetic algorithm point of view, population, and they are given a rank according to their nondomination level. New populations (solutions) are generated according to their given front rank trying to find new nondominated solutions. Readers are referred to Refs. [69,70] for more details.

NSGA-II and ParEGO are not set up to take advantage of multiple information sources, and they are employed to optimize one function. However, as these algorithms might need to query a large number of points from the function directly, optimizing the ground truth does not make sense with respect to the cost of each query. The goal is to see how they perform if the same number of resources is available for all methods. Consequently, the optimization is done over the medium-fidelity information source to allow a reasonable number of queries from the function.

The hypervolume estimations in Fig. 11 show the improvement achieved by our EHVI-based approach. Note that the starting point for every simulation is different; thus, we have included the uncertainty region even for the initial hypervolume. The EHVI approach has the advantage of coming up with a good estimation of the optimal Pareto front, meaning a larger hypervolume much faster than the other approaches. Therefore, in highly budget-constrained experiments, it finds solutions closer to the optimal Pareto front. Although it is seen that the estimated hypervolumes might converge when more resources are available, since they have enough budget to search the space, the EHVI is still suggesting better solutions to the problem. Also, the NSGA-II approach is building up the Pareto front gradually and will likely reach the EHVI and ParEGO estimations of the hypervolume at higher costs.

In Fig. 12, we demonstrate the results of using our multi-information source approach versus a single information source approach using EHVI as the acquisition function, as well as the ParEGO and NSGA-II methods. Here, a representative result from the 30 simulations is used to show the results of the different algorithms. We see from the figure that the multi-information source approach dominates the other approaches. In some cases, it is possible that ParEGO and NSGA-II provide a better solution (say, in a different choice among the 30 simulations); however, as shown in Fig. 11, our method is performing better on average.

V. Conclusions

In this paper, a methodology to optimize expensive multiobjective functions is presented. The methodology seeks to exploit all available information sources for efficiently identifying nondominated points

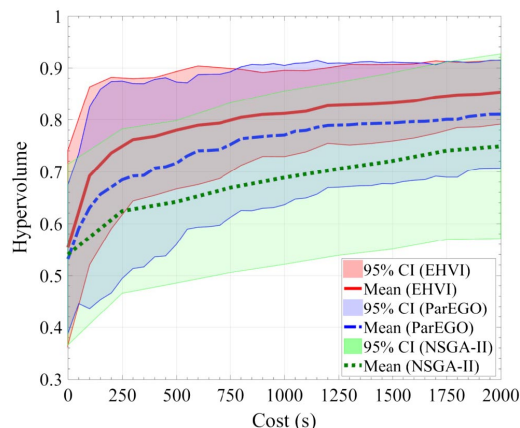


Fig. 11 Our proposed approach using EHVI outperformed other methods obtaining larger hypervolume. Values show the averaged hypervolume over 30 independent simulations.

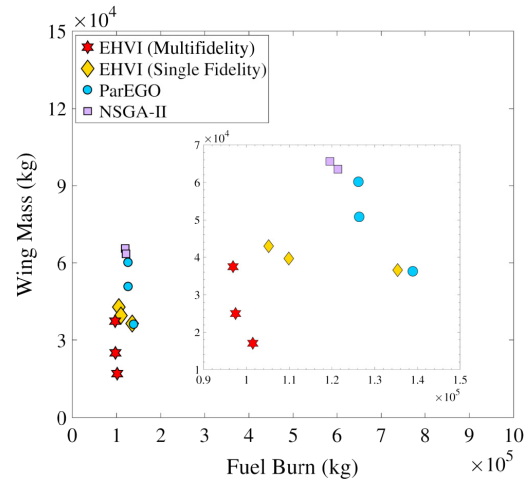


Fig. 12 Pareto fronts obtained using different methods.

in the objective space as a means of estimating the true Pareto front. The approach was based on the fast evaluation of the expected hypervolume improvement through the use of temporarily updated Gaussian process surrogate models of each information source. The process also incorporates model reification to fuse new information rigorously as it becomes available through proper accounting for correlation between the sources. The conclusion of this study is that multi-information source Bayesian optimization approaches to directing efficient querying when the budget is constrained can be effective ways of estimating the Pareto front of a multiobjective problem. In particular, the ability to rapidly query lower-fidelity sources while accounting for their correlation with higher-fidelity sources and ground truth has enabled efficient (less than 10% of the budget for the problems studied here) identification of promising regions for nondominated point searching. Then, improvement over the Pareto front estimation is shown when more information sources are available. Any information source can provide useful information about the quantity of interest that is not accessible from other information sources. Finally, the performance of the proposed approach is compared to two other well-known multiobjective optimization approaches, called ParEGO and NSGA-II. The results demonstrate the effectiveness of the current approach in budget-constrained situations.

Acknowledgments

This work was supported by the U.S. Air Force Office of Scientific Research Multidisciplinary Research Program of the University Research Initiative on multi-information sources of multiphysics systems under award number FA9550-15-1-0038 with Program Manager Fariba Fahroo and by the National Science Foundation under grant no. CMMI-1663130. The opinions expressed in this paper are of the authors and do not necessarily reflect the views of the National Science Foundation.

References

- [1] Ghoreishi, S. F., Molkeri, A., Srivastava, A., Arroyave, R., and Allaire, D., "Multi-Information Source Fusion and Optimization to Realize ICME: Application to Dual-Phase Materials," *Journal of Mechanical Design*, Vol. 140, No. 11, 2018, Paper 111409. <https://doi.org/10.1115/1.4041034>
- [2] Ariyarat, A., and Kanazaki, M., "Multi-Fidelity Multi-Objective Efficient Global Optimization Applied to Airfoil Design Problems," *Applied Sciences*, Vol. 7, No. 12, 2017, Paper 1318. <https://doi.org/10.3390/app7121318>
- [3] Amrit, A., and Leifsson, L., "Applications of Surrogate-Assisted and Multi-Fidelity Multi-Objective Optimization Algorithms to Simulation-Based Aerodynamic Design," *Engineering Computations*, Vol. 37, No. 2, 2019, pp. 430–457. <https://doi.org/10.1108/EC-12-2018-0553>

- [4] Leifsson, L., Koziel, S., and Tesfahunegn, Y. A., "Multiobjective Aerodynamic Optimization by Variable-Fidelity Models and Response Surface Surrogates," *AIAA Journal*, Vol. 54, No. 2, 2016, pp. 531–541. <https://doi.org/10.2514/1.J054128>
- [5] Coello, C. A. C., and Lamont, G. B., *Applications of Multi-Objective Evolutionary Algorithms*, Vol. 1, World Scientific, Singapore, 2004. https://doi.org/10.1142/9789812567796_0001
- [6] Fonseca, C. M. M. D., "Multiobjective Genetic Algorithms with Application to Control Engineering Problems," Ph.D. Thesis, Univ. of Sheffield, Sheffield, England, U.K., 1995.
- [7] Amrit, A., Leifsson, L., and Koziel, S., "Design Strategies for Multi-Objective Optimization of Aerodynamic Surfaces," *Engineering Computations*, Vol. 34, No. 5, 2017, pp. 1724–1753. <https://doi.org/10.1108/EC-07-2016-0239>
- [8] Allaire, D., and Willcox, K., "Fusing Information from Multifidelity Computer Models of Physical Systems," *2012 15th International Conference on Information Fusion*, Inst. of Electrical and Electronics Engineers, New York, 2012, pp. 2458–2465.
- [9] Winkler, R., "Combining Probability Distributions from Dependent Information Sources," *Management Science*, Vol. 27, No. 4, 1981, pp. 479–488. <https://doi.org/10.1287/mnsc.27.4.479>
- [10] Winkler, R. L., "Combining Probability Distributions from Dependent Information Sources," *Management Science*, Vol. 27, No. 4, 1981, pp. 479–488. <https://doi.org/10.1287/mnsc.27.4.479>
- [11] Rasmussen, C. E., and Williams, C. K. I., *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*, MIT Press, Cambridge, MA, 2005.
- [12] Allaire, D., and Willcox, K., "Surrogate Modeling for Uncertainty Assessment with Application to Aviation Environmental System Models," *AIAA Journal*, Vol. 48, No. 8, 2010, pp. 1791–1803. <https://doi.org/10.2514/1.J050247>
- [13] Ghoreishi, S. F., and Allaire, D. L., "Gaussian Process Regression for Bayesian Fusion of Multi-Fidelity Information Sources," *2018 Multidisciplinary Analysis and Optimization Conference*, AIAA Paper 2018-4176, 2018. <https://doi.org/10.2514/6.2018-4176>
- [14] Zhao, G., Arroyave, R., and Qian, X., "Fast Exact Computation of Expected HyperVolume Improvement," 2018.
- [15] Alexandrov, N. M., Lewis, R. M., Gumbert, C. R., Green, L. L., and Newman, P. A., "Approximation and Model Management in Aerodynamic Optimization with Variable-Fidelity Models," *AIAA Journal*, Vol. 38, No. 6, 2001, pp. 1093–1101. <https://doi.org/10.2514/2.2877>
- [16] Booker, A. J., Dennis, J. E., Frank, P. D., Serafini, D. B., Torczon, V., and Trosset, M. W., "A Rigorous Framework for Optimization of Expensive Functions by Surrogates," *Structural Optimization*, Vol. 17, No. 1, 1999, pp. 1–13. <https://doi.org/10.1007/BF01197708>
- [17] Lam, R., Allaire, D. L., and Willcox, K. E., "Multifidelity Optimization Using Statistical Surrogate Modeling for Non-Hierarchical Information Sources," *56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA Paper 2015-0143, 2015. <https://doi.org/10.2514/6.2015-0143>
- [18] Allaire, D., and Willcox, K., "A Mathematical and Computational Framework for Multifidelity Design and Analysis with Computer Models," *International Journal for Uncertainty Quantification*, Vol. 4, No. 1, 2014, pp. 1–20. <https://doi.org/10.1615/Int.J.UncertaintyQuantification.2013004121>
- [19] Archetti, F., and Candelieri, A., *Bayesian Optimization and Data Science*, Springer, New York, 2019. <https://doi.org/10.1007/978-3-030-24494-1>
- [20] Ghoreishi, S. F., and Allaire, D., "Multi-Information Source Constrained Bayesian Optimization," *Structural and Multidisciplinary Optimization*, Vol. 59, No. 3, 2019, pp. 977–991. <https://doi.org/10.1007/s00158-018-2115-z>
- [21] Ghoreishi, S. F., Thomison, W. D., and Allaire, D., "Sequential Information-Theoretic and Reification-Based Approach for Querying Multi-Information Sources," *Journal of Aerospace Information Systems*, Vol. 16, No. 12, 2019, pp. 575–587. <https://doi.org/10.2514/1.1010753>
- [22] Thomison, W. D., and Allaire, D. L., "A Model Reification Approach to Fusing Information from Multifidelity Information Sources," *19th AIAA Non-Deterministic Approaches Conference*, AIAA Paper 2017-1949, 2017. <https://doi.org/10.2514/6.2017-1949>
- [23] Balabanov, V., Haftka, R., Grossman, B., Mason, W., and Watson, L., "Multifidelity Response Surface Model for HSCT Wing Bending Material Weight," *7th AIAA/USAF/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization*, AIAA Paper 1998-4804, 1998. <https://doi.org/10.2514/6.1998-4804>
- [24] Balabanov, V., and Venter, G., "Multi-Fidelity Optimization with High-Fidelity Analysis and Low-Fidelity Gradients," *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, AIAA Paper 2004-4459, 2004. <https://doi.org/10.2514/6.2004-4459>
- [25] Choi, S., Alonso, J. J., and Kroo, I. M., "Two-Level Multifidelity Design Optimization Studies for Supersonic Jets," *Journal of Aircraft*, Vol. 46, No. 3, 2009, pp. 776–790. <https://doi.org/10.2514/1.34362>
- [26] Eldred, M., Giunta, A., and Collis, S., "Second-Order Corrections for Surrogate-Based Optimization with Model Hierarchies," *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, AIAA Paper 2004-4457, 2004. <https://doi.org/10.2514/6.2004-4457>
- [27] March, A., and Willcox, K., "Convergent Multifidelity Optimization Using Bayesian Model Calibration," *Structural and Multidisciplinary Optimization*, Vol. 46, No. 1, 2012, pp. 93–109. <https://doi.org/10.1007/s00158-011-0749-1>
- [28] March, A., and Willcox, K., "Provably Convergent Multifidelity Optimization Algorithm Not Requiring High-Fidelity Derivatives," *AIAA Journal*, Vol. 50, No. 5, 2012, pp. 1079–1089. <https://doi.org/10.2514/1.J051125>
- [29] Clyde, M., "Model Averaging," *Subjective and Objective Bayesian Statistics*, 2nd ed., Wiley-Interscience, New York, 2003, Chap. 13. <https://doi.org/10.1002/9780470317105.ch13>
- [30] Clyde, M., and George, E., "Model Uncertainty," *Statistical Science*, Vol. 19, No. 1, 2004, pp. 81–94. <https://doi.org/10.1214/088342304000000035>
- [31] Draper, D., "Assessment and Propagation of Model Uncertainty," *Journal of the Royal Statistical Society Series B*, Vol. 57, No. 1, 1995, pp. 45–97. <https://doi.org/10.1111/j.2517-6161.1995.tb02015.x>
- [32] Hoeting, J., Madigan, D., Raftery, A., and Volinsky, C., "Bayesian Model Averaging: A Tutorial," *Statistical Science*, Vol. 14, No. 4, 1999, pp. 382–417. <https://doi.org/10.1214/ss/1009212519>
- [33] Leamer, E., *Specification Searches: Ad Hoc Inference with Nonexperimental Data*, Wiley, New York, 1978. <https://doi.org/10.2307/1057568>
- [34] Madigan, D., and Raftery, A., "Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam's Window," *American Statistical Association*, Vol. 89, No. 428, 1994, pp. 1535–1546. <https://doi.org/10.1080/01621459.1994.10476894>
- [35] Mosleh, A., and Apostolakis, G., "The Assessment of Probability Distributions from Expert Opinions with an Application to Seismic Fragility Curves," *Risk Analysis*, Vol. 6, No. 4, 1986, pp. 447–461. <https://doi.org/10.1111/j.1539-6924.1986.tb00957.x>
- [36] Reinert, J., and Apostolakis, G., "Including Model Uncertainty in Risk-Informed Decision Making," *Annals of Nuclear Energy*, Vol. 33, No. 4, 2006, pp. 354–369. <https://doi.org/10.1016/j.anucene.2005.11.010>
- [37] Riley, M., and Grandhi, R., "Quantification of Modeling Uncertainty in Aeroelastic Analyses," *Journal of Aircraft*, Vol. 48, No. 3, 2011, pp. 866–873. <https://doi.org/10.2514/1.C031059>
- [38] Zio, E., and Apostolakis, G., "Two Methods for the Structured Assessment of Model Uncertainty by Experts in Performance Assessments of Radioactive Waste Repositories," *Reliability Engineering and System Safety*, Vol. 54, Nos. 2–3, 1996, pp. 225–241. [https://doi.org/10.1016/S0951-8320\(96\)00078-6](https://doi.org/10.1016/S0951-8320(96)00078-6)
- [39] Julier, S., and Uhlmann, J., "A Non-Divergent Estimation Algorithm in the Presence of Unknown Correlations," *Proceedings of the American Control Conference*, IEEE, New York, 1997, pp. 2369–2373. <https://doi.org/10.1109/ACC.1997.609105>
- [40] Geisser, S., "A Bayes Approach for Combining Correlated Estimates," *Journal of the American Statistical Association*, Vol. 60, No. 310, 1965, pp. 602–607. <https://doi.org/10.1080/01621459.1965.10480816>
- [41] Morris, P., "Combining Expert Judgments: A Bayesian Approach," *Management Science*, Vol. 23, No. 7, 1977, pp. 679–693. <https://doi.org/10.1287/mnsc.23.7.679>
- [42] Marler, R. T., and Arora, J. S., "The Weighted Sum Method for Multi-Objective Optimization: New Insights," *Structural and Multidisciplinary Optimization*, Vol. 41, No. 6, 2010, pp. 853–862. <https://doi.org/10.1007/s00158-009-0460-7>

- [43] Kim, I. Y., and de Weck, O. L., "Adaptive Weighted-Sum Method for Bi-Objective Optimization: Pareto Front Generation," *Structural and Multidisciplinary Optimization*, Vol. 29, No. 2, 2005, pp. 149–158. <https://doi.org/10.1007/s00158-004-0465-1>
- [44] Das, I., and Dennis, J. E., "Normal-Boundary Intersection: A New Method for Generating the Pareto Surface in Nonlinear Multicriteria Optimization Problems," *SIAM Journal on Optimization*, Vol. 8, No. 3, 1998, pp. 631–657. <https://doi.org/10.1137/S1052623496307510>
- [45] Beume, N., "S-Metric Calculation by Considering Dominated Hypervolume as Klee's Measure Problem," *Evolutionary Computation*, Vol. 17, No. 4, 2009, pp. 477–492. <https://doi.org/10.1162/evco.2009.17.4.17402>
- [46] Bradstreet, L., While, L., and Barone, L., "A Fast Many-Objective Hypervolume Algorithm Using Iterated Incremental Calculations," *IEEE Congress on Evolutionary Computation*, Inst. of Electrical and Electronics Engineers, New York, 2010, pp. 1–8. <https://doi.org/10.1109/CEC.2010.5586344>
- [47] Emmerich, M. T., Deutz, A. H., and Klöckner, J. W., "Hypervolume-Based Expected Improvement: Monotonicity Properties and Exact Computation," *2011 IEEE Congress of Evolutionary Computation (CEC)*, Inst. of Electrical and Electronics Engineers, New York, 2011, pp. 2147–2154. <https://doi.org/10.2514/6.2015-0143>
- [48] Fonseca, C. M., Paquete, L., and López-Ibáñez, M., "An Improved Dimension-Sweep Algorithm for the Hypervolume Indicator," *2006 IEEE International Conference on Evolutionary Computation*, Inst. of Electrical and Electronics Engineers, New York, 2006, pp. 1157–1163. <https://doi.org/10.1109/CEC.2006.1688440>
- [49] Russo, L. M., and Francisco, A. P., "Quick Hypervolume," *IEEE Transactions on Evolutionary Computation*, Vol. 18, No. 4, 2013, pp. 481–502. <https://doi.org/10.1109/TEVC.2013.2281525>
- [50] Yang, Q., and Ding, S., "Novel Algorithm to Calculate Hypervolume Indicator of Pareto Approximation Set," *International Conference on Intelligent Computing*, Springer, New York, 2007, pp. 235–244. https://doi.org/10.1007/978-3-540-74282-1_27
- [51] Zitzler, E., and Thiele, L., "Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach," *IEEE Transactions on Evolutionary Computation*, Vol. 3, No. 4, 1999, pp. 257–271. <https://doi.org/10.1109/4235.797969>
- [52] Moćkus, J., "On Bayesian Methods for Seeking the Extremum," *Optimization Techniques IFIP Technical Conference*, Springer, New York, 1975, pp. 400–404. https://doi.org/10.1007/978-3-662-38527-2_55
- [53] Jones, D. R., Schonlau, M., and Welch, W. J., "Efficient Global Optimization of Expensive Black-Box Functions," *Journal of Global Optimization*, Vol. 13, No. 4, 1998, pp. 455–492. <https://doi.org/10.1023/A:1008306431147>
- [54] Frazier, P. I., "A Tutorial on Bayesian Optimization," Preprint, submitted 8 July 2018, <https://arxiv.org/abs/1807.02811>.
- [55] Snoek, J., Larochelle, H., and Adams, R. P., "Practical Bayesian Optimization of Machine Learning Algorithms," *Advances in Neural Information Processing Systems*, 2012, pp. 2951–2959.
- [56] Couckuyt, I., Deschrijver, D., and Dhaene, T., "Fast Calculation of Multiobjective Probability of Improvement and Expected Improvement Criteria for Pareto Optimization," *Journal of Global Optimization*, Vol. 60, No. 3, 2014, pp. 575–594. <https://doi.org/10.1007/s10898-013-0118-2>
- [57] Feliot, P., Bect, J., and Vazquez, E., "A Bayesian Approach to Constrained Single and Multi-Objective Optimization," *Journal of Global Optimization*, Vol. 67, Nos. 1–2, 2017, pp. 97–133. <https://doi.org/10.1007/s10898-016-0427-3>
- [58] While, L., Bradstreet, L., and Barone, L., "A Fast Way of Calculating Exact Hypervolumes," *IEEE Transactions on Evolutionary Computation*, Vol. 16, No. 1, 2011, pp. 86–95. <https://doi.org/10.1109/TEVC.2010.2077298>
- [59] Jaszkiewicz, A., "Improved Quick Hypervolume Algorithm," *Computers and Operations Research*, Vol. 90, Feb. 2018, pp. 72–83. <https://doi.org/10.1016/j.cor.2017.09.016>
- [60] Khatamsaz, D., Peddareddygar, L., Friedman, S., and Allaire, D. L., "Efficient Multi-Information Source Multiobjective Bayesian Optimization," *AIAA SciTech 2020 Forum*, AIAA Paper 2020-2127, 2020. <https://doi.org/10.2514/6.2020-2127>
- [61] Deb, K., Thiele, L., Laumanns, M., and Zitzler, E., "Scalable Multi-Objective Optimization Test Problems," *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, Vol. 1, Inst. of Electrical and Electronics Engineers, New York, 2002, pp. 825–830. <https://doi.org/10.1109/CEC.2002.1007032>
- [62] Jasa, J. P., Hwang, J. T., and Martins, J. R., "Open-Source Coupled Aerostructural Optimization Using Python," *Structural and Multidisciplinary Optimization*, Vol. 57, No. 4, 2018, pp. 1815–1827. <https://doi.org/10.1007/s00158-018-1912-8>
- [63] Poloni, C., Giurgevich, A., Onesti, L., and Pediroda, V., "Hybridization of a Multi-Objective Genetic Algorithm, a Neural Network and a Classical Optimizer for a Complex Design Problem in Fluid Dynamics," *Computer Methods in Applied Mechanics and Engineering*, Vol. 186, Nos. 2–4, 2000, pp. 403–420. [https://doi.org/10.1016/S0045-7825\(99\)00394-1](https://doi.org/10.1016/S0045-7825(99)00394-1)
- [64] Riley, M. J., Peachey, T., Abramson, D., and Jenkins, K. W., "Multi-Objective Engineering Shape Optimization Using Differential Evolution Interfaced to the Nimrod/O Tool," *IOP Conference Series: Materials Science and Engineering*, Vol. 10, July 2010, Paper 012189. <https://doi.org/10.1088/1757-899X/10/1/012189>
- [65] Kelner, V., and Léonard, O., "Application of Genetic Algorithms to Lubrication Pump Stacking Design," *Journal of Computational and Applied Mathematics*, Vol. 168, Nos. 1–2, 2004, pp. 255–265. <https://doi.org/10.1016/j.cam.2003.05.022>
- [66] Gray, J. S., Hwang, J. T., Martins, J. R. R. A., Moore, K. T., and Naylor, B. A., "OpenMDAO: An Open-Source Framework for Multidisciplinary Design, Analysis, and Optimization," *Structural and Multidisciplinary Optimization*, Vol. 59, No. 4, 2019, pp. 1075–1104. <https://doi.org/10.1007/s00158-019-02211-z>
- [67] Chauhan, S. S., and Martins, J. R., "Low-Fidelity Aerostructural Optimization of Aircraft Wings with a Simplified Wingbox Model Using OpenAeroStruct," *International Conference on Engineering Optimization*, Springer, New York, 2018, pp. 418–431. https://doi.org/10.1007/978-3-319-97773-7_38
- [68] Elham, A., and van Tooren, M. J., "Coupled Adjoint Aerostructural Wing Optimization Using Quasi-Three-Dimensional Aerodynamic Analysis," *Structural and Multidisciplinary Optimization*, Vol. 54, No. 4, 2016, pp. 889–906. <https://doi.org/10.1007/s00158-016-1447-9>
- [69] Deb, K., Agrawal, S., Pratap, A., and Meyarivan, T., "A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II," *International Conference on Parallel Problem Solving from Nature*, Springer, New York, 2000, pp. 849–858. https://doi.org/10.1007/3-540-45356-3_83
- [70] Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T., "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 2, 2002, pp. 182–197. <https://doi.org/10.1109/4235.996017>
- [71] Baskar, S., Tamilselvi, S., and Varshini, P. R., "MATLAB Code for Constrained NSGA II," *MATLAB Central File Exchange* (online database), 2020, <https://www.mathworks.com/matlabcentral/fileexchange/49806-matlab-code-for-constrained-nsga-ii-dr-s-baskar-s-tamilselvi-and-p-r-varshini> [retrieved 8 Feb. 2020].
- [72] Davins-Valldaura, J., Moussaoui, S., Pita-Gil, G., and Plestan, F., "Par-EGO Extensions for Multi-Objective Optimization of Expensive Evaluation Functions," *Journal of Global Optimization*, Vol. 67, Nos. 1–2, 2017, pp. 79–96. <https://doi.org/10.1007/s10898-016-0419-3>
- [73] Knowles, J., "ParEGO: A Hybrid Algorithm with On-Line Landscape Approximation for Expensive Multiobjective Optimization Problems," *IEEE Transactions on Evolutionary Computation*, Vol. 10, No. 1, 2006, pp. 50–66. <https://doi.org/10.1109/TEVC.2005.851274>

C. Pettit
Associate Editor